Probability Theory and Stochastic Modelling 95

Shige Peng

Nonlinear Expectations and Stochastic Calculus under Uncertainty

with Robust CLT and G-Brownian Motion

Springer

Probability Theory and Stochastic Modelling

Volume 95

The Probability Theory and Stochastic Modelling series is a merger and continuation of Springer's two well established series Stochastic Modelling and Applied Probability and Probability and Its Applications. It publishes research monographs that make a significant contribution to probability theory or an applications domain in which advanced probability methods are fundamental. Books in this series are expected to follow rigorous mathematical standards, while also displaying the expository quality necessary to make them useful and accessible to advanced students as well as researchers. The series covers all aspects of modern probability theory including

- Gaussian processes
- Markov processes
- Random Fields, point processes and random sets
- Random matrices
- Statistical mechanics and random media
- Stochastic analysis

as well as applications that include (but are not restricted to):

- Branching processes and other models of population growth
- Communications and processing networks
- Computational methods in probability and stochastic processes, including simulation
- Genetics and other stochastic models in biology and the life sciences
- Information theory, signal processing, and image synthesis
- Mathematical economics and finance
- Statistical methods (e.g. empirical processes, MCMC)
- Statistics for stochastic processes
- Stochastic control
- Stochastic models in operations research and stochastic optimization
- Stochastic models in the physical sciences.

More information about this series at http://www.springer.com/series/13205

Shige Peng

Nonlinear Expectations and Stochastic Calculus under Uncertainty

with Robust CLT and G-Brownian Motion

Springer

Shige Peng
Institute of Mathematics
Shandong University
Jinan, Shandong, China

ISSN 2199-3130 ISSN 2199-3149 (electronic)
Probability Theory and Stochastic Modelling
ISBN 978-3-662-59902-0 ISBN 978-3-662-59903-7 (eBook)
https://doi.org/10.1007/978-3-662-59903-7

Mathematics Subject Classification (2010): 60-xx, 39-xx, 62-xx

This Springer imprint is published by the registered company Springer-Verlag GmbH, DE part of Springer Nature.
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Preface

In this book, we take the notion of nonlinear expectation as a fundamental notion of an axiomatical system. This enables us to get directly many new and fundamental results: such as nonlinear law of large numbers, central limit theorem, the theory of Brownian motion under nonlinear expectation and the corresponding new stochastic analysis of Itô's type. Many of them are highly non-trivial, natural but usually hard to be expected from its special case of the classical linear expectation theory, i.e., probability theory. In this new framework, the nonlinear version of the notion of independence and identical distribution plays a crucially important role.

This book is based on the author's Lecture Notes [140] and its extension [144], used in my teaching of postgraduate courses in Shandong University, and several series of lectures, among them, 2nd Workshop on Stochastic Equations and Related Topics, Jena, July 23–29, 2006; Graduate Courses given at Yantai Summer School in Finance, Yantai University, July 06–21, 2007; Graduate Courses at Wuhan Summer School, July 24–26, 2007; Mini-Course given at Institute of Applied Mathematics, AMSS, April 16–18, 2007; Mini-course at Fudan University, May 2007 and August 2009; Graduate Courses at CSFI, Osaka University, May 15–June 13, 2007; Minerva Research Foundation Lectures of Columbia University in Fall of 2008; Mini-Workshop of G-Brownian motion and G-expectations in Weihai, July 2009, a series talks in Department of Applied Mathematics, Hong Kong Polytechnic University, November–December, 2009 and an intensive course in WCU Center for Financial Engineering, Ajou University and graduate courses during 2011–2013 in Princeton University. The hospitalities and encouragements of the above institutions and the enthusiasm of the audiences are the main engine to use these lecture notes and make this book.

I thank, especially, Li Juan and Hu Mingshang for many comments and suggestions given during those courses. During the preparation of this book, a special reading group was organized with members Hu Mingshang, Li Xinpeng, Xu Xiaoming, Lin Yiqing, Su Chen, Wang Falei and Yin Yue. They proposed very helpful suggestions for the revision of the book. Hu Mingshang, Li Xinpeng, Song Yongsheng and Wang Falei have made great efforts and contributed many exercises for the final editing. Their efforts are decisively important for the realization of this book.

The author is grateful to Jordan Stoyanov for reviewing and scrutinizing the whole text. This helped to essentially minimize the inaccuracies and make the text consistent and smooth.

Jinan, China Shige Peng

Introduction

How to measure uncertain quantities is an important problem. A widely accepted mathematical tool is probability theory. But in our real world, it is rare to find an ideal situation in which the probability can be exactly determined. Often the uncertainty of the probability itself becomes a hard problem. In the theory of economics, this type of higher level uncertainty is known as Knightian uncertainty. We refer to Frank Knight [102] and the survey of Lars Hansen (2014).

This book is focused on the recent developments on problems of probability model uncertainty by using the notion of nonlinear expectations and, in particular, sublinear expectations. Roughly speaking, a nonlinear expectation $\mathbb{E}$ is a monotone and constant preserving functional defined on a linear space of random variables. We are particularly interested in sublinear expectations, i.e., $\mathbb{E}[X+Y] \leq \mathbb{E}[X] + \mathbb{E}[Y]$ for all random variables X, Y and $\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X]$ for any constant $\lambda \geq 0$.

A sublinear expectation $\mathbb{E}$ can be represented as the upper expectation of a subset of linear expectations $\{E_\theta : \theta \in \Theta\}$, i.e., $\mathbb{E}[X] = \sup_{\theta \in \Theta} E_\theta[X]$. In most cases, this subset is often treated as an uncertain model of probabilities $\{P_\theta : \theta \in \Theta\}$, and the notion of sublinear expectation provides a robust way to measure a risk loss X. In fact, the sublinear expectation theory provides many rich, flexible and elegant tools.

A remarkable point of view is that we emphasize the term "expectation" rather than the well-accepted classical notion "probability" and its non-additive counterpart "capacity". A technical reason is that in general, the information contained in a nonlinear expectation $\mathbb{E}$ will be lost if one considers only its corresponding "non-additive probability" or "capacity" $\mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A]$. Philosophically, the notion of expectation has its direct meaning of "mean", "average", which is not necessary to be derived from the corresponding "relative frequency" which is the origin of the probability measure. For example, when a person gets a sample $\{x_1, \cdots, x_N\}$ from a random variable X, he/she can directly use $\overline{X} = \frac{1}{N}\sum x_i$ to calculate its mean. In general he/she uses $\overline{\varphi(X)} = \frac{1}{N}\sum \varphi(x_i)$ for the mean of $\varphi(X)$ for a given function φ. We will discuss in detail this issue after the overview of our new law of large numbers (LLN) and central limit theorem (CLT).

A theoretical foundation of the above expectation framework is our new LLN and CLT under sublinear expectations. Classical LLN and CLT have been widely used in probability theory, statistics, data analysis, as well as many practical situations such as financial pricing and risk management. They provide a strong and convincing way to explain why in practice normal distributions are so widely utilized. But often a serious problem is that, in general, the "i.i.d." condition is difficult to be satisfied. In practice, for the most real-time processes and data for which the classical trials and samplings become impossible, the uncertainty of probabilities and distributions can not be neglected. In fact the abuse of normal distributions in finance and many other industrial or commercial domains has been criticized.

Our new CLT does not need this strong "i.i.d." assumption. Instead of fixing a probability measure P, we introduce an uncertain subset of probability measures $\{P_\theta : \theta \in \Theta\}$ and consider the corresponding sublinear expectation $\mathbb{E}[X] = \sup_{\theta\in\Theta} E_\theta[X]$. Our main assumptions are the following:

(i) The distribution of each X_i is within a subset of distributions F_θ, $\theta \in \Theta$, and we don't know which θ is the "right one".
(ii) Any realization of $X_1 = x_1, \cdots, X_n = x_n$ does not change this distributional uncertainty of X_{n+1}.

Under $\mathbb{E}$, we call $X_1, X_2, \cdots$ to be identically distributed if condition (i) is satisfied, and we call X_{n+1} to be independent from $X_1, \cdots, X_n$ if condition (ii) is fulfilled. Mainly under the above much weak "i.i.d." assumptions, we have proved that for each continuous function φ with linear growth we have the following LLN: as $n \to \infty$, $\mathbb{E}[\varphi(S_n/n)] \to \sup_{\underline{\mu}\le v\le\overline{\mu}} \varphi(v)$. Namely, the uncertain subset of distributions of S_n/n is approximately the subset of all probability distribution measures of random variables taken values in $[\underline{\mu}, \overline{\mu}]$. In particular, if $\underline{\mu} = \overline{\mu} = 0$, then S_n/n converges to 0. In this case, if we assume furthermore that $\overline{\sigma}^2 = \mathbb{E}[X_1^2]$ and $\underline{\sigma}^2 = -\mathbb{E}[-X_1^2]$, then we have the following extension of the CLT:

$$\lim_{n\to\infty} \mathbb{E}[\varphi(S_n/\sqrt{n})] = \mathbb{E}[\varphi(X)].$$

Here, the random variable X is G-normally distributed and denoted by $N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$. The value $\mathbb{E}[\varphi(X)]$ can be calculated by defining $u(t,x) := \mathbb{E}[\varphi(x + \sqrt{t}X)]$ which solves the partial differential equation (PDE) $\partial_t u = G(u_{xx})$ with $G(a) := \frac{1}{2}(\overline{\sigma}^2 a^+ - \underline{\sigma}^2 a^-)$. Our results reveal a deep and essential relation between the theory of probability and statistics under uncertainty and second order fully nonlinear parabolic equations (HJB equations). We have two interesting situations: when φ is a convex function, then the value of $\mathbb{E}[\varphi(X)]$ coincides with $\frac{1}{\sqrt{2\pi\overline{\sigma}^2}} \int_{-\infty}^{\infty} \varphi(x) \exp\left(-\frac{x^2}{2\overline{\sigma}^2}\right) \mathrm{d}x$, but if φ is a concave function, the above $\overline{\sigma}^2$ needs to be replaced by $\underline{\sigma}^2$. If $\underline{\sigma} = \overline{\sigma} = \sigma$, then $N(\{0\} \times \underline{\sigma}^2, \overline{\sigma}^2]) = N(0, \sigma^2)$, which is a classical normal distribution.

This result provides a new way to explain a well-known puzzle: many practitioners, e.g., traders and risk officials in financial markets can widely use normal distributions without serious data analysis or even with data inconsistence. In many typical situations $\mathbb{E}[\varphi(X)]$ can be calculated by using normal distributions with careful choice of parameters, but it is also a high risky quantification if the reasoning behind has not been understood.

We call $N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$ the G-normal distribution. This new type of sublinear distributions was first introduced in [138] (see also [139–142]) for a new type of "G-Brownian motion" and the related calculus of Itô's type. The main motivations were uncertainties in statistics, measures of risk and super-hedging in finance (see El Karoui, Peng and Quenez [58], Artzner, Delbaen, Eber and Heath [3], Eber and Heath, Chen and Epstein [28], Föllmer and Schied [69]). Fully nonlinear super-hedging is also a possible domain of applications (see Avellaneda, Lévy and Paras [6], Lyons [114].

Technically, we introduce a new method to prove our CLT on a sublinear expectation space. This proof is short since we have borrowed a deep interior estimate of fully nonlinear partial differential equation (PDE) from Krylov [105]. In fact, the theory of fully nonlinear parabolic PDE plays an essential role in deriving our new results of LLN and CLT. In the classical situation, the corresponding PDE becomes a heat equation which is often hidden behind its heat kernel, i.e., the normal distribution. In this book, we use the powerful notion of viscosity solutions for our nonlinear PDE initially introduced by Crandall and Lions [38]. This notion is especially useful when the equation is degenerate. For reader's convenience, we provide an introductory chapter in Appendix C. If readers are only interested in the classical non-degenerate cases, the corresponding solutions will become smooth (see the last section of Appendix C).

We define a sublinear expectation on the space of continuous paths from $\mathbb{R}^+$ to $\mathbb{R}^d$ which is an analogue of Wiener's law, by which a G-Brownian motion is formulated. Briefly speaking, a G-Brownian motion $(B_t)_{t \geq 0}$ is a continuous process with independent and stationary increments under a given sublinear expectation $\mathbb{E}$.

G-Brownian motion has a very rich and interesting new structure which non-trivially generalizes the classical one. We can develop the related stochastic calculus, especially Itô's integrals and the related quadratic variation process $\langle B \rangle$. A very interesting new phenomenon of the G-Brownian motion is that its quadratic variation process $\langle B \rangle$ is also a continuous process with independent and stationary increments, and thus can be still regarded as a Brownian motion. The corresponding G-Itô's formula is obtained. We have also established the existence and uniqueness of solutions to stochastic differential equation based on this new stochastic calculus by the same Picard iterations as in the classical situation.

New norms were introduced in the notion of G-expectation by which the corresponding stochastic calculus becomes significantly more flexible and powerful. Many interesting, attractive and challenging problems are also provided within this new framework.

In this book, we adopt a novel method to present our G-Brownian motion theory. In the first two chapters, as well as the first two sections of Chap. 3, our sublinear expectations are only assumed to be finitely sub-additive, instead of "σ-sub-additive". This is just because all the related results obtained in this part do not need the "σ-sub-additive" assumption, and readers even need not to have the background of classical probability theory. In fact, in the whole part of the first five chapters, we only use a very basic knowledge of functional analysis such as Hahn–Banach Theorem (see Appendix A). A special situation is when all the sublinear expectations treated in this book become linear. In this case, the book still can be considered as a source based on a new, simple and rigorous approach to introduce the classical Itô's stochastic calculus, since we do not need the knowledge of probability theory. This is an important advantage to use expectation as the basic notion.

The "authentic probabilistic parts", i.e., the pathwise analysis of G-Brownian motion and the corresponding random variables, viewed as functions of G-Brownian path, is presented in Chap. 6. Here just as in the classical "P-sure analysis", we introduce "$\hat{c}$-sure analysis" for G-capacity $\hat{c}$.

In Chap. 7, we present a highly nontrivial generalization of the classical martingale representation theorem. In a nonlinear G-Brownian motion framework, a G-martingale can be decomposed into two essentially different martingales, the first one is an Itô's integral with respect to the G-Brownian motion B, and the second one is a non-increasing G-martingale. The later term vanishes once G is a linear function and thus B becomes to a classical Brownian motion.

In Chap. 8, we use the quasi-surely analysis theory to develop Itô's integrals without the quasi-continuity condition. This allows us to define Itô's integral on stopping time interval. In particular, this new formulation can be applied to obtain Itô's formula for a general $C^{1,2}$-function, thus extend previously available results.

For reader's convenience, we provide some preliminary results in functional analysis, probability theory and nonlinear partial differential equations of parabolic types in Appendix A, B and C, respectively.

Contents

Part I
Basic Theory of Nonlinear Expectations

Chapter 1
Sublinear Expectations and Risk Measures

A sublinear expectation is also called the upper expectation or the upper prevision, and this notion is used in situations when the probability models have uncertainty. In this chapter, we present the basic notion of sublinear expectations and the corresponding sublinear expectation spaces. We give the representation theorem of a sublinear expectation and the notions of distributions and independence within the framework of sublinear expectations. We also introduce a natural Banach norm of a sublinear expectation in order to get the completion of a sublinear expectation space which is a Banach space. As a fundamentally important example, we introduce the notion of coherent risk measures in finance. A large part of the notions and the results in this chapter will be used throughout the book.

1.1 Sublinear Expectations and Sublinear Expectation Spaces

Let Ω be a given set and let $\mathcal{H}$ be a linear space of real valued functions defined on Ω. In this book, we suppose that $\mathcal{H}$ satisfies the following two conditions:

(1) $c \in \mathcal{H}$ for each constant c;
(2) $|X| \in \mathcal{H}$ if $X \in \mathcal{H}$.

In this book, the space $\mathcal{H}$ will be used as the space of random variables.

Definition 1.1.1 A **Sublinear expectation** $\mathbb{E}$ is a functional $\mathbb{E} : \mathcal{H} \mapsto \mathbb{R}$ satisfying
(i) Monotonicity:

$$\mathbb{E}[X] \leq \mathbb{E}[Y] \text{ if } X \leq Y.$$

(ii) Constant preserving:

$$\mathbb{E}[c] = c \quad \text{for } c \in \mathbb{R}.$$

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_1

(iii) Sub-additivity: For each $X, Y \in \mathcal{H}$,

$$\mathbb{E}[X + Y] \le \mathbb{E}[X] + \mathbb{E}[Y].$$

(iv) Positive homogeneity:

$$\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X] \quad \text{for } \lambda \ge 0.$$

The triplet $(\Omega, \mathcal{H}, \mathbb{E})$ is called a **sublinear expectation space**. If (i) and (ii) are satisfied, $\mathbb{E}$ is called a **nonlinear expectation** and the triplet $(\Omega, \mathcal{H}, \mathbb{E})$ is called a **nonlinear expectation space**.

Definition 1.1.2 Let $\mathbb{E}_1$ and $\mathbb{E}_2$ be two nonlinear expectations defined on $(\Omega, \mathcal{H})$. We say that $\mathbb{E}_1$ is **dominated** by $\mathbb{E}_2$, or $\mathbb{E}_2$ **dominates** $\mathbb{E}_1$, if

$$\mathbb{E}_1[X] - \mathbb{E}_1[Y] \le \mathbb{E}_2[X - Y] \quad \text{for } X, Y \in \mathcal{H}. \tag{1.1.1}$$

Remark 1.1.3 From (iii), a sublinear expectation is dominated by itself. In many situations, (iii) is also called the property of self-domination. If the inequality in (iii) becomes equality, then $\mathbb{E}$ is a linear expectation, i.e., $\mathbb{E}$ is a linear functional satisfying Properties (i) and (ii).

Remark 1.1.4 Properties (iii)+(iv) are called **sublinearity**. This sublinearity implies
(v) Convexity:

$$\mathbb{E}[\alpha X + (1 - \alpha)Y] \le \alpha \mathbb{E}[X] + (1 - \alpha)\mathbb{E}[Y] \quad \text{for } \alpha \in [0, 1].$$

If a nonlinear expectation $\mathbb{E}$ satisfies the convexity property, we call it a **convex expectation**.

Properties (ii)+(iii) imply
(vi) Cash translatability:

$$\mathbb{E}[X + c] = \mathbb{E}[X] + c \quad \text{for } c \in \mathbb{R}.$$

In fact, we have

$$\mathbb{E}[X] + c = \mathbb{E}[X] - \mathbb{E}[-c] \le \mathbb{E}[X + c] \le \mathbb{E}[X] + \mathbb{E}[c] = \mathbb{E}[X] + c.$$

For Property (iv), an equivalent form is

$$\mathbb{E}[\lambda X] = \lambda^+ \mathbb{E}[X] + \lambda^- \mathbb{E}[-X] \quad \text{for } \lambda \in \mathbb{R}.$$

In this book, we will systematically study the sublinear expectation spaces. In the following chapters, unless otherwise stated, we consider the following sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$: if $X_1, \cdots, X_n \in \mathcal{H}$, then $\varphi(X_1, \cdots, X_n) \in \mathcal{H}$ for each $\varphi \in C_{l.Lip}(\mathbb{R}^n)$, where $C_{l.Lip}(\mathbb{R}^n)$ denotes the linear space of functions φ satisfying the following local Lipschitz condition:

$$|\varphi(x) - \varphi(y)| \le C(1 + |x|^m + |y|^m)|x - y| \text{ for } x, y \in \mathbb{R}^n,$$
$$\text{where the constant } C > 0 \text{ and the integer } m \in \mathbb{N} \text{ depend on } \varphi.$$

Often an n-dimensional random variable $X=(X_1, \cdots, X_n)$ is called an n-dimensional random vector, denoted by $X \in \mathcal{H}^n$.

Here we mainly use $C_{l.Lip}(\mathbb{R}^n)$ in our framework, and this is only for convenience of techniques. In fact, our essential requirement is that $\mathcal{H}$ contains all constants and, moreover, $X \in \mathcal{H}$ implies $|X| \in \mathcal{H}$. In practice, $C_{l.Lip}(\mathbb{R}^n)$ can be replaced by any one of the following spaces of functions defined on $\mathbb{R}^n$:

- $L^0(\mathbb{R}^n)$: the space of Borel measurable functions;
- $\mathbb{L}^\infty(\mathbb{R}^n)$: the space of bounded Borel-measurable functions;
- $C_b(\mathbb{R}^n)$: the space of bounded and continuous functions;
- $C_b^k(\mathbb{R}^n)$: the space of bounded and k-times continuously differentiable functions with bounded derivatives of all orders less than or equal to k;
- $C_{l.Lip}^k(\mathbb{R}^n)$: the space of k-times continuously differentiable functions, whose partial derivatives of all orders less than or equal to k are in $C_{l,Lip}(\mathbb{R}^n)$;
- $C_{Lip}(\mathbb{R}^n)$: the space of Lipschitz continuous functions;
- $C_{b.Lip}(\mathbb{R}^n)$: the space of bounded and Lipschitz continuous functions;
- $C_{unif}(\mathbb{R}^n)$: the space of bounded and uniformly continuous functions.
- $USC(\mathbb{R}^n)$: the space of upper semi continuous functions on $\mathbb{R}^n$
- $LSC(\mathbb{R}^n)$: the space of lower semi continuous functions on $\mathbb{R}^n$.

Next we give two examples of sublinear expectations.

Example 1.1.5 In a game a gambler randomly pick a ball from an urn containing W white, B black and Y yellow balls. The owner of the urn, who is the banker of the game, does not tell the gambler the exact numbers of W, B and Y. He/She only ensures that $W + B + Y = 100$ and $W = B \in [20, 25]$. Let ξ be a random variable defined by

$$\xi = \begin{cases} 1 & \text{if the picked ball is white;} \\ 0 & \text{if the picked ball is yellow;} \\ -1 & \text{if the picked ball is black.} \end{cases}$$

Problem: how to conservatively measure the loss $X = \varphi(\xi)$ for a given local Lipschitz function φ on $\mathbb{R}$.

We know that the distribution of ξ is

$$\left\{ \begin{matrix} -1 & 0 & 1 \\ \frac{p}{2} & 1-p & \frac{p}{2} \end{matrix} \right\} \text{ with uncertainty: } p \in [\underline{\mu}, \overline{\mu}] = [0.4, 0.5].$$

Thus the **robust expectation** of ξ is

$$\begin{aligned}\mathbb{E}[\varphi(\xi)] &:= \sup_{P\in\mathcal{P}} E_P[\varphi(\xi)]\\ &= \sup_{p\in[\underline{\mu},\overline{\mu}]} [\tfrac{p}{2}(\varphi(1)+\varphi(-1))+(1-p)\varphi(0)].\end{aligned}$$

Notice, in this example, that ξ has distribution uncertainty.

Example 1.1.6 A more general situation is that the banker of a game can choose a distribution from a set of distributions $\{F(\theta, A)\}_{A\in\mathcal{B}(\mathbb{R}),\theta\in\Theta}$ of a random variable ξ. In this situation the robust expectation of the risk position $\varphi(\xi)$ for a given function $\varphi \in C_{l.Lip}(\mathbb{R})$ is

$$\mathbb{E}[\varphi(\xi)] := \sup_{\theta\in\Theta} \int_{\mathbb{R}} \varphi(x)F(\theta, dx).$$

1.2 Representation of a Sublinear Expectation

A sublinear expectation can be expressed as a supremum of linear expectations.

Theorem 1.2.1 *Let $\mathbb{E}$ be a functional defined on a linear space $\mathcal{H}$ satisfying sub-additivity and positive homogeneity. Then there exists a family of linear functionals $E_\theta : \mathcal{H} \mapsto \mathbb{R}$, indexed by $\theta \in \Theta$, such that*

$$\mathbb{E}[X] = \max_{\theta\in\Theta} E_\theta[X] \ \ \textit{for } X \in \mathcal{H}. \tag{1.2.1}$$

Moreover, for each $X \in \mathcal{H}$, there exists $\theta_X \in \Theta$ such that $\mathbb{E}[X] = E_{\theta_X}[X]$.

Furthermore, if $\mathbb{E}$ is a sublinear expectation, then the corresponding E_θ is a linear expectation.

Proof Let $\mathcal{Q} = \{E_\theta : \theta \in \Theta\}$ be the family of all linear functionals dominated by $\mathbb{E}$, i.e., $E_\theta[X] \le \mathbb{E}[X]$, for all $X \in \mathcal{H}$, $E_\theta \in \mathcal{Q}$.

Let us first prove that this $\mathcal{Q}$ is not an empty set. For a given $X \in \mathcal{H}$, we set $L = \{aX : a \in \mathbb{R}\}$ which is a subspace of $\mathcal{H}$. We define a linear functional $I : L \mapsto \mathbb{R}$ by $I[aX] := a\mathbb{E}[X]$, $\forall a \in \mathbb{R}$. Then $I[\cdot]$ forms a linear functional on L and $I \le \mathbb{E}$ on L. Since $\mathbb{E}[\cdot]$ is sub-additive and positively homogeneous, by Hahn–Banach theorem (see Appendix A), there exists a linear functional E defined on $\mathcal{H}$ such that $E = I$ on L and $E \le \mathbb{E}$ on $\mathcal{H}$. Thus this E is a linear functional dominated by $\mathbb{E}$ such that $\mathbb{E}[X] = E[X]$, namely (1.2.1) holds.

Furthermore, if $\mathbb{E}$ is a sublinear expectation, then for each E_θ and each nonnegative element $X \in \mathcal{H}$, we have $E_\theta[X] = -E_\theta[-X] \ge -\mathbb{E}[-X] \ge 0$. Moreover, since for each $c \in \mathbb{R}$,

$$-E_\theta[c] = E_\theta[-c] \le \mathbb{E}[-c] = -c, \quad \textit{and} \quad E_\theta[c] \le \mathbb{E}[c] = c,$$

E_θ also preserves constants. Hence it is a linear expectation. □

Observe that the above linear expectation E_θ is possibly finitely additive. We now give an important sufficient condition for the σ-additivity of such E_θ:

Theorem 1.2.2 (Robust Daniell-Stone Theorem) *Assume that* $(\Omega, \mathcal{H}, \mathbb{E})$ *is a sublinear expectation space satisfying*

$$\mathbb{E}[X_i] \to 0, \quad as \quad i \to \infty, \tag{1.2.2}$$

for each sequence $\{X_i\}_{i=1}^{\infty}$ *of random variables in* $\mathcal{H}$ *such that* $X_i(\omega) \downarrow 0$ *for each* $\omega \in \Omega$. *Then there exists a family of probability measures* $\{P_\theta\}_{\theta\in\Theta}$ *defined on the measurable space* $(\Omega, \sigma(\mathcal{H}))$ *such that*

$$\mathbb{E}[X] = \max_{\theta\in\Theta} \int_\Omega X(\omega) dP_\theta, \ for\ each\ X \in \mathcal{H}. \tag{1.2.3}$$

Here $\sigma(\mathcal{H})$ *is the smallest* σ*-algebra generated by* $\mathcal{H}$.

Proof Theorem 1.2.1 implies that there exists a family of linear expectations $\{E_\theta\}_{\theta\in\Theta}$ defined on the measurable space $(\Omega, \sigma(\mathcal{H}))$ such that $\mathbb{E}[X] = \max_{\theta\in\Theta} E_\theta[X]$ for each element $X \in \mathcal{H}$. Note that Condition (1.2.2) implies that $E_\theta[X_i] \downarrow 0$ for each $\theta \in \Theta$. It then follows from the well-known Daniell-Stone theorem (see Theorem 3.3 in Appendix B) that there exists a unique probability P_θ defined on $(\Omega, \sigma(\mathcal{H}))$ such that $E_\theta[X] = \int_\Omega X(\omega) dP_\theta$ which implies (1.2.3). □

Remark 1.2.3 We call the subset $\{P_\theta\}_{\theta\in\Theta}$ the uncertain probability measures associated to the sublinear expectation $\mathbb{E}$. $\{P_\theta : \theta \in \Theta\}$. For a given n-dimensional random vector $X \in \mathcal{H}$, we set $\{F_X(\theta, A) = P_\theta(X \in A),\ \ A \in \mathcal{B}(\mathbb{R}^n)\}_{\theta\in\Theta}$ and call it the uncertain probability distributions of X.

1.3 Distributions, Independence and Product Spaces

We now give the notion of distributions of random variables under sublinear expectations.

Let $X = (X_1, \cdots, X_n)$ be a given n-dimensional random vector on a nonlinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$. We define a functional on $C_{l.Lip}(\mathbb{R}^n)$ by

$$\mathbb{F}_X[\varphi] := \mathbb{E}[\varphi(X)] : \varphi \in C_{l.Lip}(\mathbb{R}^n) \mapsto \mathbb{R}.$$

The triplet $(\mathbb{R}^n, C_{l.Lip}(\mathbb{R}^n), \mathbb{F}_X)$ forms a nonlinear expectation space. $\mathbb{F}_X$ is called the **distribution** of X under $\mathbb{E}$. This notion is very useful for a sublinear expectation $\mathbb{E}$. In this case $\mathbb{F}_X$ is also a sublinear expectation. Furthermore we can prove that (see Theorem 1.2.2), there exists a family of probability measures $\{F_X(\theta, \cdot)\}_{\theta\in\Theta}$ defined on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that

$$\mathbb{F}_X[\varphi] = \sup_{\theta \in \Theta} \int_{\mathbb{R}^n} \varphi(x) F_X(\theta, dx), \text{ for each } \varphi \in C_{l.Lip}(\mathbb{R}^n).$$

Thus $\mathbb{F}_X[\cdot]$ characterizes the uncertainty of the distributions of X.

Definition 1.3.1 Let X_1 and X_2 be two n–dimensional random vectors defined on nonlinear expectation spaces $(\Omega_1, \mathcal{H}_1, \mathbb{E}_1)$ and $(\Omega_2, \mathcal{H}_2, \mathbb{E}_2)$, respectively. They are called **identically distributed**, denoted by $X_1 \overset{d}{=} X_2$, if

$$\mathbb{E}_1[\varphi(X_1)] = \mathbb{E}_2[\varphi(X_2)] \text{ for all } \varphi \in C_{b.Lip}(\mathbb{R}^n).$$

We say that the distribution of X_1 is stronger than that of X_2 if $\mathbb{E}_1[\varphi(X_1)] \geq \mathbb{E}_2[\varphi(X_2)]$, for each $\varphi \in C_{b.Lip}(\mathbb{R}^n)$.

Note that $X \in \mathcal{H}^n$ implies that $\varphi(X) \in \mathcal{H}$ for each $\varphi \in C_{l.Lip}(\mathbb{R}^n)$. Then in our framework, the identically distributed can also be characterized by the following.

Proposition 1.3.2 *Suppose that $X_1 \overset{d}{=} X_2$. Then*

$$\mathbb{E}_1[\varphi(X_1)] = \mathbb{E}_2[\varphi(X_2)] \textit{ for all } \varphi \in C_{l.Lip}(\mathbb{R}^n).$$

Moreover $X_1 \overset{d}{=} X_2$ if and only if their distributions coincide.

Proof For each $\varphi \in C_{l.Lip}(\mathbb{R}^n)$ and integer $N \geq 1$, we define

$$\varphi_N(x) := \varphi((x \wedge N) \vee (-N)), \ \forall x \in \mathbb{R}^n.$$

It is easy to check that $\varphi_N \in C_{b.Lip}(\mathbb{R}^n)$. Moreover, there exist some constants C and m such that

$$|\varphi_N(x) - \varphi(x)| \leq C(1 + |x|^m)(|x| - N)^+ \leq C(1 + |x|^m)\frac{|x|^2}{N},$$

from which we deduce that

$$\lim_{N \to \infty} \mathbb{E}_i[|\varphi_N(X_i) - \varphi(X_i)|] = 0, \ i = 1, 2.$$

Consequently,

$$\mathbb{E}_1[\varphi(X_1)] = \lim_{N \to \infty} \mathbb{E}_1[|\varphi_N(X_1)] = \lim_{N \to \infty} \mathbb{E}_2[|\varphi_N(X_2)] = \mathbb{E}_2[\varphi(X_2)],$$

which is the desired result. □

Remark 1.3.3 In many cases of sublinear expectations, $X_1 \overset{d}{=} X_2$ implies that the uncertainty subsets of distributions of X_1 and X_2 are the same, e.g., in view of Remark 1.2.3,

$$\{F_{X_1}(\theta_1,\cdot):\theta_1\in\Theta_1\}=\{F_{X_2}(\theta_2,\cdot):\theta_2\in\Theta_2\}.$$

Similarly, if the distribution of X_1 is stronger than that of X_2, then

$$\{F_{X_1}(\theta_1,\cdot):\theta_1\in\Theta_1\}\supset\{F_{X_2}(\theta_2,\cdot):\theta_2\in\Theta_2\}.$$

The distribution of $X\in\mathcal{H}$ has the following typical parameters:

$$\bar{\mu}:=\mathbb{E}[X],\ \underline{\mu}:=-\mathbb{E}[-X].$$

The interval $[\underline{\mu},\bar{\mu}]$ characterizes the **mean-uncertainty** of X.

A natural question is: can we find a family of distribution measures to represent the above sublinear distribution of X? The answer is affirmative.

Lemma 1.3.4 *Let $(\Omega,\mathcal{H},\mathbb{E})$ be a sublinear expectation space. Let $X\in\mathcal{H}^d$ be given. Then for each sequence $\{\varphi_n\}_{n=1}^{\infty}\subset C_{l.Lip}(\mathbb{R}^d)$ satisfying $\varphi_n\downarrow 0$, as $n\to\infty$, we have $\mathbb{E}[\varphi_n(X)]\downarrow 0$.*

Proof For each fixed $N>0$,

$$\varphi_n(x)\le k_{n,N}+\varphi_1(x)\mathbf{1}_{[|x|>N]}\le k_{n,N}+\frac{\varphi_1(x)|x|}{N}\text{ for each }x\in\mathbb{R}^{d\times m},$$

where $k_{n,N}=\max_{|x|\le N}\varphi_n(x)$. We then have

$$\mathbb{E}[\varphi_n(X)]\le k_{n,N}+\frac{1}{N}\mathbb{E}[\varphi_1(X)|X|].$$

It follows from $\varphi_n\downarrow 0$ that $k_{n,N}\downarrow 0$. Thus we have $\lim_{n\to\infty}\mathbb{E}[\varphi_n(X)]\le\frac{1}{N}\mathbb{E}[\varphi_1(X)|X|]$. Since N can be arbitrarily large, we get $\mathbb{E}[\varphi_n(X)]\downarrow 0$ as $n\to\infty$. □

Lemma 1.3.5 *Let $(\Omega,\mathcal{H},\mathbb{E})$ be a sublinear expectation space and let $\mathbb{F}_X[\varphi]:=\mathbb{E}[\varphi(X)]$ be the sublinear distribution of $X\in\mathcal{H}^d$. Then there exists a family of probability measures $\{F_\theta\}_{\theta\in\Theta}$ defined on $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ such that*

$$\mathbb{F}_X[\varphi]=\sup_{\theta\in\Theta}\int_{\mathbb{R}^d}\varphi(x)F_\theta(dx),\ \varphi\in C_{l.Lip}(\mathbb{R}^d).\qquad(1.3.1)$$

Proof From Lemma 1.3.4, the sublinear expectation $\mathbb{F}_X[\varphi]:=\mathbb{E}[X]$ defined on the space $(\mathbb{R}^d,C_{l.Lip}(\mathbb{R}^n))$ satisfies the following downwardly continuous property: for each sequence $\{\varphi_n\}_{n=1}^{\infty}\subset C_{l.Lip}(\mathbb{R}^d)$ satisfying $\varphi_n\downarrow 0$ as $n\to\infty$, we have $\mathbb{F}_X[\varphi_n]\downarrow 0$. The proof then follows directly from the robust Daniell-Stone Theorem 1.2.2. □

Remark 1.3.6 Lemma 1.3.5 tells us that in fact the sublinear distribution $\mathbb{F}_X$ of X characterizes the uncertainty of the distribution of X which is a subset of distributions $\{F_\theta\}_{\theta\in\Theta}$.

The following property is very useful in the sublinear expectation theory.

Proposition 1.3.7 *Let* $(\Omega, \mathcal{H}, \mathbb{E})$ *be a sublinear expectation space and* X, Y *be two random variables such that* $\mathbb{E}[Y] = -\mathbb{E}[-Y]$*, i.e.,* Y *has no mean-uncertainty. Then we have*

$$\mathbb{E}[X + \alpha Y] = \mathbb{E}[X] + \alpha\mathbb{E}[Y] \textit{ for } \alpha \in \mathbb{R}.$$

In particular, if $\mathbb{E}[Y] = \mathbb{E}[-Y] = 0$*, then* $\mathbb{E}[X + \alpha Y] = \mathbb{E}[X]$.

Proof We have

$$\mathbb{E}[\alpha Y] = \alpha^+\mathbb{E}[Y] + \alpha^-\mathbb{E}[-Y] = \alpha^+\mathbb{E}[Y] - \alpha^-\mathbb{E}[Y] = \alpha\mathbb{E}[Y] \text{ for } \alpha \in \mathbb{R}.$$

Thus

$$\mathbb{E}[X + \alpha Y] \le \mathbb{E}[X] + \mathbb{E}[\alpha Y] = \mathbb{E}[X] + \alpha\mathbb{E}[Y] = \mathbb{E}[X] - \mathbb{E}[-\alpha Y] \le \mathbb{E}[X + \alpha Y].$$

□

A more general form of the above proposition is:

Proposition 1.3.8 *We make the same assumptions as in the previous proposition. Let* $\tilde{\mathbb{E}}$ *be a nonlinear expectation on* $(\Omega, \mathcal{H})$ *dominated by the sublinear expectation* $\mathbb{E}$ *in the sense of Definition 1.1.2. If* $\mathbb{E}[Y] = -\mathbb{E}[-Y]$*, then we have*

$$\tilde{\mathbb{E}}[\alpha Y] = \alpha\tilde{\mathbb{E}}[Y] = \alpha\mathbb{E}[Y],\ \alpha \in \mathbb{R}, \tag{1.3.2}$$

as well as

$$\tilde{\mathbb{E}}[X + \alpha Y] = \tilde{\mathbb{E}}[X] + \alpha\tilde{\mathbb{E}}[Y], X \in \mathcal{H},\ \alpha \in \mathbb{R}. \tag{1.3.3}$$

In particular

$$\tilde{\mathbb{E}}[X + c] = \tilde{\mathbb{E}}[X] + c,\ \textit{for } c \in \mathbb{R}. \tag{1.3.4}$$

Proof We have

$$-\tilde{\mathbb{E}}[Y] = \tilde{\mathbb{E}}[0] - \tilde{\mathbb{E}}[Y] \le \mathbb{E}[-Y] = -\mathbb{E}[Y] \le -\tilde{\mathbb{E}}[Y]$$

and

$$\begin{aligned}\mathbb{E}[Y] &= -\mathbb{E}[-Y] \le -\tilde{\mathbb{E}}[-Y] \\ &= \tilde{\mathbb{E}}[0] - \tilde{\mathbb{E}}[-Y] \le \mathbb{E}[Y].\end{aligned}$$

From these relations we conclude that $\tilde{\mathbb{E}}[Y] = \mathbb{E}[Y] = -\tilde{\mathbb{E}}[-Y]$ and thus (1.3.2). Still by the domination,

$$\tilde{\mathbb{E}}[X + \alpha Y] - \tilde{\mathbb{E}}[X] \leq \mathbb{E}[\alpha Y],$$
$$\tilde{\mathbb{E}}[X] - \tilde{\mathbb{E}}[X + \alpha Y] \leq \mathbb{E}[-\alpha Y] = -\mathbb{E}[\alpha Y].$$

Therefore (1.3.3) holds. □

Definition 1.3.9 A sequence of n-dimensional random vectors $\{\eta_i\}_{i=1}^{\infty}$ defined on a nonlinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ is said to **converge in distribution** (or **converge in law**) under $\mathbb{E}$ if for each $\varphi \in C_{b.Lip}(\mathbb{R}^n)$, the sequence $\{\mathbb{E}[\varphi(\eta_i)]\}_{i=1}^{\infty}$ converges.

The following result is easy to check.

Proposition 1.3.10 *Let $\{\eta_i\}_{i=1}^{\infty}$ converge in law in the above sense. Then the mapping $\mathbb{F}[\cdot] : C_{b.Lip}(\mathbb{R}^n) \mapsto \mathbb{R}$ defined by*

$$\mathbb{F}[\varphi] := \lim_{i\to\infty} \mathbb{E}[\varphi(\eta_i)] \textit{for } \varphi \in C_{b.Lip}(\mathbb{R}^n)$$

is a sublinear expectation defined on $(\mathbb{R}^n, C_{b.Lip}(\mathbb{R}^n))$.

The following notion of independence plays a key role in the nonlinear expectation theory.

Definition 1.3.11 In a nonlinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$, a random vector $Y \in \mathcal{H}^n$ is said to be **independent** of another random vector $X \in \mathcal{H}^m$ under $\mathbb{E}$ if for each test function $\varphi \in C_{b.Lip}(\mathbb{R}^{m+n})$ we have

$$\mathbb{E}[\varphi(X, Y)] = \mathbb{E}[\mathbb{E}[\varphi(x, Y)]_{x=X}].$$

Remark 1.3.12 Note that the space $C_{b.Lip}(\mathbb{R}^{m+n})$ can be replaced by $C_{l.Lip}(\mathbb{R}^{m+n})$ due to the assumptions on $\mathcal{H}$. This is left as an exercise.

Remark 1.3.13 The situation "Y is independent of X" often appears when Y occurs after X, thus a robust expectation should take the information of X into account.

Remark 1.3.14 In a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$, Y is independent of X means that the uncertainty of distributions $\{F_Y(\theta, \cdot) : \theta \in \Theta\}$ of Y does not change after each realization of $X = x$. In other words, the "conditional sublinear expectation" of Y with respect to X is $\mathbb{E}[\varphi(x, Y)]_{x=X}$. In the case of linear expectation, this notion of independence is just the classical one.

It is important to observe that, under a nonlinear expectation, Y is independent of X does not in general imply that X is independent of Y. An illustration follows.

Example 1.3.15 We consider a case where $\mathbb{E}$ is a sublinear expectation and $X, Y \in \mathcal{H}$ are identically distributed with $\mathbb{E}[X] = \mathbb{E}[-X] = 0$ and $\overline{\sigma}^2 = \mathbb{E}[X^2] > \underline{\sigma}^2 = -\mathbb{E}[-X^2]$. We also assume that $\mathbb{E}[|X|] > 0$, thus $\mathbb{E}[X^+] = \frac{1}{2}\mathbb{E}[|X| + X] = \frac{1}{2}\mathbb{E}[|X|] > 0$. In the case of Y independent of X, we have

$$\mathbb{E}[XY^2] = \mathbb{E}[X^+\overline{\sigma}^2 - X^-\underline{\sigma}^2] = (\overline{\sigma}^2 - \underline{\sigma}^2)\mathbb{E}[X^+] > 0.$$

This contradicts the relation

$$\mathbb{E}[XY^2] = 0,$$

which is true if X is independent of Y.

The independence property of two random vectors X, Y involves only the "joint distribution" of (X, Y). The following result tells us how to construct random vectors with given "marginal distributions" and with a specific direction of independence.

Definition 1.3.16 Let $(\Omega_i, \mathcal{H}_i, \mathbb{E}_i)$, $i = 1, 2$ be two sublinear (resp. nonlinear) expectation spaces. We denote

$$\mathcal{H}_1 \otimes \mathcal{H}_2 := \{Z(\omega_1, \omega_2) = \varphi(X(\omega_1), Y(\omega_2)) : (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2, \\ X \in \mathcal{H}_1^m, \ \ Y \in \mathcal{H}_2^n, \ \ \varphi \in C_{l.Lip}(\mathbb{R}^{m+n}), \ m, n = 1, 2, \cdots, \}$$

and, for each random variable of the above form $Z(\omega_1, \omega_2) = \varphi(X(\omega_1), Y(\omega_2))$,

$$(\mathbb{E}_1 \otimes \mathbb{E}_2)[Z] := \mathbb{E}_1[\bar{\varphi}(X)], \text{ where } \bar{\varphi}(x) := \mathbb{E}_2[\varphi(x, Y)], \ x \in \mathbb{R}^m.$$

It is easy to check that the triplet $(\Omega_1 \times \Omega_2, \mathcal{H}_1 \otimes \mathcal{H}_2, \mathbb{E}_1 \otimes \mathbb{E}_2)$ forms a sublinear (resp. nonlinear) expectation space. We call it the **product space** of sublinear (resp. nonlinear) expectation spaces $(\Omega_1, \mathcal{H}_1, \mathbb{E}_1)$ and $(\Omega_2, \mathcal{H}_2, \mathbb{E}_2)$. In this way, we can define the product space

$$\left(\prod_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathcal{H}_i, \bigotimes_{i=1}^n \mathbb{E}_i\right)$$

of given sublinear (resp. nonlinear) expectation spaces $(\Omega_i, \mathcal{H}_i, \mathbb{E}_i), i = 1, 2, \cdots, n$. In particular, when $(\Omega_i, \mathcal{H}_i, \mathbb{E}_i) = (\Omega_1, \mathcal{H}_1, \mathbb{E}_1)$ we have the product space of the form $(\Omega_1^n, \mathcal{H}_1^{\otimes n}, \mathbb{E}_1^{\otimes n})$.

Let $X, \bar{X}$ be two n-dimensional random vectors on a sublinear (resp. nonlinear) expectation space $(\Omega, \mathcal{H}, \mathbb{E})$. $\bar{X}$ is called an independent copy of X if $\bar{X} \overset{d}{=} X$ and $\bar{X}$ is independent of X.

The following property is easy to check.

Proposition 1.3.17 *Let X_i be an n_i-dimensional random vector on sublinear (resp. nonlinear) expectation space $(\Omega_i, \mathcal{H}_i, \mathbb{E}_i)$ for $i = 1, \cdots, n$, respectively. We denote*

$$Y_i(\omega_1, \cdots, \omega_n) := X_i(\omega_i), \ i = 1, \cdots, n.$$

Then Y_i, $i = 1, \cdots, n$ are random vectors on $(\prod_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathcal{H}_i, \bigotimes_{i=1}^n \mathbb{E}_i)$. Moreover we have $Y_i \overset{d}{=} X_i$ and Y_{i+1} is independent of $(Y_1, \cdots, Y_i)$, for each $i = 1, 2, \cdots, n-1$.

Furthermore, if $(\Omega_i, \mathcal{H}_i, \mathbb{E}_i) = (\Omega_1, \mathcal{H}_1, \mathbb{E}_1)$ *and* $X_i \overset{d}{=} X_1$ *for all* i*, then we also have* $Y_i \overset{d}{=} Y_1$*. In this case* Y_i *is said to be an* ***independent copy*** *of* Y_1 *for* $i = 2, \cdots, n$.

Remark 1.3.18 In the above construction the integer n can also be infinity. In this case each random variable $X \in \bigotimes_{i=1}^{\infty} \mathcal{H}_i$ belongs to $(\prod_{i=1}^{k} \Omega_i, \bigotimes_{i=1}^{k} \mathcal{H}_i, \bigotimes_{i=1}^{k} \mathbb{E}_i)$ for some positive integer $k < \infty$ and, for this X, we have

$$\bigotimes_{i=1}^{\infty} \mathbb{E}_i[X] := \bigotimes_{i=1}^{k} \mathbb{E}_i[X].$$

In a sublinear expectation space we have:

Example 1.3.19 We consider a situation where two random variables X and Y in $\mathcal{H}$ are identically distributed and their common distribution is

$$\mathbb{F}_X[\varphi] = \mathbb{F}_Y[\varphi] = \sup_{\theta \in \Theta} \int_{\mathbb{R}} \varphi(y) F(\theta, dy) \text{for } \varphi \in C_{l.Lip}(\mathbb{R}),$$

where for each $\theta \in \Theta$, $\{F(\theta, A)\}_{A \in \mathcal{B}(\mathbb{R})}$ is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. In this case, "$Y$ is independent of X" means that the joint distribution of X and Y is

$$\mathbb{F}_{X,Y}[\psi] = \sup_{\theta_1 \in \Theta} \int_{\mathbb{R}} \left[\sup_{\theta_2 \in \Theta} \int_{\mathbb{R}} \psi(x, y) F(\theta_2, dy) \right] F(\theta_1, dx) \text{ for } \psi \in C_{l.Lip}(\mathbb{R}^2).$$

1.4 Completion of Sublinear Expectation Spaces

Let $(\Omega, \mathcal{H}, \mathbb{E})$ be a sublinear expectation space. Then we have the following useful inequalities.

We recall the well-known classical inequalities.

Lemma 1.4.1 *For* $r > 0$ *and* $1 < p, q < \infty$ *with* $\frac{1}{p} + \frac{1}{q} = 1$*, we have*

$$|a + b|^r \leq \max\{1, 2^{r-1}\}(|a|^r + |b|^r) \textit{ for } a, b \in \mathbb{R}; \tag{1.4.1}$$

$$|ab| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}, \quad \textit{(Young's inequality)}. \tag{1.4.2}$$

Proposition 1.4.2 *For each* $X, Y \in \mathcal{H}$*, we have*

$$\mathbb{E}[|X + Y|^r] \leq \max\{1, 2^{r-1}\}(\mathbb{E}[|X|^r] + \mathbb{E}[|Y|^r]), \quad \textit{for } r > 0; \tag{1.4.3}$$

$$\mathbb{E}[|XY|] \leq (\mathbb{E}[|X|^p])^{1/p} \cdot (\mathbb{E}[|Y|^q])^{1/q}, \quad \textit{for } 1 < p, q < \infty, \ \frac{1}{p} + \frac{1}{q} = 1; \tag{1.4.4}$$

$$(\mathbb{E}[|X+Y|^p])^{1/p} \leq (\mathbb{E}[|X|^p])^{1/p} + (\mathbb{E}[|Y|^p])^{1/p}, \quad \textit{for } p > 1. \tag{1.4.5}$$

In particular, for $1 \leq p < p'$, *we have the Lyapnov inequality:*

$$(\mathbb{E}[|X|^p])^{1/p} \leq (\mathbb{E}[|X|^{p'}])^{1/p'}.$$

Proof The inequality (1.4.3) follows from (1.4.1).

For the case $\mathbb{E}[|X|^p] \cdot \mathbb{E}[|Y|^q] > 0$, we set

$$\xi = \frac{X}{(\mathbb{E}[|X|^p])^{1/p}}, \eta = \frac{Y}{(\mathbb{E}[|Y|^q])^{1/q}}.$$

By (1.4.2) we have

$$\begin{aligned} \mathbb{E}[|\xi\eta|] &\leq \mathbb{E}[\frac{|\xi|^p}{p} + \frac{|\eta|^q}{q}] \leq \mathbb{E}[\frac{|\xi|^p}{p}] + \mathbb{E}[\frac{|\eta|^q}{q}] \\ &= \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

Thus (1.4.4) follows.

For the case $\mathbb{E}[|X|^p] \cdot \mathbb{E}[|Y|^q] = 0$, we consider $\mathbb{E}[|X|^p] + \varepsilon$ and $\mathbb{E}[|Y|^q] + \varepsilon$ for $\varepsilon > 0$. Applying the above method and letting $\varepsilon \to 0$, we get (1.4.4).

We now prove (1.4.5). We only consider the case $\mathbb{E}[|X+Y|^p] > 0$.

$$\begin{aligned} \mathbb{E}[|X+Y|^p] &= \mathbb{E}[|X+Y| \cdot |X+Y|^{p-1}] \\ &\leq \mathbb{E}[|X| \cdot |X+Y|^{p-1}] + \mathbb{E}[|Y| \cdot |X+Y|^{p-1}] \\ &\leq (\mathbb{E}[|X|^p])^{1/p} \cdot (\mathbb{E}[|X+Y|^{(p-1)q}])^{1/q} \\ &+ (\mathbb{E}[|Y|^p])^{1/p} \cdot (\mathbb{E}[|X+Y|^{(p-1)q}])^{1/q}. \end{aligned}$$

Since $(p-1)q = p$, we obtain (1.4.5).

By (1.4.4), it is easy to deduce that $(\mathbb{E}[|X|^p])^{1/p} \leq (\mathbb{E}[|X|^{p'}])^{1/p'}$ for $1 \leq p < p'$. □

For each fixed $p \geq 1$, we observe that $\mathcal{H}_0^p = \{X \in \mathcal{H}, \mathbb{E}[|X|^p] = 0\}$ is a linear subspace of $\mathcal{H}$. Taking $\mathcal{H}_0^p$ as our null space, we introduce the quotient space $\mathcal{H}/\mathcal{H}_0^p$. Observe that, for every $\{X\} \in \mathcal{H}/\mathcal{H}_0^p$ with a representation $X \in \mathcal{H}$, we can define an expectation $\mathbb{E}[\{X\}] := \mathbb{E}[X]$ which is still a sublinear expectation. We set $\|X\|_p := (\mathbb{E}[|X|^p])^{\frac{1}{p}}$. By Proposition 1.4.2, $\|\cdot\|_p$ defines a Banach norm on $\mathcal{H}/\mathcal{H}_0^p$. We extend $\mathcal{H}/\mathcal{H}_0^p$ to its completion $\hat{\mathcal{H}}_p$ under this norm, then $(\hat{\mathcal{H}}_p, \|\cdot\|_p)$ is a Banach space. In particular, when $p = 1$, we denote it by $(\hat{\mathcal{H}}, \|\cdot\|)$.

For each $X \in \mathcal{H}$, the mappings

$$X^+(\omega) : \mathcal{H} \mapsto \mathcal{H} \text{ and } X^-(\omega) : \mathcal{H} \mapsto \mathcal{H}$$

satisfy

$$|X^+ - Y^+| \leq |X - Y| \text{ and } |X^- - Y^-| = |(-X)^+ - (-Y)^+| \leq |X - Y|.$$

Thus they are both contraction mappings under $\|\cdot\|_p$ and can be continuously extended to the Banach space $(\hat{\mathcal{H}}_p, \|\cdot\|_p)$.

We can define the partial order "$\geq$" in this Banach space.

Definition 1.4.3 An element X in $(\hat{\mathcal{H}}, \|\cdot\|)$ is said to be nonnegative, or $X \geq 0$, $0 \leq X$, if $X = X^+$. We also write $X \geq Y$, or $Y \leq X$, if $X - Y \geq 0$.

It is easy to check that $X \geq Y$ and $Y \geq X$ imply $X = Y$ on $(\hat{\mathcal{H}}_p, \|\cdot\|_p)$.

For each $X, Y \in \mathcal{H}$, note that

$$|\mathbb{E}[X] - \mathbb{E}[Y]| \leq \mathbb{E}[|X - Y|] \leq ||X - Y||_p.$$

We discuss below further properties of sublinear expectations.

Definition 1.4.4 The sublinear expectation $\mathbb{E}[\cdot]$ can be continuously extended to $(\hat{\mathcal{H}}_p, \|\cdot\|_p)$, on which it is still a sublinear expectation. We still denote it by $(\Omega, \hat{\mathcal{H}}_p, \mathbb{E})$.

Let $(\Omega, \mathcal{H}, \mathbb{E}_1)$ be a nonlinear expectation spac1e. We say that $\mathbb{E}_1$ is dominated by $\mathbb{E}$, or $\mathbb{E}$ dominates $\mathbb{E}_1$, if

$$\mathbb{E}_1[X] - \mathbb{E}_1[Y] \leq \mathbb{E}[X - Y] \text{ for } X, Y \in \mathcal{H}.$$

From this we can easily deduce that $|\mathbb{E}_1[X] - \mathbb{E}_1[Y]| \leq \mathbb{E}[|X - Y|]$, thus the nonlinear expectation $\mathbb{E}_1[\cdot]$ can be continuously extended to $(\hat{\mathcal{H}}_p, \|\cdot\|_p)$, on which it is still a nonlinear expectation. We still denote it by $(\Omega, \hat{\mathcal{H}}_p, \mathbb{E}_1)$.

Remark 1.4.5 It is important to note that $X_1, \cdots, X_n \in \hat{\mathcal{H}}$ does not imply in general that $\varphi(X_1, \cdots, X_n) \in \hat{\mathcal{H}}$ for each $\varphi \in C_{l.Lip}(\mathbb{R}^n)$. Thus, when we talk about the notions of distributions, independence and product spaces on $(\Omega, \hat{\mathcal{H}}, \mathbb{E})$, the space $C_{b.Lip}(\mathbb{R}^n)$ cannot be replaced by $C_{l.Lip}(\mathbb{R}^n)$.

Remark 1.4.6 If the linear space $\mathcal{H}$ just satisfies that $|X| \in \mathcal{H}$ if $X \in \mathcal{H}$ and $\varphi(X) \in \mathcal{H}$ for each $X \in \mathcal{H}^n$, $\varphi \in C_{b.Lip}(\mathbb{R}^n)$, we could also get the Banach space $(\hat{\mathcal{H}}, \|\cdot\|_1)$.

1.5 Examples of i.i.d Sequences Under Uncertainty of Probabilities

We give three typical examples of i.i.d sequence under sublinear expectations.

Example 1.5.1 (Nonlinear version of Bernoulli random sequence) In an urn there are totally b_1 black balls and w_1 white balls such that $b_1 + w_1 = 100$. But what

I know about the number b_1 is only $b_1 \in [\underline{\mu}, \overline{\mu}]$ with a pair of fixed and known number $0 < \underline{\mu} < \overline{\mu} < 100$. Now we are allowed to completely randomly mix and then choose a ball from the urn, then get 1 dollar if the chosen ball is black, and -1 dollar if it is white. Our gain ξ_1 for this game is a random number

$$\xi_1(\omega) = \mathbf{1}_{\{\text{ black ball }\}} - \mathbf{1}_{\{\text{ white ball }\}}.$$

Now we repeat this game, but after each time i the random number ξ_i is output and a new game begins, the number of the balls b_{i+1} can be changed by our counterpart within the fixed range $[\underline{\mu}, \overline{\mu}]$ without informing us. What we can do is, again, to sufficiently mixed the 100 balls in the urn, then completely randomly choose a ball and thus get the random variable ξ_{i+1}. In this way a random sequence $\{\xi_i\}_{i=1}^{\infty}$ is then produced from this game.

Now at the starting time $i = 0$, if we sell a contract $\varphi(\xi_i)$ based on the ith output ξ_i, then, in considering the worst case scenario, the robust expectation is

$$\hat{\mathbb{E}}[\varphi(\xi_i)] = \hat{\mathbb{E}}[\varphi(\xi_1)] = \max_{p \in [\underline{\mu}, \overline{\mu}]} [p\varphi(1) + (1-p)\varphi(-1)], \quad i = 1, 2, \cdots,$$

It is clear that $\hat{\mathbb{E}}[\varphi(\xi_i)] = \hat{\mathbb{E}}[\varphi(\xi_j)]$ for each $i, j = 1, 2, \cdots$. Namely the sequence is identically distributed. We can also check that ξ_{i+1} is independent from $(\xi_1, \cdots, \xi_i)$. In general, if a path-dependent loss function is $X(\omega) = \varphi(\xi_1, \ldots, \xi_i)$, then the robust expected loss is:

$$\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[\varphi(\xi_1, \ldots, \xi_i, \xi_{i+1})] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x_1, \ldots, x_i, \xi_{i+1})]_{\{\xi_j = x_j, 1 \le j \le i\}}].$$

This sequence is a typical Bernoulli random sequence under uncertainty of probabilities, which is a more realistic simulation of of real world uncertainty.

Example 1.5.2 Let us consider another Bernoulli random sequence in which, at each time i, there are b_i black balls, w_i white balls and y_i yellow balls in the urn, satisfying $b_i = w_i$, $b_i + w_i + y_i = 100$ and $b_i \in [\underline{\mu}, \overline{\mu}]$. Here the bounds $0 < \underline{\mu} < \overline{\mu} < 50$ are known and fixed. The game is repeated similarly as for the last one with the following random output ξ_i at time i:

$$\xi_i(\omega) = \mathbf{1}_{\{\text{ black ball }\}} - \mathbf{1}_{\{\text{ white ball }\}}.$$

Namely, $\xi = 0$ if a yellow ball is chosen at the time i. We then have a different i.i.d sequence $\{\xi_i\}_{i=1}^{\infty}$:

$$\hat{\mathbb{E}}[\xi_i] = -\hat{\mathbb{E}}[-\xi_i] = 0.$$

We call this type of Bernoulli sequence under uncertainty a symmetric sequence.

Example 1.5.3 A very general situation is to replace the above urn to a generator of random vectors. At each time i this generator, follows a probability distribution $F_\theta(x)$, outputs a random vector ξ_i completely randomly. But the rule is that there is a given subset $\{F_\theta\}_{\theta\in\Theta}$ of probability distributions and, we do not know which one is chosen. Moreover, at the next time $i+1$, this generator can follow another completely different distribution $F_{\theta'}$ from the subset $\{F_\theta\}_{\theta\in\Theta}$. Similarly as in the first example, $\{\xi_i\}_{i=1}^\infty$ constitutes an i.i.d. sequence of random vectors with distribution uncertainty. We have

$$\hat{\mathbb{E}}[\varphi(\xi_i)] = \max_{\theta\in\Theta} \int_{\mathbb{R}^d} \varphi(x) F_\theta(dx).$$

1.6 Relation with Coherent Measures of Risk

Let the pair $(\Omega, \mathcal{H})$ be such that Ω is a set of scenarios and $\mathcal{H}$ is the collection of all possible risk positions in a financial market.

If $X \in \mathcal{H}$, then for each constant c, $X \vee c$, $X \wedge c$ are all in $\mathcal{H}$. One typical example in finance is that X is the tomorrow's price of a stock. In this case, any European call or put options with strike price K of the forms $(S-K)^+$, $(K-S)^+$, are in $\mathcal{H}$.

A risk supervisor is responsible for taking a rule to tell traders, securities companies, banks or other institutions under his supervision, which kind of risk positions is unacceptable and thus a minimum amount of risk capitals should be deposited in order to make the positions acceptable. The collection of acceptable positions is defined by

$$\mathcal{A} = \{X \in \mathcal{H} : X \text{ is acceptable}\}.$$

This set has meaningful properties in economy.

Definition 1.6.1 A set $\mathcal{A}$ is called a **coherent acceptable set** if it satisfies:
(i) Monotonicity:

$$X \in \mathcal{A},\ Y \geq X \ \text{ imply } \ Y \in \mathcal{A}.$$

(ii) $0 \in \mathcal{A}$ but $-1 \notin \mathcal{A}$.
(iii) Positive homogeneity:

$$X \in \mathcal{A} \ \text{ implies } \ \lambda X \in \mathcal{A} \ \text{ for } \ \lambda \geq 0.$$

(iv) Convexity:

$$X, Y \in \mathcal{A} \text{ imply } \alpha X + (1-\alpha)Y \in \mathcal{A} \ \text{ for } \alpha \in [0, 1].$$

Remark 1.6.2 Properties (iii) and (iv) imply
(v) Sublinearity:

$$X, Y \in \mathcal{A} \Rightarrow \mu X + \nu Y \in \mathcal{A} \ \text{ for constants } \ \mu, \nu \geq 0.$$

Remark 1.6.3 If the set $\mathcal{A}$ only satisfies (i), (ii) and (iv), then $\mathcal{A}$ is called a **convex acceptable set**.

In this section we mainly study the coherent case. Once the rule of the acceptable set is fixed, the minimum requirement of risk deposit is then automatically determined.

Definition 1.6.4 Given a coherent acceptable set $\mathcal{A}$, the functional $\rho(\cdot)$ defined by

$$\rho(X) = \rho_{\mathcal{A}}(X) := \inf\{m \in \mathbb{R} : \ m + X \in \mathcal{A}\}, \ X \in \mathcal{H}$$

is called the **coherent risk measure** related to $\mathcal{A}$.

It is easy to see that

$$\rho(X + \rho(X)) = 0.$$

Proposition 1.6.5 *A coherent risk measure $\rho(\cdot)$ satisfies the following four properties:*

(i) ***Monotonicity:*** *If $X \geq Y$, then $\rho(X) \leq \rho(Y)$.*
(ii) ***Constant preserving:*** $\rho(1) = -\rho(-1) = -1$.
(iii) ***Sub-additivity:*** *For each $X, Y \in \mathcal{H}$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$.*
(iv) ***Positive homogeneity:*** $\rho(\lambda X) = \lambda\rho(X)$*for* $\lambda \geq 0$.

Proof (i) and (ii) are obvious.

We now prove (iii). Indeed,

$$\begin{aligned}
\rho(X + Y) &= \inf\{m \in \mathbb{R} : \ m + (X + Y) \in \mathcal{A}\} \\
&= \inf\{m + n : m, n \in \mathbb{R}, \ (m + X) + (n + Y) \in \mathcal{A}\} \\
&\leq \inf\{m \in \mathbb{R} : \ m + X \in \mathcal{A}\} + \inf\{n \in \mathbb{R} : \ n + Y \in \mathcal{A}\} \\
&= \rho(X) + \rho(Y).
\end{aligned}$$

The case $\lambda = 0$ for (iv) is trivial; when $\lambda > 0$,

$$\begin{aligned}
\rho(\lambda X) &= \inf\{m \in \mathbb{R} : \ m + \lambda X \in \mathcal{A}\} \\
&= \lambda \inf\{n \in \mathbb{R} : \ n + X \in \mathcal{A}\} = \lambda\rho(X),
\end{aligned}$$

where $n = m/\lambda$. □

Obviously, if $\mathbb{E}$ is a sublinear expectation, we define $\rho(X) := \mathbb{E}[-X]$, then ρ is a coherent risk measure. Conversely, if ρ is a coherent risk measure, we define $\mathbb{E}[X] := \rho(-X)$, then $\mathbb{E}$ is a sublinear expectation.

1.7 Exercises

Exercise 1.7.1 Prove that a functional $\mathbb{E}$ is sublinear if and only if it satisfies convexity and positive homogeneity.

Exercise 1.7.2 Suppose that all elements in $\mathcal{H}$ are bounded. Prove that the strongest sublinear expectation on $\mathcal{H}$ is

$$\mathbb{E}^{\infty}[X] := X^* = \sup_{\omega \in \Omega} X(\omega).$$

Namely, all other sublinear expectations are dominated by $\mathbb{E}^{\infty}[\cdot]$.

Exercise 1.7.3 Suppose that the sublinear expectation $\mathbb{E}$ is given by

$$\mathbb{E}[X] = \max_{\theta \in \Theta} E_\theta[X], \forall X \in \mathcal{H},$$

where E_θ is a family of linear expectations. Prove that $\mathbb{E}[X] = \mathbb{E}[-X] = 0$ if and only if $E_\theta[X] = 0$ for each $\theta \in \Theta$.

Exercise 1.7.4 Suppose that $\mathbb{L}^{\infty}(\Omega, \mathcal{F})$ is the collection of all bounded random variables on a measurable space $(\Omega, \mathcal{F})$. Given a finitely additive probability measure Q on $(\Omega, \mathcal{F})$, define

$$E_Q[X] := \sum_{i=1}^{N} x_i Q(A_i), \quad \text{for all } X \in \mathbb{L}_0^{\infty}(\Omega, \mathcal{F}).$$

Here $\mathbb{L}_0^{\infty}(\Omega, \mathcal{F})$ is the collection of all random variables X of the form

$$X(\omega) = \sum_{i=1}^{N} x_i \mathbf{1}_{A_i}(\omega), \ x_i \in \mathbb{R}, \ A_i \in \mathcal{F}, i = 1, \cdots, N.$$

Show that:

(i) $E_Q : \mathbb{L}_0^{\infty}(\Omega, \mathcal{F}) \mapsto \mathbb{R}$ is a linear expectation.
(ii) E_Q is continuous under the norm $\|\cdot\|_\infty$ and it can be extended from $\mathbb{L}_0^{\infty}(\Omega, \mathcal{F})$ to a linear continuous functional on $\mathbb{L}^{\infty}(\Omega, \mathcal{F})$.

Furthermore, suppose that $E : \mathbb{L}^{\infty}(\Omega, \mathcal{F}) \mapsto \mathbb{R}$ is a linear expectation. Prove that there exists a finitely additive probability measure Q on $(\Omega, \mathcal{F})$ such that $E = E_Q$.

Exercise 1.7.5 Suppose $X, Y \in \mathcal{H}^d$ and Y is an independent copy of X. Prove that, for each $a \in \mathbb{R}$ and $b \in \mathbb{R}^d$, $a + \langle b, Y\rangle$ is an independent copy of $a + \langle b, X\rangle$.

Exercise 1.7.6 Prove that the space $C_{l.Lip}(\mathbb{R}^{m+n})$ can be replaced by the space $C_{b.Lip}(\mathbb{R}^{m+n})$ in Definition 1.3.11.

Exercise 1.7.7 Let $(\Omega, \mathcal{H}, \mathbb{E})$ be a sublinear expectation space and X_i be in $\mathcal{H}$, $i = 1, 2, 3, 4$.

(i) Suppose that X_2 is independent of X_1 and X_3 is independent of (X_1, X_2). Prove that (X_2, X_3) is independent of X_1.

(ii) Suppose that $X_1 \overset{d}{=} X_2$ and $X_3 \overset{d}{=} X_4$. Prove that if X_3 is independent of X_1 and X_4 is independent of X_2, then $X_1 + X_3 \overset{d}{=} X_2 + X_4$.

Exercise 1.7.8 Suppose that $\mathbb{E}$ is a linear expectation on $(\Omega, \mathcal{H})$ and $X, Y \in \mathcal{H}$. Show that X is independent of Y if and only if Y is independent of X.

Exercise 1.7.9 Let X and Y be two non-constant random variables in a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ and satisfy $\mathbb{E}[X] \neq -\mathbb{E}[-X]$. Suppose that X is independent of Y and Y is independent of X. Show that:

(i) $\mathbb{E}[(\varphi(Y) - \mathbb{E}[\varphi(Y)])^+] = 0$ for each $\varphi \in C_{b.Lip}(\mathbb{R})$.

(ii) there exists a closed subset B of $\mathbb{R}$ such that

$$\mathbb{E}[(\phi(Y)] = \sup_{y \in B} \phi(y), \quad \text{for all } \phi \in C_{b.Lip}(\mathbb{R}).$$

Exercise 1.7.10 Prove that the inequalities (1.4.3), (1.4.4), (1.4.5) still hold for $(\Omega, \hat{\mathcal{H}}, \mathbb{E})$.

Exercise 1.7.11 Suppose that $X \in \hat{\mathcal{H}}_p$ satisfies $\lim_{\lambda \to \infty} \mathbb{E}[(|X|^p - \lambda)^+] = 0$. Prove that $\varphi(X) \in \hat{\mathcal{H}}_p$ for each $\varphi \in C(\mathbb{R})$ satisfying $|\varphi(x)| \leq C(1 + |x|^p)$.

Exercise 1.7.12 Suppose that $X \in \hat{\mathcal{H}}$ is bounded. Prove that $XY \in \hat{\mathcal{H}}$, for each $Y \in \hat{\mathcal{H}}$ satisfying that $\lim_{\lambda \to \infty} \mathbb{E}[(|Y| - \lambda)^+] = 0$.

Exercise 1.7.13 Suppose that $\{X_n\}_{n=1}^{\infty}$ converges to X and $\{Y_n\}_{n=1}^{\infty}$ converges to Y under $\|\cdot\|_p$ for $p \geq 1$.

(i) Prove that if $Y_n \overset{d}{=} X_n$ for each n, then $Y \overset{d}{=} X$.

(ii) Prove that if Y_n is independent of X_n for each n, then Y is independent from X.

Exercise 1.7.14 Let $\rho(\cdot)$ be a coherent risk measure. We define

$$\mathcal{A}_\rho := \{X \in \mathcal{H} : \rho(X) \leq 0\}.$$

Prove that $\mathcal{A}_\rho$ is a coherent acceptable set.

Notes and Comments

The sublinear expectation is also called the upper expectation (see Huber [88] in robust statistics), or the upper prevision in the theory of imprecise probabilities (see Walley [172] and a rich literature provided in the Notes of this book). To our knowledge, the Representation Theorem 1.2.1 was firstly obtained in [88] for the case where Ω is a finite set, and this theorem was rediscovered independently by Heath Artzner, Delbaen, Eber and Heath [3] and then by Delbaen [45] for a general Ω. A typical dynamic nonlinear expectation is the so-called g–expectation (small g), which was introduced in the the years of 90s, see e.g., [130] in the framework of backward stochastic differential equations. For the further development of this theory and it's applications, readers are referred to Briand et al. [20], Chen [26], Chen and Epstein [28], Chen et al. [29], Chen and Peng [30, 31], Coquet et al. [35, 36], Jiang [96], Jiang and Chen [97, 98], Peng [132, 135], Peng and Xu [148] and Rosazza-Gianin [152]. It seems that the notions of distributions and independence under nonlinear expectations are new. We believe that these notions are perfectly adapted for the further development of dynamic nonlinear expectations. For other types of the related notions of distributions and independence under nonlinear expectations or non-additive probabilities, we refer to the Notes of the book [172] and the references listed in Marinacci [115], Maccheroni and Marinacci [116]. Coherent risk measures can also be regarded as sublinear expectations defined on the space of risk positions in financial market. This notion was firstly introduced in [3]. Readers can also be referred to the well-known book of Föllmer and Schied [69] for the systematical presentation of coherent risk measures and convex risk measures. For the dynamic risk measure in continuous time, see [135] or [152], Barrieu and El Karoui [10] using g-expectations. The notion of super-hedging and super pricing (see El Karoui and Quenez [57] and El Karoui, Peng and Quenez [58]) are also closely related to this formulation.

Chapter 2
Law of Large Numbers and Central Limit Theorem Under Probability Uncertainty

In this chapter, we first introduce two types of fundamentally important distributions, namely, maximal distribution and a new type of nonlinear normal distribution—G-normal distribution in the theory of sublinear expectations. The former corresponds to constants and the latter corresponds to normal distribution in the classical probability theory. We then present the law of large numbers (LLN) and central limit theorem (CLT) under sublinear expectations. It is worth pointing out that the limit in LLN is a maximal distribution and the limit in CLT is a G-normal distribution.

2.1 Some Basic Results of Parabolic Partial Differential Equations

We recall some basic results from parabolic partial differential equations (PDEs), defined on the time-space $[0, T] \times \mathbb{R}^d$, of the following type:

$$\partial_t u(t,x) - G(Du(t,x), D^2u(t,x)) = 0, \quad t \in (0,T), \tag{2.1.1}$$

$$u(0,x) = \varphi(x) \quad \text{for } x \in \mathbb{R}^d, \tag{2.1.2}$$

where $D := (\partial_{x_i})_{i=1}^d$, $D^2 := (\partial^2_{x_i x_j})_{i,j=1}^d$ and $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$, $\varphi \in C(\mathbb{R}^d)$ are given functions.

Let Q be a subset of $[0, \infty) \times \mathbb{R}^d$. We denote by $C(Q)$ all continuous functions u defined on Q, in the relative topology on Q, with a finite norm

$$\|u\|_{C(Q)} = \sup_{(t,x)\in Q} |u(t,x)|.$$

Given $\alpha, \beta \in (0, 1)$, let $C^{\alpha,\beta}(Q)$ be the set of functions in $C(Q)$ such that following norm is finite:

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_2

$$\|u\|_{C^{\alpha,\beta}(Q)} = \|u\|_{C(Q)} + \sup_{(t,x),(s,y)\in Q,(t,x)\neq(s,y)} \frac{|u(t,x)-u(t,y)|}{|s-t|^{\alpha}+|x-y|^{\beta}}.$$

We also introduce the norms

$$\begin{aligned}\|u\|_{C^{1,1}(Q)} =& \|u\|_{C(Q)} + \|\partial_t u\|_{C(Q)} + \sum_{i=1}^{d} \|\partial_{x_i} u\|_{C(Q)},\\ \|u\|_{C^{1,2}(Q)} =& \|u\|_{C^{1,1}(Q)} + \sum_{i,j=1}^{d} \|\partial_{x_i x_j} u\|_{C(Q)},\end{aligned}$$

and

$$\begin{aligned}\|u\|_{C^{1+\alpha,1+\beta}(Q)} =& \|u\|_{C^{\alpha,\beta}(Q)} + \|\partial_t u\|_{C^{\alpha,\beta}(Q)} + \sum_{i=1}^{d} \|\partial_{x_i} u\|_{C^{\alpha,\beta}(Q)},\\ \|u\|_{C^{1+\alpha,2+\beta}(Q)} =& \|u\|_{C^{1+\alpha,1+\beta}(Q)} + \sum_{i,j=1}^{d} \|\partial_{x_i x_j} u\|_{C^{\alpha,\beta}(Q)}.\end{aligned}$$

The corresponding subspaces of $C(Q)$ in which the correspondent derivatives exist and the above norms are finite are denoted respectively by

$$C^{1,1}(Q), \quad C^{1,2}(Q), \quad C^{1+\alpha,1+\beta}(Q) \text{ and } \quad C^{1+\alpha,2+\beta}(Q).$$

We always suppose that the given function $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ is continuous and satisfies the following degenerate ellipticity condition:

$$G(p, A) \leq G(p, A') \quad \text{whenever} \quad A \leq A'. \tag{2.1.3}$$

In many situations, we assume that G is a sublinear function, i.e., it satisfies

$$G(p, A) - G(p', A') \leq G(p - p', A - A'), \qquad \text{for all } p, p' \in \mathbb{R}^d, \quad A, A' \in \mathbb{S}(d). \tag{2.1.4}$$

and

$$G(\alpha p, \alpha A) = \alpha G(p, A), \qquad \text{for all } \alpha \geq 0, \quad p \in \mathbb{R}^d, \quad A \in \mathbb{S}(d). \tag{2.1.5}$$

Sometimes we need the following strong ellipticity condition: there exists a constant $\lambda > 0$ such that

$$G(p, A) - G(p, \bar{A}) \geq \lambda \mathrm{tr}[A - \bar{A}], \quad \text{for all } p \in \mathbb{R}^d, \;\; A, A' \in \mathbb{S}(d) \tag{2.1.6}$$

Definition 2.1.1 A viscosity subsolution of (2.1.1) defined on $(0, T) \times \mathbb{R}^d$, is a function $u \in USC((0, T) \times \mathbb{R}^d)$ such that for all $(t, x) \in (0, T) \times \mathbb{R}^d$, $\phi \in C^2((0, T) \times \mathbb{R}^d)$ with $u(t, x) = \phi(t, x)$ and $u < \phi$ on $(0, T) \times \mathbb{R}^d \backslash (t, x)$, we have

$$\partial_t \phi(t, x) - G(D\phi(t, x), D^2\phi(t, x)) \leq 0.$$

Likewise, a viscosity supersolution of (2.1.1) defined on $(0, T) \times \mathbb{R}^d$ is a function $v \in LSC((0, T) \times \mathbb{R}^d)$ such that for all $(t, x) \in (0, T) \times \mathbb{R}^d$, $\phi \in C^2((0, T) \times \mathbb{R}^d)$ with $u(t, x) = \phi(t, x)$ and $u > \phi$ on $(0, T) \times \mathbb{R}^d \backslash (t, x)$, we have

$$\partial_t \phi(t, x) - G(D\phi(t, x), D^2\phi(t, x)) \geq 0.$$

A viscosity solution of (2.1.1) defined on $(0, T) \times \mathbb{R}^d$ is a function which is simultaneously a viscosity subsolution and a viscosity supersolution of (2.1.1) on $(0, T) \times \mathbb{R}^d$.

Theorem 2.1.2 *Suppose (2.1.3), (2.1.4) and (2.1.5) hold. Then for any* $\varphi \in C_{l.Lip}(\mathbb{R}^d)$*, there exists a unique* $u \in C([0, \infty) \times \mathbb{R}^d)$*, bounded by*

$$|u(t, x) - u(t, \bar{x})| \leq C(1 + |x|^k + |\bar{x}|^k)(|x - \bar{x}|) \tag{2.1.7}$$

and

$$|u(t, x) - u(t + s, x)| \leq C(1 + |x|^k)(|s| + |s|^{1/2}), \tag{2.1.8}$$

and satisfying initial condition (2.1.2) such that, u *is a viscosity solution of the PDE (2.1.1) on* $(0, T) \times \mathbb{R}^d$ *for each* $T > 0$*. Moreover, setting* $u^\varphi(t, x) = u(t, x)$ *to indicate its dependence of the initial condition* $u(0, \cdot) = \varphi(\cdot)$*, we have, for each* $(t, x) \in [0, \infty) \times \mathbb{R}^d$*,*

$$u^\varphi(t, x) \leq u^\psi(t, x), \quad \text{if } \varphi \leq \psi; \tag{2.1.9}$$

$$u^c(t, x) \equiv c, \quad \text{for each constant } c; \tag{2.1.10}$$

$$u^{\alpha\varphi}(t, x) = \alpha u^\varphi(t, x), \quad \text{for all constant } \alpha \geq 0; \tag{2.1.11}$$

$$u^\varphi(t, x) - u^\psi(t, x) \leq u^{\varphi - \psi}(t, x), \quad \text{for each } \varphi, \psi \in C_b(\mathbb{R}^d). \tag{2.1.12}$$

If moreover, the strong ellipticity condition (2.1.6) holds and the initial condition φ *is uniformly bounded, then for each* $0 < \kappa < T$*, there is a number* $\alpha \in (0, 1)$ *such that* $u \in C^{1+\alpha/2, 2+\alpha}([\kappa, T] \times \mathbb{R}^d)$*, namely,*

$$\|u\|_{u \in C^{1+\alpha/2, 2+\alpha}([\kappa, T] \times \mathbb{R}^d)} < \infty. \tag{2.1.13}$$

Proof The proofs of (2.1.7)–(2.1.12) are provided in Appendix C, Theorems C.2.5, C.2.6, C.2.8 and C.3.4. The property of smoothness (2.1.13) is due to Krylov (see Appendix C, Theorem C.4.5 for details). □

Remark 2.1.3 It is easy to check that, if $u \in C^{1,2}(0, T)$, then u is a viscosity solution of (2.1.1)–(2.1.2) if and only if u satisfies

$$\partial_t u(t, x) - G(Du(t, x), D^2u(t, x)) = 0, \quad \text{for each} \ \ t \in (0, T), \ \ x \in \mathbb{R}^d.$$

In this book we will mainly use the notion of viscosity solution to describe the solution of this PDE. For reader's convenience, we give a systematic introduction of the notion of viscosity solution and its related properties used in this book (see Appendix C, Sect. C1–C3). It is worth to mention here that for G satisfying strongly elliptic condition (2.1.6), the viscosity solution of the PDE (2.1.1) with initial condition (2.1.2) becomes a classical $C^{1,2}$-solution. Readers without knowledge of viscosity solutions can simply understand solutions of this PDE in the classical sense.

2.2 Maximal Distribution and *G*-Normal Distribution

Let us define a special type of very simple distributions which are frequently used in practice, known as "worst case risk measure".

Definition 2.2.1 (*maximal distribution*) A d-dimensional random vector $\eta = (\eta_1, \cdots, \eta_d)$ on a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ is called **maximally distributed** if there exists a bounded, closed and convex subset $\Gamma \subset \mathbb{R}^d$ such that

$$\mathbb{E}[\varphi(\eta)] = \max_{y \in \Gamma} \varphi(y), \quad \varphi \in C_{l.Lip}(\mathbb{R}^d).$$

Remark 2.2.2 Here Γ gives the degree of uncertainty of η. It is easy to check that this maximally distributed random vector η satisfies the relation

$$a\eta + b\bar{\eta} \overset{d}{=} (a + b)\eta \quad \text{for } a, b \geq 0,$$

where $\bar{\eta}$ is an independent copy of η. We will see later that in fact this relation characterizes a maximal distribution. Maximal distribution is also called "worst case risk measure" in finance.

Remark 2.2.3 When $d = 1$ we have $\Gamma = [\underline{\mu}, \overline{\mu}]$, where $\overline{\mu} = \mathbb{E}[\eta]$ and $\underline{\mu} = -\mathbb{E}[-\eta]$. The distribution of η is

$$\mathbb{F}_\eta[\varphi] = \mathbb{E}[\varphi(\eta)] = \sup_{\underline{\mu} \leq y \leq \bar{\mu}} \varphi(y) \quad \text{for } \varphi \in C_{l.Lip}(\mathbb{R}).$$

Recall a well-known classical characterization: $X \overset{d}{=} N(0, \Sigma)$ if and only if

$$aX + b\bar{X} \overset{d}{=} \sqrt{a^2 + b^2}X \ \ \text{for } a, b \geq 0, \tag{2.2.1}$$

where $\bar{X}$ is an independent copy of X. The covariance matrix Σ is defined by $\Sigma = E[XX^T]$. We will see that, within the framework of sublinear distributions, this normal distribution is just a special type of normal distributions. Let us consider the so-called G-normal distribution in probability model under uncertainty situation. The existence, uniqueness and characterization will be given later.

Definition 2.2.4 (G*-normal distribution*) We say that a d-dimensional random vector $X = (X_1, \cdots, X_d)$ on a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ is called (centered) G**-normally distributed** if

$$aX + b\bar{X} \overset{d}{=} \sqrt{a^2 + b^2}X \quad \text{for } a, b \geq 0,$$

where $\bar{X}$ is an independent copy of X.

Remark 2.2.5 Noting that, for G-normally distributed X, $\mathbb{E}[X + \bar{X}] = 2\mathbb{E}[X]$ and $\mathbb{E}[X + \bar{X}] = \mathbb{E}[\sqrt{2}X] = \sqrt{2}\mathbb{E}[X]$, we then have $\mathbb{E}[X] = 0$. Similarly, we can prove that $\mathbb{E}[-X] = 0$. Therefore such X has no mean-uncertainty.

The following property is easy to be proved by the definition.

Proposition 2.2.6 *Let X be G-normally distributed. Then for each $A \in \mathbb{R}^{m\times d}$, AX is also G-normally distributed. In particular, for each* $\mathbf{a} \in \mathbb{R}^d$, $\langle \mathbf{a}, X\rangle$ *is a* 1*-dimensional G-normally distributed random variable. The converse is not true in general (see Exercise 2.5.1).*

We denote by $\mathbb{S}(d)$ the collection of all $d \times d$ symmetric matrices. Let X be G-normally distributed and η be maximally distributed d-dimensional random vectors on $(\Omega, \mathcal{H}, \mathbb{E})$. The following function is basically important to characterize their distributions:

$$G(p, A) := \mathbb{E}\left[\tfrac{1}{2}\langle AX, X\rangle + \langle p, \eta\rangle\right], \quad (p, A) \in \mathbb{R}^d \times \mathbb{S}(d). \tag{2.2.2}$$

It is easy to check that G is a sublinear function, monotone in $A \in \mathbb{S}(d)$ in the following sense: for each $p, \bar{p} \in \mathbb{R}^d$ and $A, \bar{A} \in \mathbb{S}(d)$

$$\begin{cases} G(p + \bar{p}, A + \bar{A}) \leq G(p, A) + G(\bar{p}, \bar{A}), \\ G(\lambda p, \lambda A) = \lambda G(p, A), \quad \forall \lambda \geq 0, \\ G(p, A) \leq G(p, \bar{A}), \quad \text{if } A \leq \bar{A}. \end{cases} \tag{2.2.3}$$

Clearly, G is also a continuous function. By Theorem 1.2.1 in Chap. 1, there exists a bounded and closed subset $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^{d\times d}$ such that

$$G(p, A) = \sup_{(q,Q)\in\Gamma} \left[\tfrac{1}{2}\mathrm{tr}[AQQ^T] + \langle p, q\rangle\right] \quad \text{for } (p, A) \in \mathbb{R}^d \times \mathbb{S}(d). \tag{2.2.4}$$

In Sects. 1.1–2.1 some main properties of such special type of parabolic PDE are provided in Theorem 2.1.2.

We have the following result, which will be proved in the next section.

Proposition 2.2.7 *Let $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a given sublinear and continuous function, monotone in $A \in \mathbb{S}(d)$ in the sense of (2.2.3). Then there exists a G-normally distributed d-dimensional random vector X and a maximally distributed d-dimensional random vector η on some sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ satisfying (2.2.2) and*

$$(aX + b\bar{X}, a^2\eta + b^2\bar{\eta}) \stackrel{d}{=} (\sqrt{a^2+b^2}X, (a^2+b^2)\eta), \quad \text{for } a, b \geq 0, \tag{2.2.5}$$

where $(\bar{X}, \bar{\eta})$ is an independent copy of (X, η).

Definition 2.2.8 The pair (X, η) satisfying (2.2.5) is called G-**distributed** associated to the function G given in (2.2.2).

Remark 2.2.9 In fact, if the pair (X, η) satisfies (2.2.5), then

$$aX + b\bar{X} \stackrel{d}{=} \sqrt{a^2+b^2}X, \ a\eta + b\bar{\eta} \stackrel{d}{=} (a+b)\eta \ \text{ for } a, b \geq 0.$$

Thus X is G-normally distributed and η is maximally distributed.

The above pair (X, η) is characterized by the following parabolic partial differential equation (PDE for short) defined on $[0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$:

$$\partial_t u - G(D_y u, D_x^2 u) = 0, \tag{2.2.6}$$

with Cauchy condition $u|_{t=0} = \varphi$, where $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ is defined by (2.2.2) and $D_x^2 u = (\partial^2_{x_i x_j} u)^d_{i,j=1}$, $D_x u = (\partial_{x_i} u)^d_{i=1}$. The PDE (2.2.6) is called a G**-equation**.

Proposition 2.2.10 *Assume that the pair (X, η) satisfies (2.2.5). For any given function $\varphi \in C_{l.Lip}(\mathbb{R}^d \times \mathbb{R}^d)$, we define*

$$u(t, x, y) := \mathbb{E}[\varphi(x + \sqrt{t}X, y + t\eta)], \ (t, x, y) \in [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d.$$

Then we have

$$u(t+s, x, y) = \mathbb{E}[u(t, x + \sqrt{s}X, y + s\eta)], \quad s \geq 0. \tag{2.2.7}$$

We also have the estimates: for each $T > 0$, there exist constants $C, k > 0$ such that, for all $t, s \in [0, T]$ and $x, \bar{x}, y, \bar{y} \in \mathbb{R}^d$,

$$|u(t, x, y) - u(t, \bar{x}, \bar{y})| \leq C(1 + |x|^k + |\bar{x}|^k + |y|^k + |\bar{y}|^k)(|x - \bar{x}| + |y - \bar{y}|) \tag{2.2.8}$$

and

$$|u(t, x, y) - u(t+s, x, y)| \leq C(1 + |x|^k + |y|^k)(s + |s|^{1/2}). \tag{2.2.9}$$

Moreover, u is the unique viscosity solution, continuous in the sense of (2.2.8) and (2.2.9), of the PDE (2.2.6).

Proof Since $\varphi \in C_{l.Lip}(\mathbb{R}^n)$, we can find some constants C, C_1 and k so that

$$\begin{aligned} u(t,x,y) - u(t,\bar{x},\bar{y}) &= \mathbb{E}[\varphi(x+\sqrt{t}X, y+t\eta)] - \mathbb{E}[\varphi(\bar{x}+\sqrt{t}X, \bar{y}+t\eta)] \\ &\leq \mathbb{E}[\varphi(x+\sqrt{t}X, y+t\eta) - \varphi(\bar{x}+\sqrt{t}X, \bar{y}+t\eta)] \\ &\leq \mathbb{E}[C_1(1+|X|^k+|\eta|^k+|x|^k+|y|^k+|\bar{x}|^k+|\bar{y}|^k)] \\ &\quad \times (|x-\bar{x}|+|y-\bar{y}|) \\ &\leq C(1+|x|^k+|y|^k+|\bar{x}|^k+|\bar{y}|^k)(|x-\bar{x}|+|y-\bar{y}|). \end{aligned}$$

This is (2.2.8).

Let $(\bar{X}, \bar{\eta})$ be an independent copy of (X, η). By (2.2.5),

$$\begin{aligned} u(t+s,x,y) &= \mathbb{E}[\varphi(x+\sqrt{t+s}X, y+(t+s)\eta)] \\ &= \mathbb{E}[\varphi(x+\sqrt{s}X+\sqrt{t}\bar{X}, y+s\eta+t\bar{\eta})] \\ &= \mathbb{E}[\mathbb{E}[\varphi(x+\sqrt{s}\widetilde{x}+\sqrt{t}\bar{X}, y+s\widetilde{y}+t\bar{\eta})]_{(\widetilde{x},\widetilde{y})=(X,\eta)}] \\ &= \mathbb{E}[u(t, x+\sqrt{s}X, y+s\eta)], \end{aligned}$$

we thus obtain (2.2.7). From this and (2.2.8) it follows that

$$\begin{aligned} &u(t+s,x,y) - u(t,x,y) = \mathbb{E}[u(t,x+\sqrt{s}X, y+s\eta) - u(t,x,y)] \\ &\leq \mathbb{E}[C_1(1+|x|^k+|y|^k+|X|^k+|\eta|^k)(\sqrt{s}|X|+s|\eta|)], \end{aligned}$$

thus we conclude (2.2.9).

Now, for a fixed $(t,x,y) \in (0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$, let $\psi \in C^{2,3}_{l.Lip}([0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d)$ be such that $\psi \geq u$ and $\psi(t,x,y) = u(t,x,y)$. By (2.2.7) and Taylor's expansion, it follows that, for $\delta \in (0,t)$,

$$\begin{aligned} 0 &\leq \mathbb{E}[\psi(t-\delta, x+\sqrt{\delta}X, y+\delta\eta) - \psi(t,x,y)] \\ &\leq \bar{C}(1+|x|^m+|y|^m)(\delta^{3/2}+\delta^2) - \partial_t\psi(t,x,y)\delta \\ &\quad + \mathbb{E}[\langle D_x\psi(t,x,y), X\rangle\sqrt{\delta} + \langle D_y\psi(t,x,y),\eta\rangle\delta + \tfrac{1}{2}\langle D_x^2\psi(t,x,y)X, X\rangle\delta] \\ &= -\partial_t\psi(t,x,y)\delta + \mathbb{E}[\langle D_y\psi(t,x,y),\eta\rangle + \tfrac{1}{2}\langle D_x^2\psi(t,x,y)X,X\rangle]\delta \\ &\qquad + \bar{C}(1+|x|^m+|y|^m)(\delta^{3/2}+\delta^2) \\ &= -\partial_t\psi(t,x,y)\delta + \delta G(D_y\psi, D_x^2\psi)(t,x,y) + \bar{C}(1+|x|^m+|y|^m)(\delta^{3/2}+\delta^2), \end{aligned}$$

where the constants $\bar{C}$ and m depend on ψ. Consequently, it is easy to check that

$$[\partial_t\psi - G(D_y\psi, D_x^2\psi)](t,x,y) \leq 0.$$

Thus u is a viscosity subsolution of (2.2.6). Similarly we can prove that u is a viscosity supersolution of (2.2.6). □

Remark 2.2.11 Note that in Proposition 2.2.10, we assume for convenience that, all moments of (X, η) exist. In fact, this condition can be weakened, see Exercise 2.5.4.

Corollary 2.2.12 *If both (X, η) and $(\bar{X}, \bar{\eta})$ satisfy (2.2.5) with the same G, i.e.,*

$$G(p, A) = \mathbb{E}\left[\frac{1}{2}\langle AX, X\rangle + \langle p, \eta\rangle\right] = \mathbb{E}\left[\frac{1}{2}\langle A\bar{X}, \bar{X}\rangle + \langle p, \bar{\eta}\rangle\right] \quad \textit{for } (p, A) \in \mathbb{R}^d \times \mathbb{S}(d),$$

then $(X, \eta) \overset{d}{=} (\bar{X}, \bar{\eta})$. In particular, $X \overset{d}{=} -X$.

Proof For each $\varphi \in C_{l.Lip}(\mathbb{R}^d \times \mathbb{R}^d)$, we set

$$\begin{aligned} u(t, x, y) &:= \mathbb{E}[\varphi(x + \sqrt{t}X, y + t\eta)], \\ \bar{u}(t, x, y) &:= \mathbb{E}[\varphi(x + \sqrt{t}\bar{X}, y + t\bar{\eta})], \ (t, x, y) \in [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d. \end{aligned}$$

By Proposition 2.2.10, both u and $\bar{u}$ are viscosity solutions of the G-equation (2.2.6) with the same Cauchy condition $u|_{t=0} = \bar{u}|_{t=0} = \varphi$. It follows from the uniqueness of the viscosity solution that $u \equiv \bar{u}$. In particular,

$$\mathbb{E}[\varphi(X, \eta)] = \mathbb{E}[\varphi(\bar{X}, \bar{\eta})].$$

Thus $(X, \eta) \overset{d}{=} (\bar{X}, \bar{\eta})$. □

Corollary 2.2.13 *Let (X, η) satisfy (2.2.5). For each $\varphi \in C_{l.Lip}(\mathbb{R}^d)$ we define*

$$v(t, x) := \mathbb{E}[\varphi(x + \sqrt{t}X + t\eta)], \ (t, x) \in [0, \infty) \times \mathbb{R}^d. \tag{2.2.10}$$

Then v is the unique viscosity solution of the following parabolic PDE:

$$\partial_t v - G(D_x v, D_x^2 v) = 0, \quad v|_{t=0} = \varphi. \tag{2.2.11}$$

Moreover, we have $v(t, x + y) \equiv u(t, x, y)$, where u is the solution of the PDE (2.2.6) with initial condition $u(t, x, y)|_{t=0} = \varphi(x + y)$.

Example 2.2.14 Let X be G-normally distributed. The distribution of X is characterized by the function

$$u(t, x) = \mathbb{E}[\varphi(x + \sqrt{t}X)], \ \ \varphi \in C_{l.Lip}(\mathbb{R}^d).$$

In particular, $\mathbb{E}[\varphi(X)] = u(1, 0)$, where u is the solution of the following parabolic PDE defined on $[0, \infty) \times \mathbb{R}^d$:

$$\partial_t u - G(D^2 u) = 0, \ \ u|_{t=0} = \varphi, \tag{2.2.12}$$

where $G = G_X(A) : \mathbb{S}(d) \mapsto \mathbb{R}$ is defined by

$$G(A) := \tfrac{1}{2}\mathbb{E}[\langle AX, X\rangle], \quad A \in \mathbb{S}(d).$$

The parabolic PDE (2.2.12) is called a ***G*-heat equation**.

It is easy to check that G is a sublinear function defined on $\mathbb{S}(d)$. By Theorem 1.2.1 in Chap. 1, there exists a bounded, convex and closed subset $\Theta \subset \mathbb{S}(d)$ such that

$$\tfrac{1}{2}\mathbb{E}[\langle AX, X\rangle] = G(A) = \tfrac{1}{2}\sup_{Q\in\Theta} \mathrm{tr}[AQ], \quad A \in \mathbb{S}(d). \tag{2.2.13}$$

Since $G(A)$ is monotone: $G(A_1) \geq G(A_2)$, for $A_1 \geq A_2$, it follows that

$$\Theta \subset \mathbb{S}_+(d) = \{\theta \in \mathbb{S}(d) : \theta \geq 0\} = \left\{BB^T : B \in \mathbb{R}^{d\times d}\right\}.$$

Here $\mathbb{R}^{d\times d}$ is the set of all $d \times d$ matrices. If Θ is a singleton: $\Theta = \{Q\}$, then X is classical zero-mean normally distributed with covariance Q. In general, Θ characterizes the covariance uncertainty of X. We denote $X \overset{d}{=} N(\{0\} \times \Theta)$ (Recall Eq. (2.2.4), we can set $(q, Q) \in \{0\} \times \Theta$).

When $d = 1$, we have $X \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$ (We also denote by $X \overset{d}{=} N(0, [\underline{\sigma}^2, \overline{\sigma}^2])$), where $\overline{\sigma}^2 = \mathbb{E}[X^2]$ and $\underline{\sigma}^2 = -\mathbb{E}[-X^2]$. The corresponding G-heat Eq. (2.2.12) becomes

$$\partial_t u - \tfrac{1}{2}(\overline{\sigma}^2(\partial^2_{xx}u)^+ - \underline{\sigma}^2(\partial^2_{xx}u)^-) = 0, \; u|_{t=0} = \varphi. \tag{2.2.14}$$

In the case $\underline{\sigma}^2 > 0$, this equation is also called the Barenblatt equation.

In the following two typical situations, the calculation of $\mathbb{E}[\varphi(X)]$ is quite easy.

Proposition 2.2.15 *Let $X \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$. Then, for each convex (resp. concave) function φ in $C_{l.Lip}(\mathbb{R})$, we have*

$$\mathbb{E}[\varphi(X)] = \frac{1}{\sqrt{2\pi\overline{\sigma}^2}} \int_{-\infty}^{\infty} \varphi(y) \exp\left(-\frac{y^2}{2\overline{\sigma}^2}\right) dy, \tag{2.2.15}$$

$$(\textit{resp. } \frac{1}{\sqrt{2\pi\underline{\sigma}^2}} \int_{-\infty}^{\infty} \varphi(y) \exp\left(-\frac{y^2}{2\underline{\sigma}^2}\right) dy).$$

Proof We only consider the non-trivial case of $\underline{\sigma}^2 > 0$. It is easy to check that

$$\overline{u}(t, x) := \frac{1}{\sqrt{2\pi\overline{\sigma}^2 t}} \int_{-\infty}^{\infty} \varphi(x + y) \exp\left(-\frac{y^2}{2\overline{\sigma}^2 t}\right) dy$$

is the unique smooth solution of the following classical linear heat equation

$$\partial_t \overline{u}(t, x) = \frac{\overline{\sigma}^2}{2} \partial^2_{xx} \overline{u}(t, x), \quad (t, x) \in (0, \infty) \times \mathbb{R},$$

with $\lim_{t\to 0}\overline{u}(t,x)=\varphi(x)$. It is also easy to check that, if φ is a convex function, then $\overline{u}(t,x)$ is also a convex function in x, thus $\partial_{xx}\overline{u}(t,x)\geq 0$. Consequently, $\overline{u}$ is also the unique smooth solution of the G-heat Eq. (2.2.14). We then have $\overline{u}(t,x)=\mathbb{E}[\varphi(x+\sqrt{t}X)]$ and thus (2.2.15) holds. The proof for the concave case is similar. □

Example 2.2.16 Let η be maximally distributed. The distribution of η is characterized by the solutions of the following parabolic PDEs defined on $[0,\infty)\times\mathbb{R}^d$:

$$\partial_t u - g_\eta(Du)=0,\quad u|_{t=0}=\varphi, \tag{2.2.16}$$

where $g_\eta(p):\mathbb{R}^d\mapsto\mathbb{R}$ is defined by

$$g_\eta(p):=\mathbb{E}[\langle p,\eta\rangle],\quad p\in\mathbb{R}^d.$$

It is easy to check that g_η is a sublinear function defined on $\mathbb{R}^d$. By Theorem 1.2.1 in Chap. 1, there exists a bounded, convex and closed subset $\bar{\Theta}\subset R^d$ such that

$$g_\eta(p)=\sup_{q\in\bar{\Theta}}\langle p,q\rangle,\quad p\in\mathbb{R}^d. \tag{2.2.17}$$

By this characterization, we can prove that the distribution of η is given by

$$\hat{\mathbb{F}}_\eta[\varphi]=\mathbb{E}[\varphi(\eta)]=\sup_{v\in\bar{\Theta}}\varphi(v)=\sup_{v\in\bar{\Theta}}\int_{\mathbb{R}^d}\varphi(x)\delta_v(dx),\quad \varphi\in C_{l.Lip}(\mathbb{R}^d), \tag{2.2.18}$$

where δ_v is the Dirac measure centered at v. This means that the maximal distribution with the uncertainty subset of probabilities as Dirac measures concentrated at all $v\in\bar{\Theta}$. We denote $\eta\overset{d}{=}N(\bar{\Theta}\times\{0\})$ (Recall Eq. (2.2.4), we can set $(q,Q)\in\bar{\Theta}\times\{0\}$).

In particular, for $d=1$,

$$g_\eta(p):=\mathbb{E}[p\eta]=\bar{\mu}p^+-\underline{\mu}p^-,\quad p\in\mathbb{R},$$

where $\bar{\mu}=\mathbb{E}[\eta]$ and $\underline{\mu}=-\hat{\mathbb{E}}[-\eta]$. The distribution of η is given by (2.2.18). Sometimes we also denote $\eta\overset{d}{=}N([\underline{\mu},\bar{\mu}]\times\{0\})$.

2.3 Existence of G-Distributed Random Variables

In this section, we give the proof of the existence of G-distributed random variables, namely, the proof of Proposition 2.2.7.

Let $G:\mathbb{R}^d\times\mathbb{S}(d)\mapsto\mathbb{R}$ be a given sublinear function, monotone in $A\in\mathbb{S}(d)$ in the sense of (2.2.3). We now construct a pair of d-dimensional random vectors (X,η) on some sublinear expectation space $(\Omega,\mathcal{H},\mathbb{E})$ satisfying (2.2.2) and (2.2.5).

For each $\varphi \in C_{l.Lip}(\mathbb{R}^{2d})$, let $u = u^{\varphi}$ be the unique viscosity solution of the G-equation (2.2.6) with $u^{\varphi}|_{t=0} = \varphi$. We take $\widetilde{\Omega} = \mathbb{R}^{2d}$, $\widetilde{\mathcal{H}} = C_{l.Lip}(\mathbb{R}^{2d})$ and $\widetilde{\omega} = (x, y) \in \mathbb{R}^{2d}$. The corresponding sublinear expectation $\widetilde{\mathbb{E}}[\cdot]$ is defined by $\widetilde{\mathbb{E}}[\xi] = u^{\varphi}(1, 0, 0)$, for each $\xi \in \widetilde{\mathcal{H}}$ of the form $\xi(\widetilde{\omega}) = (\varphi(x, y))_{(x,y)\in\mathbb{R}^{2d}} \in C_{l.Lip}(\mathbb{R}^{2d})$. The monotonicity and sub-additivity of u^{φ} with respect to φ are provided in (2.1.9)–(2.1.12) of Theorem 1.1.2 in Chap. 1. The property of constant preserving of $\widetilde{\mathbb{E}}[\cdot]$ are easy to check. Thus the functional $\widetilde{\mathbb{E}}[\cdot] : \widetilde{\mathcal{H}} \mapsto \mathbb{R}$ forms a sublinear expectation.

We now consider a pair of d-dimensional random vectors $(\widetilde{X}, \widetilde{\eta})(\widetilde{\omega}) = (x, y)$. We have

$$\widetilde{\mathbb{E}}[\varphi(\widetilde{X}, \widetilde{\eta})] = u^{\varphi}(1, 0, 0) \quad \text{for } \varphi \in C_{l.Lip}(\mathbb{R}^{2d}).$$

In particular, just setting $\varphi_0(x, y) = \frac{1}{2}\langle Ax, x\rangle + \langle p, y\rangle$, we can check that

$$u^{\varphi_0}(t, x, y) = G(p, A)t + \frac{1}{2}\langle Ax, x\rangle + \langle p, y\rangle .$$

We thus have

$$\widetilde{\mathbb{E}}\left[\tfrac{1}{2}\left\langle A\widetilde{X}, \widetilde{X}\right\rangle + \langle p, \widetilde{\eta}\rangle\right] = u^{\varphi_0}(1, 0, 0) = G(p, A), \quad (p, A) \in \mathbb{R}^d \times \mathbb{S}(d).$$

We construct a product space

$$(\Omega, \mathcal{H}, \mathbb{E}) = (\widetilde{\Omega} \times \widetilde{\Omega}, \widetilde{\mathcal{H}} \otimes \widetilde{\mathcal{H}}, \widetilde{\mathbb{E}} \otimes \widetilde{\mathbb{E}}),$$

and introduce two pairs of random vectors

$$(X, \eta)(\widetilde{\omega}_1, \widetilde{\omega}_2) = \widetilde{\omega}_1, \ (\bar{X}, \bar{\eta})(\widetilde{\omega}_1, \widetilde{\omega}_2) = \widetilde{\omega}_2, \quad (\widetilde{\omega}_1, \widetilde{\omega}_2) \in \widetilde{\Omega} \times \widetilde{\Omega}.$$

By Proposition 1.3.17 in Chap, 1, $(\bar{X}, \bar{\eta})$ is an independent copy of (X, η).

We now prove that the distribution of (X, η) satisfies condition (2.2.5). For each $\varphi \in C_{l.Lip}(\mathbb{R}^{2d})$ and for each fixed $\lambda > 0$, $(\bar{x}, \bar{y}) \in \mathbb{R}^{2d}$, since the function v defined by $v(t, x, y) := u^{\varphi}(\lambda t, \bar{x} + \sqrt{\lambda}x, \bar{y} + \lambda y)$ solves exactly the same Eq. (2.2.6), with Cauchy condition

$$v(t, \cdot, \cdot)|_{t=0} = \varphi(\bar{x} + \sqrt{\lambda} \times \cdot, \bar{y} + \lambda \times \cdot),$$

we have

$$\mathbb{E}[\varphi(\bar{x} + \sqrt{\lambda}X, \bar{y} + \lambda\eta)] = v(1, 0, 0) = u^{\varphi}(\lambda, \bar{x}, \bar{y}).$$

By the definition of $\mathbb{E}$, for each $t > 0$ and $s > 0$,

$$\begin{aligned}\mathbb{E}[\varphi(\sqrt{t}X + \sqrt{s}\bar{X}, t\eta + s\bar{\eta})] &= \mathbb{E}[\mathbb{E}[\varphi(\sqrt{t}x + \sqrt{s}\bar{X}, ty + s\bar{\eta})]_{(x,y)=(X,\eta)}] \\ &= \mathbb{E}[u^{\varphi}(s, \sqrt{t}X, t\eta)] = u^{u^{\varphi}(s,\cdot,\cdot)}(t, 0, 0) \\ &= u^{\varphi}(t + s, 0, 0)\end{aligned}$$

$$= \mathbb{E}[\varphi(\sqrt{t+s}X, (t+s)\eta)].$$

This means that $(\sqrt{t}X + \sqrt{s}\bar{X}, t\eta + s\bar{\eta}) \overset{d}{=} (\sqrt{t+s}X, (t+s)\eta)$. Thus the distribution of (X, η) satisfies condition (2.2.5).

Remark 2.3.1 From now on, when we mention the sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$, we suppose that there exists a pair of random vectors (X, η) on $(\Omega, \mathcal{H}, \mathbb{E})$ such that (X, η) is G-distributed.

2.4 Law of Large Numbers and Central Limit Theorem

In this section we present two most important results in the limit theory of nonlinear expectations: the law of large numbers and central limit theorem.

For the universality of theory, in the sequel we always assume that the space of random variables $\mathcal{H}$ is a linear space such that $|X| \in \mathcal{H}$ if $X \in \mathcal{H}$ and $\varphi(X) \in \mathcal{H}$ for each $X \in \mathcal{H}^n$, $\varphi \in C_{b.Lip}(\mathbb{R}^n)$. In fact, the condition that all moments of random variables exist is not necessary for our results.

Theorem 2.4.1 (Law of large numbers) *Let $\{Y_i\}_{i=1}^{\infty}$ be a sequence of $\mathbb{R}^d$-valued random variables on $(\Omega, \mathcal{H}, \mathbb{E})$. We assume that $Y_{i+1} \overset{d}{=} Y_i$ and Y_{i+1} is independent from $\{Y_1, \cdots, Y_i\}$ for each $i = 1, 2, \cdots$. We assume further the following uniform integrability condition:*

$$\lim_{\lambda\to+\infty} \mathbb{E}[(|Y_1| - \lambda)^+] = 0. \tag{2.4.1}$$

Then the sequence $\{\frac{1}{n}(Y_1 + \cdots + Y_n)\}_{n=1}^{\infty}$ converges in law to a maximal distribution, i.e.,

$$\lim_{n\to\infty} \mathbb{E}\Big[\varphi\big(\tfrac{1}{n}\big(Y_1 + \cdots + Y_n\big)\Big] = \max_{\theta\in\bar{\Theta}} \varphi(\theta), \tag{2.4.2}$$

for all functions $\varphi \in C(\mathbb{R}^d)$ satisfying linear growth condition, i.e., $|\varphi(x)| \leq C(1 + |x|)$, where $\bar{\Theta}$ is the (unique) bounded, closed and convex subset of $\mathbb{R}^d$ satisfying

$$\max_{\theta\in\bar{\Theta}} \langle p, \theta\rangle = \mathbb{E}[\langle p, Y_1\rangle], \quad p \in \mathbb{R}^d.$$

Remark 2.4.2 Note that in general $\varphi(Y_1)$ is in the completion space $\hat{\mathcal{H}}$, see Exercise 1.7.11 in Chap. 1.

The convergence result (2.4.2) means that the sequence $\{\frac{1}{n}\sum_{i=1}^{n} Y_i\}$ converges in law to a d-dimensional maximal distributed random vector η and the corresponding sublinear function $g : \mathbb{R}^d \mapsto \mathbb{R}$ is defined by

$$g(p) := \mathbb{E}[\langle p, Y_1\rangle], \;\; p \in \mathbb{R}^d.$$

Remark 2.4.3 When $d = 1$, the sequence $\{\frac{1}{n}\sum_{i=1}^{n} Y_i\}$ converges in law to $N([\underline{\mu}, \bar{\mu}] \times \{0\})$, where $\bar{\mu} = \mathbb{E}[Y_1]$ and $\underline{\mu} = -\mathbb{E}[-Y_1]$. In the general case, the averaged sum $\frac{1}{n}\sum_{i=1}^{n} Y_i$ converges in law to $N(\bar{\Theta} \times \{0\})$, where $\bar{\Theta} \subset \mathbb{R}^d$ is the bounded, convex and closed subset defined in Example 2.2.16. If we take in particular $\varphi(y) = d_{\bar{\Theta}}(y) = \inf\{|x - y| : x \in \bar{\Theta}\}$, then by (2.4.2) we have the following generalized law of large numbers:

$$\lim_{n\to\infty} \mathbb{E}\Big[d_{\bar{\Theta}}(\tfrac{1}{n}(Y_1 + \cdots + Y_n)\Big] = \sup_{\theta\in\bar{\Theta}} d_{\bar{\Theta}}(\theta) = 0. \tag{2.4.3}$$

If Y_i has no mean-uncertainty, or in other words, $\bar{\Theta}$ is a singleton: $\bar{\Theta} = \{\theta_0\}$, then (2.4.3) becomes

$$\lim_{n\to\infty} \mathbb{E}\Big[|\tfrac{1}{n}(Y_1 + \cdots + Y_n) - \theta_0|\Big] = 0.$$

The above LLN can be directly obtained from a more general limit theorem, namely, Theorem 2.4.7, which also contain the following central limit theorem (CLT) as a special case.

Theorem 2.4.4 (Central limit theorem with zero-mean) *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of $\mathbb{R}^d$-valued random variables on a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$. We assume that $X_{i+1} \overset{d}{=} X_i$ and X_{i+1} is independent from $\{X_1, \cdots, X_i\}$ for each $i = 1, 2, \cdots$. We further assume that $\mathbb{E}[X_1] = \mathbb{E}[-X_1] = 0$ and*

$$\lim_{\lambda\to+\infty} \mathbb{E}[(|X_1|^2 - \lambda)^+] = 0. \tag{2.4.4}$$

Then the sequence $\{\frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)\}_{n=1}^{\infty}$ converges in law to X:

$$\lim_{n\to\infty} \mathbb{E}\Big[\varphi\Big(\tfrac{1}{\sqrt{n}}(X_1 + \cdots + X_n)\Big)\Big] = \mathbb{E}[\varphi(X)],$$

for all functions $\varphi \in C(\mathbb{R}^d)$ with linear growth condition, where X is a G-normally distributed random vector and the corresponding sublinear function $G : \mathbb{S}(d) \mapsto \mathbb{R}$ is defined by

$$G(A) := \mathbb{E}\left[\tfrac{1}{2}\langle AX_1, X_1\rangle\right], \quad A \in \mathbb{S}(d).$$

Remark 2.4.5 A sufficient condition for (2.4.1) (resp. (2.4.4)) is

$$\mathbb{E}[|X_i|^{2+\delta}] < \infty \quad (\textit{resp.} \quad \mathbb{E}[|Y_i|^{1+\delta}] < \infty) \tag{2.4.5}$$

for some $\delta > 0$.

Remark 2.4.6 When $d = 1$, the sequence $\{\frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)\}_{n=1}^{\infty}$ converges in law to $N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$, where $\overline{\sigma}^2 = \mathbb{E}[X_1^2]$ and $\underline{\sigma}^2 = -\mathbb{E}[-X_1^2]$. In particular, if $\overline{\sigma}^2 = \underline{\sigma}^2$, we have the classical central limit theorem.

In this section we will prove our main theorem that nontrivially generalizes the above LLN and CLT.

Theorem 2.4.7 (Central limit theorem with law of large numbers) *Let* $\{(X_i, Y_i)\}_{i=1}^{\infty}$ *be a sequence of* $\mathbb{R}^d \times \mathbb{R}^d$*-valued random vectors on a sublinear expectation space* $(\Omega, \mathcal{H}, \mathbb{E})$. *We assume that* $(X_{i+1}, Y_{i+1}) \overset{d}{=} (X_i, Y_i)$ *and* (X_{i+1}, Y_{i+1}) *is independent from* $(X_1, Y_1), \cdots, (X_i, Y_i)$ *for each* $i = 1, 2, \cdots$. *We further assume that* $\mathbb{E}[X_1] = \mathbb{E}[-X_1] = 0$ *and*

$$\lim_{\lambda \to +\infty} \mathbb{E}[(|X_1|^2 - \lambda)^+] = 0, \quad \lim_{\lambda \to +\infty} \mathbb{E}[(|Y_1| - \lambda)^+] = 0. \tag{2.4.6}$$

Then the sequence $\{\sum_{i=1}^{n} (\frac{X_i}{\sqrt{n}} + \frac{Y_i}{n})\}_{n=1}^{\infty}$ *converges in law to* $X + \eta$, *i.e.,*

$$\lim_{n \to \infty} \mathbb{E}\left[\varphi\Big(\sum_{i=1}^{n} \Big(\frac{X_i}{\sqrt{n}} + \frac{Y_i}{n}\Big)\Big)\right] = \mathbb{E}[\varphi(X + \eta)], \tag{2.4.7}$$

for all functions $\varphi \in C(\mathbb{R}^d)$ *with a linear growth condition, where the pair of random vectors* (X, η) *is* G*-distributed. The corresponding sublinear function* $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ *is defined by*

$$G(p, A) := \mathbb{E}\Big[\langle p, Y_1 \rangle + \tfrac{1}{2} \langle AX_1, X_1 \rangle\Big], \ A \in \mathbb{S}(d), \quad p \in \mathbb{R}^d.$$

Thus the limit $\mathbb{E}[\varphi(X + \eta)]$ *equals to* $u(1, 0)$, *where* u *is the solution of the PDE (2.2.11) in Corollary 2.2.13.*

The following result can be immediately obtained from the above central limit theorem.

Theorem 2.4.8 *We make the same assumptions as in Theorem 2.4.7. Then for each function* $\varphi \in C(\mathbb{R}^d \times \mathbb{R}^d)$ *satisfying linear growth condition, we have*

$$\lim_{n \to \infty} \mathbb{E}\left[\varphi\Big(\sum_{i=1}^{n} \frac{X_i}{\sqrt{n}}, \sum_{i=1}^{n} \frac{Y_i}{n}\Big)\right] = \mathbb{E}[\varphi(X, \eta)].$$

Proof It is easy to prove Theorem 2.4.7 by Theorem 2.4.8. To prove Theorem 2.4.8 from Theorem 2.4.7, it suffices to define a pair of $2d$-dimensional random vectors

$$\bar{X}_i = (X_i, 0), \quad \bar{Y}_i = (0, Y_i) \ \text{ for } i = 1, 2, \cdots .$$

We have

$$\lim_{n\to\infty} \mathbb{E}\left[\varphi\Big(\sum_{i=1}^{n}\tfrac{X_i}{\sqrt{n}}, \sum_{i=1}^{n}\tfrac{Y_i}{n}\Big)\right] = \lim_{n\to\infty} \mathbb{E}\left[\varphi\Big(\sum_{i=1}^{n}\Big(\tfrac{\bar{X}_i}{\sqrt{n}} + \tfrac{\bar{Y}_i}{n}\Big)\Big)\right] = \mathbb{E}[\varphi(\bar{X}+\bar{\eta})]$$
$$= \mathbb{E}[\varphi(X,\eta)]$$

with $\bar{X}=(X,0)$ and $\bar{\eta}=(0,\eta)$. □

The following result is a direct consequence of Theorem 2.4.8.

Corollary 2.4.9 *Given a fixed $\varphi \in C(\mathbb{R}^d \times \mathbb{R}^d)$ satisfying a linear growth condition, we make the same assumptions as in Theorem 2.4.7. If there exists a sequence of $\mathbb{R}^d \times \mathbb{R}^d$-valued random vectors $\{(\bar{X}_i, \bar{Y}_i)\}_{i=1}^{\infty}$ on a sublinear expectation space $(\bar{\Omega}, \bar{\mathcal{H}}, \bar{\mathbb{E}})$ such that*

$$\bar{\mathbb{E}}\left[\varphi\Big(\sum_{i=1}^{n}\tfrac{\bar{X}_i}{\sqrt{n}}, \sum_{i=1}^{n}\tfrac{\bar{Y}_i}{n}\Big)\right] = \mathbb{E}\left[\varphi\Big(\sum_{i=1}^{n}\tfrac{X_i}{\sqrt{n}}, \sum_{i=1}^{n}\tfrac{Y_i}{n}\Big)\right]$$

for each $n \geq 1$, then

$$\lim_{n\to\infty} \mathbb{E}\left[\varphi\Big(\sum_{i=1}^{n}\tfrac{\bar{X}_i}{\sqrt{n}}, \sum_{i=1}^{n}\tfrac{\bar{Y}_i}{n}\Big)\right] = \mathbb{E}[\varphi(X,\eta)].$$

The following lemma tells us that the claim of Theorem 2.4.7 holds in a non-degenerate situation with test function $\varphi \in C_{b.Lip}(\mathbb{R}^d)$.

Lemma 2.4.10 *We keep the same assumptions as in Theorem 2.4.7. We further assume that there exists a constant $\beta > 0$ such that, for each A, $\bar{A} \in \mathbb{S}(d)$ with $A \geq \bar{A}$, we have*

$$\mathbb{E}[\langle AX_1, X_1\rangle] - \mathbb{E}[\langle \bar{A}X_1, X_1\rangle] \geq \beta\, \mathrm{tr}[A-\bar{A}], \tag{2.4.8}$$

where tr*[A] is the trace operator for $A \in \mathbb{S}(d)$. Then our main result (2.4.7) holds for test function $\varphi \in C_{b.Lip}(\mathbb{R}^d)$.*

Proof For a small but fixed $h > 0$, let V be the unique viscosity solution of

$$\partial_t V + G(DV, D^2V) = 0,\ (t,x) \in [0, 1+h) \times \mathbb{R}^d,\ V|_{t=1+h} = \varphi. \tag{2.4.9}$$

Since (X,η) satisfies (2.2.5), we have

$$V(h,0) = \mathbb{E}[\varphi(X+\eta)],\quad V(1+h,x) = \varphi(x). \tag{2.4.10}$$

Since (2.4.9) is a uniformly parabolic PDE and G is a convex function, by the interior regularity of V (see (2.1.13) in Chap. 1, or Theorem C.4.5 in Appendix C), we have

$$\|V\|_{C^{1+\alpha/2,2+\alpha}([0,1]\times\mathbb{R}^d)} < \infty \text{ for some } \alpha \in (0,1). \tag{2.4.11}$$

We set $\delta = \frac{1}{n}$ and $S_0^n = 0$ and

$$\bar{S}_i^n := \sum_{k=1}^{i} \left(\frac{X_k}{\sqrt{n}} + \frac{Y_k}{n} \right), \quad i = 1, 2, \cdots, n.$$

Then

$$V(1, \bar{S}_n^n) - V(0,0) = \sum_{i=0}^{n-1} \{V((i+1)\delta, \bar{S}_{i+1}^n) - V(i\delta, \bar{S}_i^n)\}$$
$$= \sum_{i=0}^{n-1} \{[V((i+1)\delta, \bar{S}_{i+1}^n) - V(i\delta, \bar{S}_{i+1}^n)] + [V(i\delta, \bar{S}_{i+1}^n) - V(i\delta, \bar{S}_i^n)]\} = \sum_{i=0}^{n-1} \left\{ I_\delta^i + J_\delta^i \right\}.$$

Here we regroup the main terms of Taylor's expansion of $V(i\delta + \delta, \bar{S}_{i+1}^n) - V(i\delta, \bar{S}_{i+1}^n)$ and $V(i\delta, \bar{S}_{i+1}^n) - V(i\delta, \bar{S}_i^n)$ into

$$J_\delta^i = \partial_t V(i\delta, \bar{S}_i^n)\delta + \left\langle DV(i\delta, \bar{S}_i^n), X_{i+1}\sqrt{\delta} + Y_{i+1}\delta \right\rangle + \frac{1}{2} \left\langle D^2 V(i\delta, \bar{S}_i^n) X_{i+1}, X_{i+1} \right\rangle \delta.$$

Their residue terms are put into I_δ^i:

$$I_\delta^i = \delta K_\delta^{0,i}(\omega) + \left\langle K_\delta^{1,i}, \delta Y_{i+1} \right\rangle + \left\langle K_\delta^{2,i} \sqrt{\delta} X_{i+1}, \sqrt{\delta} X_{i+1} \right\rangle$$

with

$$K_\delta^{0,i}(\omega) = \int_0^1 [\partial_t V(i\delta + \alpha\delta, \bar{S}_{i+1}^n) - \partial_t V(i\delta, \bar{S}_i^n)] d\alpha,$$
$$K_\delta^{1,i}(\omega) = \int_0^1 [DV(i\delta, \bar{S}_i^n + X_{i+1}\sqrt{\delta} + \alpha Y_{i+1}\delta) - DV(i\delta, \bar{S}_i^n)] d\alpha,$$
$$K_\delta^{2,i}(\omega) = \int_0^1 \int_0^1 [D^2 V(i\delta, \bar{S}_i + \alpha\beta X_{i+1}\sqrt{\delta})\beta - D^2 V(i\delta, \bar{S}_i^n)] d\alpha d\beta.$$

It then follows from the estimate (2.4.11) that, for all $i = 0, 1, \cdots, n-1$ with $n = 1, 2, \cdots$,

$$|K_\delta^{j,i}(\omega)| \le C, \quad j = 0, 1, 2,$$
$$\mathbb{E}[|K_\delta^{0,i}|] \le C\mathbb{E}[(1 + |X_{i+1}|^{\alpha/2} + |Y_{i+1}|^{\alpha/2})]\delta^{\alpha/2} = c_0 C \delta^{\alpha/2},$$
$$\mathbb{E}[|K_\delta^{j,i}|] \le C\mathbb{E}[(1 + |X_{i+1}|^{\alpha} + |Y_{i+1}|^{\alpha})]\delta^{\alpha} = c_1 C \delta^{\alpha}, \quad j = 1, 2,$$

where $C = \|V\|_{C^{1+\alpha/2,\alpha}[0,1]\times\mathbb{R}^d}$, $c_0 = \mathbb{E}[(1 + |X_1|^{\alpha/2} + |Y_1|^{\alpha/2})]$ and $c_1 = \mathbb{E}[(1 + |X_1|^{\alpha} + |Y_1|^{\alpha})]$. It follows from Condition (2.4.6) that, for each $\varepsilon > 0$, we can find a large enough number c such that $C\mathbb{E}[(|X_1|^2 - c)^+] + C\mathbb{E}[(|Y_1| - c)^+] < \varepsilon/2$, and for this fixed c, we can find a sufficiently small $\delta_0 > 0$, such that $c_0 C\delta^{\alpha/2} + 2cc_1 C\delta^{\alpha} < \varepsilon/2$, for all $\delta < \delta_0$. It follows that $\mathbb{E}[|I^i_\delta|]/\delta \to 0$, as $\delta \to 0$, uniformly. Therefore $\sum_{i=0}^{n-1} \mathbb{E}[|I^i_\delta|] \to 0$ as $n \to \infty$.

Observe that

$$\mathbb{E}\Big[\sum_{i=0}^{n-1} J^i_\delta\Big] - \sum_{i=0}^{n-1} \mathbb{E}\Big[|I^i_\delta|\Big] \le \mathbb{E}\big[V(1, \bar{S}^n_n)\big] - V(0, 0) \le \mathbb{E}\Big[\sum_{i=0}^{n-1} J^i_\delta\Big] + \sum_{i=0}^{n-1} \mathbb{E}[|I^i_\delta|]. \tag{2.4.12}$$

Now let us prove that

$$\mathbb{E}\Big[\sum_{i=0}^{n-1} J^i_\delta\Big] = 0.$$

Indeed, since

$$\mathbb{E}\Big[\Big\langle DV(i\delta, \bar{S}^n_i), X_{i+1}\sqrt{\delta}\Big\rangle\Big] = \mathbb{E}\Big[-\Big\langle DV(i\delta, \bar{S}^n_i), X_{i+1}\sqrt{\delta}\Big\rangle\Big] = 0,$$

we derive directly from the definition of the function G that

$$\mathbb{E}[J^i_\delta] = \mathbb{E}[\partial_t V(i\delta, \bar{S}^n_i) + G(DV(i\delta, \bar{S}^n_i), D^2V(i\delta, \bar{S}^n_i))]\delta.$$

Combining the above two equalities with $\partial_t V + G(DV, D^2V) = 0$ and use the independence of (X_{i+1}, Y_{i+1}) from $\{(X_1, Y_1), \cdots, (X_i, Y_i)\}$, we obtain

$$\mathbb{E}\Big[\sum_{i=0}^{n-1} J^i_\delta\Big] = \mathbb{E}\Big[\sum_{i=0}^{n-2} J^i_\delta\Big] = \cdots = 0.$$

Thus (2.4.12) can be rewritten as

$$|\mathbb{E}[V(1, \bar{S}^n_n)] - V(0, 0)| \le \sum_{i=0}^{n-1} \mathbb{E}[|I^i_\delta|].$$

As $n \to \infty$, we have

$$\lim_{n\to\infty} \mathbb{E}[V(1, \bar{S}^n_n)] = V(0, 0). \tag{2.4.13}$$

On the other hand, for each $t, t' \in [0, 1 + h]$ and $x \in \mathbb{R}^d$, we have

$$|V(t, x) - V(t', x)| \le C(\sqrt{|t - t'|} + |t - t'|).$$

Thus $|V(0,0) - V(h,0)| \le C\,(\sqrt{h} + h)$ and, by (2.4.13),

$$|\mathbb{E}[V(1, \bar{S}_n^n)] - \mathbb{E}[\varphi(\bar{S}_n^n)]| = |\mathbb{E}[V(1, \bar{S}_n^n)] - \mathbb{E}[V(1+h, \bar{S}_n^n)]| \le C(\sqrt{h} + h).$$

It follows from (2.4.10) and (2.4.13) that

$$\limsup_{n\to\infty} |\mathbb{E}[\varphi(\bar{S}_n^n)] - \mathbb{E}[\varphi(X+\eta)]| \le 2C(\sqrt{h} + h).$$

Since h can be arbitrarily small, we have

$$\lim_{n\to\infty} \mathbb{E}[\varphi(\bar{S}_n^n)] = \mathbb{E}[\varphi(X+\eta)].$$

□

Remark 2.4.11 From the proof we can check that the main assumption of identical distribution of $\{X_i, Y_i\}_{i=1}^\infty$ can be weakened to

$$\mathbb{E}\Big[\langle p, Y_i\rangle + \tfrac{1}{2}\langle AX_i, X_i\rangle\Big] = G(p, A), \quad i = 1, 2, \cdots, \ p \in \mathbb{R}^d, \ A \in \mathbb{S}(d).$$

We are now in the position to give the proof of Theorem 2.4.7.

Proof of Theorem 2.4.7 In the case when the uniform ellipticity condition (2.4.8) does not hold, we first introduce a perturbation to prove the above convergence for $\varphi \in C_{b.Lip}(\mathbb{R}^d)$. According to Definition 1.3.16 and Proposition 1.3.17 in Chap. 1, we can construct a sublinear expectation space $(\bar{\Omega}, \bar{\mathcal{H}}, \bar{\mathbb{E}})$ and a sequence of random vectors $\{(\bar{X}_i, \bar{Y}_i, \bar{\kappa}_i)\}_{i=1}^\infty$ such that, for each $n = 1, 2, \cdots$, $\{(\bar{X}_i, \bar{Y}_i)\}_{i=1}^n \overset{d}{=} \{(X_i, Y_i)\}_{i=1}^n$ and $(\bar{X}_{n+1}, \bar{Y}_{n+1}, \bar{\kappa}_{n+1})$ is independent from $\{(\bar{X}_i, \bar{Y}_i, \bar{\kappa}_i)\}_{i=1}^n$ and, moreover,

$$\bar{\mathbb{E}}[\psi(\bar{X}_i, \bar{Y}_i, \bar{\kappa}_i)] = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \mathbb{E}[\psi(X_i, Y_i, x)] e^{-|x|^2/2} dx \ \text{ for } \psi \in C_{b.Lip}(\mathbb{R}^{3\times d}).$$

We then use the perturbation $\bar{X}_i^\varepsilon = \bar{X}_i + \varepsilon\bar{\kappa}_i$ for a fixed $\varepsilon > 0$. It is easy to see that the sequence $\{(\bar{X}_i^\varepsilon, \bar{Y}_i)\}_{i=1}^\infty$ satisfies all conditions in the above CLT, in particular,

$$G_\varepsilon(p, A) := \bar{\mathbb{E}}\Big[\tfrac{1}{2}\big\langle A\bar{X}_1^\varepsilon, \bar{X}_1^\varepsilon\big\rangle + \big\langle p, \bar{Y}_1\big\rangle\Big] = G(p, A) + \tfrac{\varepsilon^2}{2}\mathrm{tr}[A].$$

This function G_ε is striongly elliptic. We then can apply Lemma 2.4.10 to

$$\bar{S}_n^\varepsilon := \sum_{i=1}^n \Big(\tfrac{\bar{X}_i^\varepsilon}{\sqrt{n}} + \tfrac{\bar{Y}_i}{n}\Big) = \sum_{i=1}^n \Big(\tfrac{\bar{X}_i}{\sqrt{n}} + \tfrac{\bar{Y}_i}{n}\Big) + \varepsilon J_n, \quad J_n = \sum_{i=1}^n \frac{\bar{\kappa}_i}{\sqrt{n}}$$

and obtain

$$\lim_{n\to\infty} \bar{\mathbb{E}}[\varphi(\bar{S}_n^\varepsilon)] = \bar{\mathbb{E}}[\varphi(\bar{X} + \bar{\eta} + \varepsilon\bar{\kappa})],$$

where $((\bar{X}, \bar{\kappa}), (\bar{\eta}, 0))$ is $\bar{G}$-distributed under $\bar{\mathbb{E}}[\cdot]$ and

$$\bar{G}(\bar{p}, \bar{A}) := \bar{\mathbb{E}}\Big[\tfrac{1}{2}\big\langle \bar{A}(\bar{X}_1, \bar{\kappa}_1)^T, (\bar{X}_1, \bar{\kappa}_1)^T\big\rangle + \big\langle \bar{p}, (\bar{Y}_1, 0)^T\big\rangle\Big], \quad \bar{A} \in \mathbb{S}(2d), \quad \bar{p} \in \mathbb{R}^{2d}.$$

By Proposition 2.2.6, it is easy to prove that $(\bar{X} + \varepsilon\bar{\kappa}, \bar{\eta})$ is G_ε-distributed and $(\bar{X}, \bar{\eta})$ is G-distributed. But we have

$$\begin{aligned}\big|\mathbb{E}[\varphi(\bar{S}_n)] - \bar{\mathbb{E}}[\varphi(\bar{S}_n^\varepsilon)]\big| &= \big|\bar{\mathbb{E}}[\varphi(\bar{S}_n^\varepsilon - \varepsilon J_n)] - \bar{\mathbb{E}}[\varphi(\bar{S}_n^\varepsilon)]\big| \\ &\le \varepsilon C \bar{\mathbb{E}}[|J_n|] \le C'\varepsilon\end{aligned}$$

and similarly,

$$\big|\mathbb{E}[\varphi(X + \eta)] - \bar{\mathbb{E}}[\varphi(\bar{X} + \bar{\eta} + \varepsilon\bar{\kappa})]\big| = \big|\bar{\mathbb{E}}[\varphi(\bar{X} + \bar{\eta})] - \bar{\mathbb{E}}[\varphi(\bar{X} + \bar{\eta} + \varepsilon\bar{\kappa})]\big| \le C\varepsilon.$$

Since ε can be arbitrarily small, it follows that

$$\lim_{n\to\infty} \mathbb{E}[\varphi(\bar{S}_n)] = \mathbb{E}[\varphi(X + \eta)] \ \text{ for all } \varphi \in C_{b.Lip}(\mathbb{R}^d).$$

On the other hand, it is easy to check that $\sup_n \mathbb{E}[|\bar{S}_n|^2] + \mathbb{E}[|X + \eta|^2] < \infty$. We then can apply the following lemma to prove that the above convergence holds for $\varphi \in C(\mathbb{R}^d)$ with linear growth condition. The proof is complete. □

Lemma 2.4.12 *Let* $(\Omega, \mathcal{H}, \mathbb{E})$ *and* $(\widetilde{\Omega}, \widetilde{\mathcal{H}}, \widetilde{\mathbb{E}})$ *be two sublinear expectation spaces and let* $Y_n \in \mathcal{H}$ *and* $Y \in \widetilde{\mathcal{H}}$, $n = 1, 2, \cdots$, *be given. We assume that, for a given* $p \ge 1$, $\sup_n \mathbb{E}[|Y_n|^p] + \widetilde{\mathbb{E}}[|Y|^p] < \infty$. *If the convergence* $\lim_{n\to\infty} \mathbb{E}[\varphi(Y_n)] = \widetilde{\mathbb{E}}[\varphi(Y)]$ *holds for each* $\varphi \in C_{b.Lip}(\mathbb{R}^d)$, *then it also holds for all functions* $\varphi \in C(\mathbb{R}^d)$ *with the growth condition* $|\varphi(x)| \le C(1 + |x|^{p-1})$.

Proof We first prove that the stated convergence holds for $\varphi \in C_b(\mathbb{R}^d)$ with a compact support. In this case, for each $\varepsilon > 0$, we can find $\bar{\varphi} \in C_{b.Lip}(\mathbb{R}^d)$ such that $\sup_{x\in\mathbb{R}^d} |\varphi(x) - \bar{\varphi}(x)| \le \frac{\varepsilon}{2}$. We have

$$\begin{aligned}&|\mathbb{E}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| \le |\mathbb{E}[\varphi(Y_n)] - \mathbb{E}[\bar{\varphi}(Y_n)]| + |\widetilde{\mathbb{E}}[\varphi(Y)] - \widetilde{\mathbb{E}}[\bar{\varphi}(Y)]| \\ &\quad + |\mathbb{E}[\bar{\varphi}(Y_n)] - \widetilde{\mathbb{E}}[\bar{\varphi}(Y)]| \le \varepsilon + |\mathbb{E}[\bar{\varphi}(Y_n)] - \widetilde{\mathbb{E}}[\bar{\varphi}(Y)]|.\end{aligned}$$

Thus $\limsup_{n\to\infty} |\mathbb{E}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| \le \varepsilon$. The convergence must hold since ε can be arbitrarily small.

Now let φ be an arbitrary $C(\mathbb{R}^d)$-function with growth condition $|\varphi(x)| \le C(1 + |x|^{p-1})$. For each $N > 0$ we can find $\varphi_1, \varphi_2 \in C(\mathbb{R}^d)$ such that $\varphi = \varphi_1 + \varphi_2$, where φ_1 has a compact support and $\varphi_2(x) = 0$ for $|x| \le N$, and $|\varphi_2(x)| \le |\varphi(x)|$ for all x. It is clear that

$$|\varphi_2(x)| \le \frac{2C(1 + |x|^p)}{N} \quad \text{for } x \in \mathbb{R}^d.$$

Thus

$$\begin{aligned}|\mathbb{E}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| &= |\mathbb{E}[\varphi_1(Y_n) + \varphi_2(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y) + \varphi_2(Y)]| \\ &\leq |\mathbb{E}[\varphi_1(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y)]| + \mathbb{E}[|\varphi_2(Y_n)|] + \widetilde{\mathbb{E}}[|\varphi_2(Y)|] \\ &\leq |\mathbb{E}[\varphi_1(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y)]| + \frac{2C}{N}(2 + \mathbb{E}[|Y_n|^p] + \widetilde{\mathbb{E}}[|Y|^p]) \\ &\leq |\mathbb{E}[\varphi_1(Y_n)] - \widetilde{\mathbb{E}}[\varphi_1(Y)]| + \frac{\bar{C}}{N},\end{aligned}$$

where $\bar{C} = 2C(2 + \sup_n \mathbb{E}[|Y_n|^p] + \widetilde{\mathbb{E}}[|Y|^p])$. We thus have $\limsup_{n\to\infty} |\mathbb{E}[\varphi(Y_n)] - \widetilde{\mathbb{E}}[\varphi(Y)]| \leq \frac{\bar{C}}{N}$. Since N can be arbitrarily large, $\mathbb{E}[\varphi(Y_n)]$ must converge to $\widetilde{\mathbb{E}}[\varphi(Y)]$. □

2.5 Exercises

Exercise 2.5.1 We consider $X = (X_1, X_2)$, where $X_1 \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$ with $\overline{\sigma} > \underline{\sigma}$, X_2 is an independent copy of X_1. Show that:

(i) For each $a \in \mathbb{R}^2$, $\langle a, X\rangle$ is a 1-dimensional G-normally distributed random variable.
(ii) X is not G-normally distributed.

Exercise 2.5.2 Let X be G-normally distributed. For each $\varphi \in C_{l.Lip}(\mathbb{R}^d)$, we define a function

$$u(t, x) := \mathbb{E}[\varphi(x + \sqrt{t}X)],\ \ (t, x) \in [0, \infty) \times \mathbb{R}^d.$$

Show that u is the unique viscosity solution of the PDE (2.2.12) with $u\big|_{t=0} = \varphi$.

Exercise 2.5.3 Let η be maximally distributed. For each $\varphi \in C_{l.Lip}(\mathbb{R}^d)$, we define a function

$$u(t, y) := \mathbb{E}[\varphi(y + t\eta)],\ \ (t, y) \in [0, \infty) \times \mathbb{R}^d.$$

Show that u is the unique viscosity solution of the PDE (2.2.16) with Cauchy condition $u|_{t=0} = \varphi$.

Exercise 2.5.4 Given a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$, where $\mathcal{H}$ is a linear space such that $|X| \in \mathcal{H}$ if $X \in \mathcal{H}$ and $\varphi(X) \in \mathcal{H}$ for each $X \in \mathcal{H}^n$, $\varphi \in C_{b.Lip}(\mathbb{R}^n)$. Suppose the random vector $\xi \in \mathcal{H}^d$ and satisfies the condition:

(i) $|\xi|^2 \in \mathcal{H}$ and $\lim_{\lambda\to\infty} \mathbb{E}[(|\xi|^2 - \lambda)^+] = 0$.
(ii) $\mathbb{E}[\varphi(a\xi + b\bar{\xi})] = \mathbb{E}[\varphi(\sqrt{a^2 + b^2}\xi)]$ for each $a, b \geq 0$ and $\varphi(x) \in C_{b.Lip}(\mathbb{R}^d)$, where $\bar{\xi}$ is an independent copy of ξ.

Show that for each $\psi(x) \in C_{l.Lip}(\mathbb{R}^d)$:

(1) $\psi(\xi) \in \hat{\mathcal{H}}$, where $\hat{\mathcal{H}}$ is the completion space of $\mathcal{H}$ under the norm $\mathbb{E}[|\cdot|]$.
(2) $\mathbb{E}[\psi(a\xi + b\bar{\xi})] = \mathbb{E}[\psi(\sqrt{a^2+b^2}\xi)]$ for each $a, b \geq 0$.

(Hint: use viscosity solution approach.)

Exercise 2.5.5 Prove that $\mathbb{E}[X^3] > 0$ and $\mathbb{E}[X^4 - 3X^2] > 0$ for $X \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$ with $\underline{\sigma}^2 < \overline{\sigma}^2$.

Exercise 2.5.6 Prove that

$$\mathbb{E}[\varphi(X)] \geq \sup_{\underline{\sigma} \leq \sigma \leq \overline{\sigma}} E_\sigma[\varphi(X)], \quad \text{for } \varphi \in C_{l,Lip}(\mathbb{R}).$$

where E_σ denotes the linear expectation corresponding to the classical normal distribution $N(0, \sigma^2)$. An interesting problem is to prove that, if φ is neither convex nor concave, the above inequality becomes strict.

Exercise 2.5.7 Let $\{X_i\}_{i=1}^\infty$ be a sequence of i.i.d. random variables with $\mathbb{E}[X_1] = \mathbb{E}[-X_1] = 0$ and $\mathbb{E}[|X_1|^q] < \infty$ for some $q \geq 2$. Prove that for each $2 \leq p \leq q$,

$$\mathbb{E}[|X_1 + \cdots + X_n|^p] \leq C_p n^{p/2},$$

where C_p is a constant depending on p.

Exercise 2.5.8 Let $X_i \in \mathcal{H}$, $i = 1, 2, \cdots$, be such that X_{i+1} is independent from $\{X_1, \cdots, X_i\}$, for each $i = 1, 2, \cdots$. We further assume that

$$\mathbb{E}[X_i] = -\mathbb{E}[-X_i] = 0,$$

$$\lim_{i\to\infty} \mathbb{E}[X_i^2] = \overline{\sigma}^2 < \infty, \ \lim_{i\to\infty} -\mathbb{E}[-X_i^2] = \underline{\sigma}^2,$$

$$\mathbb{E}[|X_i|^{2+\delta}] \leq M \ \text{ for some } \delta > 0 \text{ and a constant } M.$$

Prove that the sequence $\{\bar{S}_n\}_{n=1}^\infty$ defined by

$$\bar{S}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$$

converges in law to X, i.e.,

$$\lim_{n\to\infty} \mathbb{E}[\varphi(\bar{S}_n)] = \mathbb{E}[\varphi(X)] \ \text{ for } \varphi \in C_{b.lip}(\mathbb{R}),$$

where $X \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$.

In particular, if $\overline{\sigma}^2 = \underline{\sigma}^2$, it becomes a classical central limit theorem.

Exercise 2.5.9 Let the d-dimensional random vector X be defined on a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$. Suppose that

$$X + \bar{X} \overset{d}{=} \sqrt{2}X,$$

where $\bar{X}$ is an independent copy of X. Prove that X is G-normally distributed.

Notes and Comments

The material of this chapter is mainly from [139, 140, 142] (see also in the Notes [144] in which a stronger condition of the form

$$\mathbb{E}[|X_1|^p] < \infty, \quad \mathbb{E}[|Y_1|^q] < \infty, \quad \text{for a large } p > 2 \text{ and } q > 1$$

was in the place of the actual more general Condition (2.4.6). Condition (2.4.6) was proposed in Zhang [180]. The actual proof is a minor technique modification of the original one. We also mention that Chen considered strong laws of large numbers for sublinear expectations in [27] .

The notion of G-normal distribution was firstly introduced for in [138] for 1-dimen-sional case, and then in [141] for multi-dimensional case. In the classical situation, it is known that a distribution satisfying relation (2.2.1) is stable (see Lévy [107, 108]). In this sense, G-normal distribution, together with maximal-distribution, are most typical stable distributions in the framework of sublinear expectations.

Marinacci [115] proposed different notions of distributions and independence via capacity and the corresponding Choquet expectation to obtain a law of large numbers and a central limit theorem for non-additive probabilities (see also Maccheroni and Marinacci [116]). In fact, our results show that the limit in CLT, under uncertainty, is a G-normal distribution in which the distribution uncertainty cannot be just a family of classical normal distributions with different parameters (see Exercise 2.5.5).

The notion of viscosity solutions plays a basic role in the definitions and properties of G-normal distribution and maximal distribution. This notion was initially introduced by Crandall and Lions [38]. This is a fundamentally important notion in the theory of nonlinear parabolic and elliptic PDEs. Readers are referred to Crandall et al. [39] for rich references of the beautiful and powerful theory of viscosity solutions. Regarding books on the theory of viscosity solutions and the related HJB equations, see Barles [9], Fleming and Soner [66] as well as Yong and Zhou [177].

We note that, in the case when the uniform ellipticity condition holds, the viscosity solution (2.2.10) becomes a classical $C^{1+\frac{\alpha}{2},2+\alpha}$-solution (see the very profound result of Krylov [105] and in Cabre and Caffarelli [24] and Wang [171]). In 1-dimensional situation, when $\underline{\sigma}^2 > 0$, the G-equation becomes the following Barenblatt equation:

$$\partial_t u + \gamma|\partial_t u| = \triangle u, \ |\gamma| < 1.$$

This equation was first introduced by Barenblatt [8] (see also Avellaneda et al. [6]).

The rate of convergence of LLN and CLT under sublinear expectation plays a crucially important role in the statistical analysis for random data under uncertainty. We refer to Fang et al. [63], and Song [166, 167], in which a nonlinear generalization of Stein method, obtained by Hu et al. [83], is applied as a sharp tool to attack this problem. A very recent important contribution to the convergence rate of G-CLT is Krylov [106].

We also refer to Jin and Peng [99] for a design of unbiased optimal estimators, as well as [149] for the construction of G-VaR.

Part II
Stochastic Analysis Under G-Expectations

Chapter 3
G-Brownian Motion and Itô's Calculus

The aim of this chapter is to introduce the concept of G-Brownian motion, study its properties and construct Itô's integral with respect to G-Brownian motion. We emphasize here that this G-Brownian motion $B_t, t \geq 0$ is consistent with the classical one. In fact once its mean uncertainty and variance uncertainty vanish, namely

$$\hat{\mathbb{E}}[B_1] = -\hat{\mathbb{E}}[-B_1] \quad \text{and} \quad \hat{\mathbb{E}}[B_1^2] = -\hat{\mathbb{E}}[-B_1^2],$$

then B becomes a classical Brownian motion. This G-Brownian motion also has independent and stable increments. G-Brownian motion has a very rich and interesting new structure which non-trivially generalizes the classical one. Thus we can develop the related stochastic calculus, especially Itô's integrals and the related quadratic variation process. A very interesting feature of the G-Brownian motion is that its quadratic process also has independent increments which are identically distributed. The corresponding G-Itô's formula is also presented.

We emphasize that the above construction of G-Brownian motion and the establishment of the corresponding stochastic analysis of generalized Itô's type, from this chapter to Chap. 5, have been rigorously realized without firstly constructing a probability space or its generalization, whereas its special situation of linear expectation corresponds in fact to the classical Brownian motion under a Wiener probability measure space. This is an important advantage of the expectation-based framework. The corresponding path-wise analysis of G-Brownian motion functional will be established in Chap. 6, after the introduction of the corresponding G-capacity. We can see that all results obtained in this chapter to Chap. 5 still hold true in G-capacity surely analysis.

3.1 Brownian Motion on a Sublinear Expectation Space

Definition 3.1.1 Let $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ be a sublinear expectation space. $(X_t)_{t\geq 0}$ is called a d-dimensional **stochastic process** if for each $t \geq 0$, X_t is a d-dimensional random vector in $\mathcal{H}$.

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_3

We now give the definition of Brownian motion on sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$.

Definition 3.1.2 A d-dimensional stochastic process $(B_t)_{t\geq 0}$ on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called a ***G*-Brownian motion** if the following properties are satisfied:

(i) $B_0(\omega)=0$;

(ii) For each $t, s \geq 0$, $B_{t+s} - B_t$ and B_s are identically distributed and $B_{t+s} - B_t$ is independent from $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$, for each $n \in \mathbb{N}$ and $0 \leq t_1 \leq \cdots \leq t_n \leq t$.

(iii) $\lim_{t\downarrow 0} \hat{\mathbb{E}}[|B_t|^3]t^{-1} = 0$.

Moreover, if $\hat{\mathbb{E}}[B_t] = \hat{\mathbb{E}}[-B_t] = 0$, then $(B_t)_{t\geq 0}$ is called a **symmetric *G*-Brownian motion**.

In the sublinear expectation space, symmetric G-Brownian motion is an important case of Brownian motion. From now on up to Sect. 3.6, we will study its properties, which are needed in stochastic analysis of G-Brownian motion. The following theorem gives a characterization of the symmetric Brownian motion.

Theorem 3.1.3 *Let $(B_t)_{t\geq 0}$ be a given $\mathbb{R}^d$–valued symmetric G-Brownian motion on a sublinear expectation $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$. Then, for each fixed $\varphi \in C_{b.Lip}(\mathbb{R}^d)$, the function*

$$u(t,x) := \hat{\mathbb{E}}[\varphi(x + B_t)], \quad (t,x) \in [0,\infty) \times \mathbb{R}^d$$

is the viscosity solution of the following parabolic PDE:

$$\partial_t u - G(D^2 u) = 0, \quad u|_{t=0} = \varphi. \tag{3.1.1}$$

where

$$G(A) = \frac{1}{2}\hat{\mathbb{E}}[\langle AB_1, B_1\rangle], \ A \in \mathbb{S}(d). \tag{3.1.2}$$

In particular, B_1 is G-normally distributed and $B_t \overset{d}{=} \sqrt{t}B_1$.

Proof We only need to prove that u is the viscosity solution. We first show that

$$\hat{\mathbb{E}}[\langle AB_t, B_t\rangle] = 2G(A)t, \quad A \in \mathbb{S}(d).$$

For each given $A \in \mathbb{S}(d)$, we set $b(t) = \hat{\mathbb{E}}[\langle AB_t, B_t\rangle]$. Then $b(0) = 0$ and $|b(t)| \leq |A|(\hat{\mathbb{E}}[|B_t|^3])^{2/3} \to 0$ as $t \to 0$. Note that $\hat{\mathbb{E}}[B_t] = \hat{\mathbb{E}}[-B_t] = 0$, we have for each $t, s \geq 0$,

$$\begin{aligned} b(t+s) &= \hat{\mathbb{E}}[\langle AB_{t+s}, B_{t+s}\rangle] = \hat{\mathbb{E}}[\langle A(B_{t+s} - B_s + B_s), B_{t+s} - B_s + B_s\rangle] \\ &= \hat{\mathbb{E}}[\langle A(B_{t+s} - B_s), (B_{t+s} - B_s)\rangle + \langle AB_s, B_s\rangle + 2\langle A(B_{t+s} - B_s), B_s\rangle] \\ &= b(t) + b(s), \end{aligned}$$

thus $b(t) = b(1)t = 2G(A)t$.

Then we show that u is Lipschitz in x and $\frac{1}{2}$-Hölder continuous in t. In fact, for each fixed t, $u(t,\cdot) \in C_{b.Lip}(\mathbb{R}^d)$ since

$$\begin{aligned}|u(t,x)-u(t,y)| &= |\hat{\mathbb{E}}[\varphi(x+B_t)] - \hat{\mathbb{E}}[\varphi(y+B_t)]| \\ &\le \hat{\mathbb{E}}[|\varphi(x+B_t)-\varphi(y+B_t)|] \\ &\le C|x-y|,\end{aligned}$$

where C is the Lipschitz constant of φ.

For each $\delta \in [0,t]$, since $B_t - B_\delta$ is independent from B_δ, we also have

$$\begin{aligned}u(t,x) &= \hat{\mathbb{E}}[\varphi(x+B_\delta+(B_t-B_\delta)] \\ &= \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(y+(B_t-B_\delta))]_{y=x+B_\delta}],\end{aligned}$$

hence

$$u(t,x) = \hat{\mathbb{E}}[u(t-\delta, x+B_\delta)]. \tag{3.1.3}$$

Thus

$$\begin{aligned}|u(t,x)-u(t-\delta,x)| &= |\hat{\mathbb{E}}[u(t-\delta,x+B_\delta)-u(t-\delta,x)]| \\ &\le \hat{\mathbb{E}}[|u(t-\delta,x+B_\delta)-u(t-\delta,x)|] \\ &\le \hat{\mathbb{E}}[C|B_\delta|] \le C\sqrt{2G(I)}\sqrt{\delta}.\end{aligned}$$

To show that u is a viscosity solution of (3.1.1), we fix $(t,x) \in (0,\infty)\times\mathbb{R}^d$ and let $v \in C_b^{2,3}([0,\infty)\times\mathbb{R}^d)$ be such that $v \ge u$ and $v(t,x) = u(t,x)$. From (3.1.3) we have

$$v(t,x) = \hat{\mathbb{E}}[u(t-\delta,x+B_\delta)] \le \hat{\mathbb{E}}[v(t-\delta,x+B_\delta)].$$

Therefore by Taylor's expansion,

$$\begin{aligned}0 &\le \hat{\mathbb{E}}[v(t-\delta,x+B_\delta)-v(t,x)] \\ &= \hat{\mathbb{E}}[v(t-\delta,x+B_\delta)-v(t,x+B_\delta)+(v(t,x+B_\delta)-v(t,x))] \\ &= \hat{\mathbb{E}}[-\partial_t v(t,x)\delta + \langle Dv(t,x),B_\delta\rangle + \tfrac{1}{2}\langle D^2v(t,x)B_\delta,B_\delta\rangle + I_\delta] \\ &\le -\partial_t v(t,x)\delta + \tfrac{1}{2}\hat{\mathbb{E}}[\langle D^2v(t,x)B_\delta,B_\delta\rangle] + \hat{\mathbb{E}}[I_\delta] \\ &= -\partial_t v(t,x)\delta + G(D^2v(t,x))\delta + \hat{\mathbb{E}}[I_\delta],\end{aligned}$$

where

$$I_\delta = \int_0^1 -[\partial_t v(t-\beta\delta,x+B_\delta)-\partial_t v(t,x)]\delta d\beta$$

$$+\int_0^1\int_0^1 \langle (D^2v(t,x+\alpha\beta B_\delta)-D^2v(t,x))B_\delta, B_\delta\rangle \alpha d\beta d\alpha.$$

In view of condition (iii) in Definition 3.1.2, we can check that $\lim_{\delta\downarrow 0}\delta^{-1}\hat{\mathbb{E}}[|I_\delta|]=0$, from which we get $\partial_t v(t,x)-G(D^2v(t,x))\le 0$, hence u is a viscosity subsolution of (3.1.1). We can analogously prove that u is a viscosity supersolution. Thus u is a viscosity solution. □

For simplicity, symmetric Brownian motion is also called ***G*-Brownian motion**, associated with the generator G given by (3.1.2).

Remark 3.1.4 We can prove that, for each $t_0>0$, $(B_{t+t_0}-B_{t_0})_{t\ge 0}$ is a G-Brownian motion. For each $\lambda>0$, $(\lambda^{-\frac{1}{2}}B_{\lambda t})_{t\ge 0}$ is also a symmetric G-Brownian motion. This is the scaling property of G-Brownian motion, which is the same as that for the classical Brownian motion.

In the rest of this book we will use the notation

$$B_t^{\mathbf{a}}=\langle \mathbf{a}, B_t\rangle \quad \text{for each } \mathbf{a}=(a_1,\cdots,a_d)^T\in\mathbb{R}^d.$$

By the above definition we have the following proposition which is important in stochastic calculus.

Proposition 3.1.5 *Let $(B_t)_{t\ge 0}$ be a d-dimensional G-Brownian motion on a sublinear expectation space $(\Omega,\mathcal{H},\mathbb{E})$. Then $(B_t^{\mathbf{a}})_{t\ge 0}$ is a 1-dimensional $G_{\mathbf{a}}$-Brownian motion for each $\mathbf{a}\in\mathbb{R}^d$, where $G_{\mathbf{a}}(\alpha)=\frac{1}{2}(\sigma^2_{\mathbf{aa}^T}\alpha^+-\sigma^2_{-\mathbf{aa}^T}\alpha^-)$, $\sigma^2_{\mathbf{aa}^T}=2G(\mathbf{aa}^T)=\hat{\mathbb{E}}[\langle \mathbf{a},B_1\rangle^2]$, $\sigma^2_{-\mathbf{aa}^T}=-2G(-\mathbf{aa}^T)=-\hat{\mathbb{E}}[-\langle \mathbf{a},B_1\rangle^2]$.*

In particular, for each $t,s\ge 0$, $B^{\mathbf{a}}_{t+s}-B^{\mathbf{a}}_t\overset{d}{=}N(\{0\}\times[s\sigma^2_{-\mathbf{aa}^T},s\sigma^2_{\mathbf{aa}^T}])$.

Proposition 3.1.6 *For each convex function $\varphi\in C_{l.Lip}(\mathbb{R})$, we have*

$$\hat{\mathbb{E}}[\varphi(B^{\mathbf{a}}_{t+s}-B^{\mathbf{a}}_t)]=\frac{1}{\sqrt{2\pi s\sigma^2_{\mathbf{aa}^T}}}\int_{-\infty}^{\infty}\varphi(x)\exp(-\frac{x^2}{2s\sigma^2_{\mathbf{aa}^T}})dx.$$

For each concave function $\varphi\in C_{l.Lip}(\mathbb{R})$ and $\sigma^2_{-\mathbf{aa}^T}>0$, we have

$$\hat{\mathbb{E}}[\varphi(B^{\mathbf{a}}_{t+s}-B^{\mathbf{a}}_t)]=\frac{1}{\sqrt{2\pi s\sigma^2_{-\mathbf{aa}^T}}}\int_{-\infty}^{\infty}\varphi(x)\exp(-\frac{x^2}{2s\sigma^2_{-\mathbf{aa}^T}})dx.$$

In particular, the following relations are true:

$$\hat{\mathbb{E}}[(B^{\mathbf{a}}_t-B^{\mathbf{a}}_s)^2]=\sigma^2_{\mathbf{aa}^T}(t-s),\quad \hat{\mathbb{E}}[(B^{\mathbf{a}}_t-B^{\mathbf{a}}_s)^4]=3\sigma^4_{\mathbf{aa}^T}(t-s)^2,$$
$$\hat{\mathbb{E}}[-(B^{\mathbf{a}}_t-B^{\mathbf{a}}_s)^2]=-\sigma^2_{-\mathbf{aa}^T}(t-s),\quad \hat{\mathbb{E}}[-(B^{\mathbf{a}}_t-B^{\mathbf{a}}_s)^4]=-3\sigma^4_{-\mathbf{aa}^T}(t-s)^2.$$

3.2 Existence of *G*-Brownian Motion

In the rest of this book, we use the notation $\Omega = C_0^d(\mathbb{R}^+)$ for the space of all $\mathbb{R}^d$–valued continuous paths $(\omega_t)_{t\in\mathbb{R}^+}$, with $\omega_0 = 0$, equipped with the distance

$$\rho(\omega^{(1)}, \omega^{(2)}) := \sum_{i=1}^{\infty} 2^{-i}[(\max_{t\in[0,i]} |\omega_t^{(1)} - \omega_t^{(2)}|) \wedge 1], \quad \omega^{(1)}, \omega^{(2)} \in \Omega.$$

For each fixed $T \in [0, \infty)$, we set $\Omega_T := \{\omega_{\cdot\wedge T} : \omega \in \Omega\}$. We will consider the canonical process $B_t(\omega) = \omega_t$, $t \in [0, \infty)$, for $\omega \in \Omega$.

For each fixed $T \in [0, \infty)$, we set also

$$Lip(\Omega_T) := \{\varphi(B_{t_1\wedge T}, \cdots, B_{t_n\wedge T}) : n \in \mathbb{N}, t_1, \cdots, t_n \in [0, \infty),\ \varphi \in C_{l.Lip}(\mathbb{R}^{d\times n})\ \}.$$

It is clear that $Lip(\Omega_t)\subseteq Lip(\Omega_T)$, for $t \leq T$. We set

$$Lip(\Omega) := \bigcup_{n=1}^{\infty} Lip(\Omega_n).$$

Remark 3.2.1 It is clear that $C_{l.Lip}(\mathbb{R}^{d\times n})$, $Lip(\Omega_T)$ and $Lip(\Omega)$ are vector lattices. Moreover, note that $\varphi, \psi \in C_{l.Lip}(\mathbb{R}^{d\times n})$ implies $\varphi \cdot \psi \in C_{l.Lip}(\mathbb{R}^{d\times n})$, then $X, Y \in Lip(\Omega_T)$ implies $X \cdot Y \in Lip(\Omega_T)$. In particular, for each $t \in [0, \infty)$, $B_t \in Lip(\Omega)$.

Let $G(\cdot) : \mathbb{S}(d) \mapsto \mathbb{R}$ be a given monotone and sublinear function. By Theorem 1.2.1 in Chap. 1, there exists a bounded, convex and closed subset $\Sigma \subset \mathbb{S}_+(d)$ such that

$$G(A) = \frac{1}{2} \sup_{B\in\Sigma} (A, B), \quad A \in \mathbb{S}(d).$$

By Sect. 2.3 in Chap. 2, we know that the G-normal distribution $N(\{0\} \times \Sigma)$ exists.

Let us construct a sublinear expectation on $(\Omega, Lip(\Omega))$ such that the canonical process $(B_t)_{t\geq 0}$ is a G-Brownian motion. For this, we first construct a sequence of d-dimensional random vectors $(\xi_i)_{i=1}^{\infty}$ on a sublinear expectation space $(\widetilde{\Omega}, \widetilde{\mathcal{H}}, \widetilde{\mathbb{E}})$ such that ξ_i is G-normally distributed and ξ_{i+1} is independent from $(\xi_1, \cdots, \xi_i)$ for each $i = 1, 2, \cdots$.

We now construct a sublinear expectation $\hat{\mathbb{E}}$ defined on $Lip(\Omega)$ via the following procedure: for each $X \in Lip(\Omega)$ with

$$X = \varphi(B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$$

for some $\varphi \in C_{l.Lip}(\mathbb{R}^{d\times n})$ and $0 = t_0 < t_1 < \cdots < t_n < \infty$, we set

$$\hat{\mathbb{E}}[\varphi(B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})] := \widetilde{\mathbb{E}}[\varphi(\sqrt{t_1 - t_0}\xi_1, \cdots, \sqrt{t_n - t_{n-1}}\xi_n)].$$

The related conditional expectation of $X = \varphi(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$ under Ω_{t_j} is defined by

$$\hat{\mathbb{E}}[X|\Omega_{t_j}] = \hat{\mathbb{E}}[\varphi(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})|\Omega_{t_j}] := \psi(B_{t_1}, \cdots, B_{t_j} - B_{t_{j-1}}),$$

where

$$\psi(x_1, \cdots, x_j) = \widetilde{\mathbb{E}}[\varphi(x_1, \cdots, x_j, \sqrt{t_{j+1} - t_j}\xi_{j+1}, \cdots, \sqrt{t_n - t_{n-1}}\xi_n)].$$

$\hat{\mathbb{E}}[\cdot]$ consistently defines a sublinear expectation on $Lip(\Omega)$ and $(B_t)_{t\geq 0}$ is a G-Brownian motion. Since $Lip(\Omega_T)\subseteq Lip(\Omega)$, $\hat{\mathbb{E}}[\cdot]$ is also a sublinear expectation on $Lip(\Omega_T)$.

Definition 3.2.2 The sublinear expectation $\hat{\mathbb{E}}[\cdot]$: $Lip(\Omega) \mapsto \mathbb{R}$ defined through the above procedure is called a G–**expectation**. The corresponding canonical process $(B_t)_{t\geq 0}$ on the sublinear expectation space $(\Omega, Lip(\Omega), \hat{\mathbb{E}})$ is called a G–Brownian motion.

In the rest of the book, when we talk about G–Brownian motion, we mean that the canonical process $(B_t)_{t\geq 0}$ is under G-expectation.

Proposition 3.2.3 *We list the properties of* $\hat{\mathbb{E}}[\cdot|\Omega_t]$ *that hold for each* $X, Y \in Lip(\Omega)$*:*

(i) *If* $X \geq Y$, *then* $\hat{\mathbb{E}}[X|\Omega_t] \geq \hat{\mathbb{E}}[Y|\Omega_t]$.
(ii) $\hat{\mathbb{E}}[\eta|\Omega_t] = \eta$, *for each* $t \in [0, \infty)$ *and* $\eta \in Lip(\Omega_t)$.
(iii) $\hat{\mathbb{E}}[X|\Omega_t] - \hat{\mathbb{E}}[Y|\Omega_t] \leq \hat{\mathbb{E}}[X - Y|\Omega_t]$.
(iv) $\hat{\mathbb{E}}[\eta X|\Omega_t] = \eta^+\hat{\mathbb{E}}[X|\Omega_t] + \eta^-\hat{\mathbb{E}}[-X|\Omega_t]$, *for each* $\eta \in Lip(\Omega_t)$.
(v) $\hat{\mathbb{E}}[\hat{\mathbb{E}}[X|\Omega_t]|\Omega_s] = \hat{\mathbb{E}}[X|\Omega_{t\wedge s}]$, *in particular*, $\hat{\mathbb{E}}[\hat{\mathbb{E}}[X|\Omega_t]] = \hat{\mathbb{E}}[X]$.

For each $X \in Lip(\Omega^t)$, $\hat{\mathbb{E}}[X|\Omega_t] = \hat{\mathbb{E}}[X]$, *where* $Lip(\Omega^t)$ *is the linear space of random variables with the form*

$$\varphi(B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \cdots, B_{t_{n+1}} - B_{t_n}),$$
$$n = 1, 2, \cdots, \ \varphi \in C_{l.Lip}(\mathbb{R}^{d\times n}), \ t_1, \cdots, t_n, t_{n+1} \in [t, \infty).$$

Remark 3.2.4 Properties (ii) and (iii) imply

$$\hat{\mathbb{E}}[X + \eta|\Omega_t] = \hat{\mathbb{E}}[X|\Omega_t] + \eta \ \text{ for } \eta \in Lip(\Omega_t).$$

We now consider the completion of sublinear expectation space $(\Omega, Lip(\Omega), \hat{\mathbb{E}})$. For $p \geq 1$, we denote by

$$L_G^p(\Omega) := \{ \textit{the completion of the space } Lip(\Omega) \textit{ under the norm } \|X\|_p := (\hat{\mathbb{E}}[|X|^p])^{1/p}\}.$$

Similarly, we can define $L_G^p(\Omega_T)$, $L_G^p(\Omega_T^t)$ and $L_G^p(\Omega^t)$. It is clear that for each $0 \leq t \leq T < \infty$, $L_G^p(\Omega_t) \subseteq L_G^p(\Omega_T) \subseteq L_G^p(\Omega)$.

According to Sect. 1.4 in Chap. 1, $\hat{\mathbb{E}}[\cdot]$ can be continuously extended to a sublinear expectation on $(\Omega, L_G^1(\Omega))$ and still denoted by $\hat{\mathbb{E}}[\cdot]$. We now consider the extension of conditional expectations. For each fixed $t \leq T$, the conditional G-expectation $\hat{\mathbb{E}}[\cdot|\Omega_t] : Lip(\Omega_T) \mapsto Lip(\Omega_t)$ is a continuous mapping under $\|\cdot\|$. Indeed, we have

$$\hat{\mathbb{E}}[X|\Omega_t] - \hat{\mathbb{E}}[Y|\Omega_t] \leq \hat{\mathbb{E}}[X - Y|\Omega_t] \leq \hat{\mathbb{E}}[|X - Y||\Omega_t],$$

then

$$|\hat{\mathbb{E}}[X|\Omega_t] - \hat{\mathbb{E}}[Y|\Omega_t]| \leq \hat{\mathbb{E}}[|X - Y||\Omega_t].$$

We thus obtain

$$\left\|\hat{\mathbb{E}}[X|\Omega_t] - \hat{\mathbb{E}}[Y|\Omega_t]\right\| \leq \|X - Y\|.$$

It follows that $\hat{\mathbb{E}}[\cdot|\Omega_t]$ can also be extended as a continuous mapping

$$\hat{\mathbb{E}}[\cdot|\Omega_t] : L_G^1(\Omega_T) \mapsto L_G^1(\Omega_t).$$

If the above T is not fixed, then we can obtain $\hat{\mathbb{E}}[\cdot|\Omega_t] : L_G^1(\Omega) \mapsto L_G^1(\Omega_t)$.

Remark 3.2.5 Proposition 3.2.3 also holds for $X, Y \in L_G^1(\Omega)$. But in (iv), $\eta \in L_G^1(\Omega_t)$ should be bounded, since $X, Y \in L_G^1(\Omega)$ does not imply that $X \cdot Y \in L_G^1(\Omega)$.

In particular, we have the following independence property:

$$\hat{\mathbb{E}}[X|\Omega_t] = \hat{\mathbb{E}}[X], \quad \forall X \in L_G^1(\Omega^t).$$

We give the following definition similar to the classical one:

Definition 3.2.6 An n-dimensional random vector $Y \in (L_G^1(\Omega))^n$ is said to be independent from Ω_t for some given t if for each $\varphi \in C_{b.Lip}(\mathbb{R}^n)$ we have

$$\hat{\mathbb{E}}[\varphi(Y)|\Omega_t] = \hat{\mathbb{E}}[\varphi(Y)].$$

Remark 3.2.7 Just as in the classical situation, the increments of G–Brownian motion $(B_{t+s} - B_t)_{s \geq 0}$ are independent from Ω_t, for each $t \geq 0$.

Example 3.2.8 For each fixed $\mathbf{a} \in \mathbb{R}^d$ and for each $0 \leq s \leq t$, we have

$$\begin{aligned}
&\hat{\mathbb{E}}[B_t^{\mathbf{a}} - B_s^{\mathbf{a}}|\Omega_s] = 0, \quad \hat{\mathbb{E}}[-(B_t^{\mathbf{a}} - B_s^{\mathbf{a}})|\Omega_s] = 0,\\
&\hat{\mathbb{E}}[(B_t^{\mathbf{a}} - B_s^{\mathbf{a}})^2|\Omega_s] = \sigma_{\mathbf{aa}^T}^2(t - s), \quad \hat{\mathbb{E}}[-(B_t^{\mathbf{a}} - B_s^{\mathbf{a}})^2|\Omega_s] = -\sigma_{-\mathbf{aa}^T}^2(t - s),\\
&\hat{\mathbb{E}}[(B_t^{\mathbf{a}} - B_s^{\mathbf{a}})^4|\Omega_s] = 3\sigma_{\mathbf{aa}^T}^4(t - s)^2, \quad \hat{\mathbb{E}}[-(B_t^{\mathbf{a}} - B_s^{\mathbf{a}})^4|\Omega_s] = -3\sigma_{-\mathbf{aa}^T}^4(t - s)^2,
\end{aligned}$$

where $\sigma_{\mathbf{aa}^T}^2 = 2G(\mathbf{aa}^T)$ and $\sigma_{-\mathbf{aa}^T}^2 = -2G(-\mathbf{aa}^T)$.

The following property is very useful.

Proposition 3.2.9 *Let* $X, Y \in L_G^1(\Omega)$ *be such that* $\hat{\mathbb{E}}[Y|\Omega_t] = -\hat{\mathbb{E}}[-Y|\Omega_t]$, *for some* $t \in [0, T]$. *Then we have*

$$\hat{\mathbb{E}}[X + Y|\Omega_t] = \hat{\mathbb{E}}[X|\Omega_t] + \hat{\mathbb{E}}[Y|\Omega_t].$$

In particular, if $\hat{\mathbb{E}}[Y|\Omega_t] = \hat{\mathbb{E}}[-Y|\Omega_t] = 0$, *then* $\hat{\mathbb{E}}[X + Y|\Omega_t] = \hat{\mathbb{E}}[X|\Omega_t]$.

Proof This follows from the following two inequalities:

$$\hat{\mathbb{E}}[X + Y|\Omega_t] \le \hat{\mathbb{E}}[X|\Omega_t] + \hat{\mathbb{E}}[Y|\Omega_t],$$
$$\hat{\mathbb{E}}[X + Y|\Omega_t] \ge \hat{\mathbb{E}}[X|\Omega_t] - \hat{\mathbb{E}}[-Y|\Omega_t] = \hat{\mathbb{E}}[X|\Omega_t] + \hat{\mathbb{E}}[Y|\Omega_t].$$

□

Example 3.2.10 For each $\mathbf{a} \in \mathbb{R}^d$, $0 \le t \le T$, $X \in L_G^1(\Omega_t)$ and $\varphi \in C_{l.Lip}(\mathbb{R})$, we have

$$\begin{aligned}\hat{\mathbb{E}}[X\varphi(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})|\Omega_t] &= X^+\hat{\mathbb{E}}[\varphi(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})|\Omega_t] + X^-\hat{\mathbb{E}}[-\varphi(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})|\Omega_t]\\ &= X^+\hat{\mathbb{E}}[\varphi(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})] + X^-\hat{\mathbb{E}}[-\varphi(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})].\end{aligned}$$

In particular,

$$\hat{\mathbb{E}}[X(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})|\Omega_t] = X^+\hat{\mathbb{E}}[(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})] + X^-\hat{\mathbb{E}}[-(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})] = 0.$$

This, together with Proposition 3.2.9, yields

$$\hat{\mathbb{E}}[Y + X(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})|\Omega_t] = \hat{\mathbb{E}}[Y|\Omega_t], \quad Y \in L_G^1(\Omega).$$

We also have

$$\begin{aligned}\hat{\mathbb{E}}[X(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^2|\Omega_t] &= X^+\hat{\mathbb{E}}[(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^2] + X^-\hat{\mathbb{E}}[-(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^2]\\ &= [X^+\sigma^2_{\mathbf{aa}^T} - X^-\sigma^2_{-\mathbf{aa}^T}](T - t).\end{aligned}$$

For $n \in \mathbb{N}$,

$$\begin{aligned}\hat{\mathbb{E}}[X(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^{2n-1}|\Omega_t] &= X^+\hat{\mathbb{E}}[(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^{2n-1}] + X^-\hat{\mathbb{E}}[-(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^{2n-1}]\\ &= |X|\hat{\mathbb{E}}[(B_{T-t}^{\mathbf{a}})^{2n-1}].\end{aligned}$$

Example 3.2.11 Since

$$\hat{\mathbb{E}}[2B_s^{\mathbf{a}}(B_t^{\mathbf{a}} - B_s^{\mathbf{a}})|\Omega_s] = \hat{\mathbb{E}}[-2B_s^{\mathbf{a}}(B_t^{\mathbf{a}} - B_s^{\mathbf{a}})|\Omega_s] = 0,$$

we have

$$\begin{aligned}\hat{\mathbb{E}}[(B_t^{\mathbf{a}})^2-(B_s^{\mathbf{a}})^2|\Omega_s] &= \hat{\mathbb{E}}[(B_t^{\mathbf{a}}-B_s^{\mathbf{a}}+B_s^{\mathbf{a}})^2-(B_s^{\mathbf{a}})^2|\Omega_s]\\ &= \hat{\mathbb{E}}[(B_t^{\mathbf{a}}-B_s^{\mathbf{a}})^2+2(B_t^{\mathbf{a}}-B_s^{\mathbf{a}})B_s^{\mathbf{a}}|\Omega_s]\\ &= \sigma^2_{\mathbf{aa}^T}(t-s).\end{aligned}$$

3.3 Itô's Integral with Respect to G-Brownian Motion

For $T\in\mathbb{R}^+$, a partition π_T of $[0,T]$ is a finite ordered subset $\pi_T=\{t_0,t_1,\cdots,t_N\}$ such that $0=t_0<t_1<\cdots<t_N=T$. Set

$$\mu(\pi_T):=\max\{|t_{i+1}-t_i|\ :\ i=0,1,\cdots,N-1\}.$$

We use $\pi_T^N=\{t_0^N,t_1^N,\cdots,t_N^N\}$ to denote a sequence of partitions of $[0,T]$ such that $\lim_{N\to\infty}\mu(\pi_T^N)=0$.

Let $p\geq 1$ be fixed. We consider the following type of simple processes: for a given partition $\pi_T=\{t_0,\cdots,t_N\}$ of $[0,T]$ we set

$$\eta_t(\omega)=\sum_{k=0}^{N-1}\xi_k(\omega)\mathbf{1}_{[t_k,t_{k+1})}(t),$$

where $\xi_k\in L_G^p(\Omega_{t_k})$, $k=0,1,2,\cdots,N-1$ are given. The collection of these processes is denoted by $M_G^{p,0}(0,T)$.

Definition 3.3.1 For an $\eta\in M_G^{p,0}(0,T)$ with $\eta_t(\omega)=\sum_{k=0}^{N-1}\xi_k(\omega)\mathbf{1}_{[t_k,t_{k+1})}(t)$, the related **Bochner integral** is

$$\int_0^T\eta_t(\omega)dt:=\sum_{k=0}^{N-1}\xi_k(\omega)(t_{k+1}-t_k).$$

For each $\eta\in M_G^{p,0}(0,T)$, we set

$$\widetilde{\mathbb{E}}_T[\eta]:=\frac{1}{T}\hat{\mathbb{E}}\left[\int_0^T\eta_t dt\right]=\frac{1}{T}\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}\xi_k(\omega)(t_{k+1}-t_k)\right].$$

It is easy to check that $\widetilde{\mathbb{E}}_T:M_G^{p,0}(0,T)\mapsto\mathbb{R}$ forms a sublinear expectation. We then can introduce a natural norm $\|\cdot\|_{M_G^p}$, under which, $M_G^{p,0}(0,T)$ can be extended to $M_G^p(0,T)$ which is a Banach space.

Definition 3.3.2 For each $p \geq 1$, we denote by $M_G^p(0,T)$ the completion of $M_G^{p,0}(0,T)$ under the norm

$$\|\eta\|_{M_G^p(0,T)} := \left\{\hat{\mathbb{E}}\left[\int_0^T |\eta_t|^p dt\right]\right\}^{1/p}.$$

It is clear that $M_G^p(0,T) \supset M_G^q(0,T)$ for $1 \leq p \leq q$. We also denote by $M_G^p(0,T;\mathbb{R}^d)$ the space of all d-dimensional stochastic processes $\eta_t = (\eta_t^1, \cdots, \eta_t^d)$, $t \geq 0$ such that $\eta_t^i \in M_G^p(0,T)$, $i = 1, 2, \cdots, d$.

We now give the definition of Itô's integral. For simplicity, we first introduce Itô's integral with respect to 1-dimensional G–Brownian motion.

Let $(B_t)_{t\geq 0}$ be a 1-dimensional G–Brownian motion with $G(\alpha) = \frac{1}{2}(\overline{\sigma}^2\alpha^+ - \underline{\sigma}^2\alpha^-)$, where $0 \leq \underline{\sigma} \leq \overline{\sigma} < \infty$.

Definition 3.3.3 For each $\eta \in M_G^{2,0}(0,T)$ of the form

$$\eta_t(\omega) = \sum_{j=0}^{N-1} \xi_j \mathbf{1}_{[t_j,t_{j+1})}(t),$$

we define

$$I(\eta) = \int_0^T \eta_t dB_t := \sum_{j=0}^{N-1} \xi_j (B_{t_{j+1}} - B_{t_j}).$$

Lemma 3.3.4 *The mapping $I : M_G^{2,0}(0,T) \mapsto L_G^2(\Omega_T)$ is a continuous linear mapping and thus can be continuously extended to $I : M_G^2(0,T) \mapsto L_G^2(\Omega_T)$. In particular, we have*

$$\hat{\mathbb{E}}\left[\int_0^T \eta_t dB_t\right] = 0, \tag{3.3.1}$$

$$\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t\right)^2\right] \leq \overline{\sigma}^2 \hat{\mathbb{E}}\left[\int_0^T \eta_t^2 dt\right]. \tag{3.3.2}$$

Proof From Example 3.2.10, for each j,

$$\hat{\mathbb{E}}[\xi_j(B_{t_{j+1}} - B_{t_j})|\Omega_{t_j}] = \hat{\mathbb{E}}[-\xi_j(B_{t_{j+1}} - B_{t_j})|\Omega_{t_j}] = 0.$$

We have

$$\begin{aligned}\hat{\mathbb{E}}\left[\int_0^T \eta_t dB_t\right] &= \hat{\mathbb{E}}\left[\int_0^{t_{N-1}} \eta_t dB_t + \xi_{N-1}(B_{t_N} - B_{t_{N-1}})\right] \\ &= \hat{\mathbb{E}}\left[\int_0^{t_{N-1}} \eta_t dB_t + \hat{\mathbb{E}}[\xi_{N-1}(B_{t_N} - B_{t_{N-1}})|\Omega_{t_{N-1}}]\right]\end{aligned}$$

$$= \hat{\mathbb{E}}\left[\int_0^{t_{N-1}} \eta_t dB_t\right].$$

Then we can repeat this procedure to obtain (3.3.1).

We now give the proof of (3.3.2). First, from Example 3.2.10, we have

$$\begin{aligned}
\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t\right)^2\right] &= \hat{\mathbb{E}}\left[\left(\int_0^{t_{N-1}} \eta_t dB_t + \xi_{N-1}(B_{t_N} - B_{t_{N-1}})\right)^2\right] \\
&= \hat{\mathbb{E}}\left[\left(\int_0^{t_{N-1}} \eta_t dB_t\right)^2 + \xi_{N-1}^2(B_{t_N} - B_{t_{N-1}})^2\right. \\
&\quad \left. +2\left(\int_0^{t_{N-1}} \eta_t dB_t\right)\xi_{N-1}(B_{t_N} - B_{t_{N-1}})\right] \\
&= \hat{\mathbb{E}}\left[\left(\int_0^{t_{N-1}} \eta_t dB_t\right)^2 + \xi_{N-1}^2(B_{t_N} - B_{t_{N-1}})^2\right] \\
&= \cdots = \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \xi_i^2(B_{t_{i+1}} - B_{t_i})^2\right]
\end{aligned}$$

Then, for each $i = 0, 1, \cdots, N-1$, the following relations hold:

$$\begin{aligned}
&\hat{\mathbb{E}}[\xi_i^2(B_{t_{i+1}} - B_{t_i})^2 - \overline{\sigma}^2\xi_i^2(t_{i+1} - t_i)] \\
=&\hat{\mathbb{E}}[\hat{\mathbb{E}}[\xi_i^2(B_{t_{i+1}} - B_{t_i})^2 - \overline{\sigma}^2\xi_i^2(t_{i+1} - t_j)|\Omega_{t_i}]] \\
=&\hat{\mathbb{E}}[\overline{\sigma}^2\xi_i^2(t_{i+1} - t_i) - \overline{\sigma}^2\xi_i^2(t_{i+1} - t_i)] = 0.
\end{aligned}$$

Finally, we obtain

$$\begin{aligned}
&\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t\right)^2\right] = \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \xi_i^2(B_{t_{i+1}} - B_{t_i})^2\right] \\
\leq&\hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \xi_i^2(B_{t_{i+1}} - B_{t_i})^2 - \sum_{i=0}^{N-1} \overline{\sigma}^2\xi_i^2(t_{i+1} - t_i)\right] + \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \overline{\sigma}^2\xi_i^2(t_{i+1} - t_i)\right] \\
\leq&\sum_{i=0}^{N-1} \hat{\mathbb{E}}[\xi_i^2(B_{t_{i+1}} - B_{t_i})^2 - \overline{\sigma}^2\xi_i^2(t_{i+1} - t_i)] + \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \overline{\sigma}^2\xi_i^2(t_{i+1} - t_i)\right] \\
=&\hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \overline{\sigma}^2\xi_i^2(t_{i+1} - t_i)\right] = \overline{\sigma}^2\hat{\mathbb{E}}\left[\int_0^T \eta_t^2 dt\right].
\end{aligned}$$

□

Definition 3.3.5 For a fixed $\eta \in M_G^2(0,T)$, we define the stochastic integral

$$\int_0^T \eta_t dB_t := I(\eta).$$

It is clear that (3.3.1) and (3.3.2) still hold for $\eta \in M_G^2(0,T)$.

We list below the main properties of Itô's integral with respect to G-Brownian motion. We denote, for some $0 \leq s \leq t \leq T$,

$$\int_s^t \eta_u dB_u := \int_0^T \mathbf{1}_{[s,t]}(u)\eta_u dB_u.$$

Proposition 3.3.6 *Let $\eta, \theta \in M_G^2(0,T)$ and let $0 \leq s \leq r \leq t \leq T$. Then we have*

(i) $\int_s^t \eta_u dB_u = \int_s^r \eta_u dB_u + \int_r^t \eta_u dB_u$.

(ii) $\int_s^t (\alpha\eta_u + \theta_u) dB_u = \alpha \int_s^t \eta_u dB_u + \int_s^t \theta_u dB_u$, *if α is bounded and in $L_G^1(\Omega_s)$.*

(iii) $\hat{\mathbb{E}}\left[X + \int_r^T \eta_u dB_u \Big| \Omega_s\right] = \hat{\mathbb{E}}[X|\Omega_s]$ *for all $X \in L_G^1(\Omega)$.*

We now consider the multi-dimensional case. Let $G(\cdot) : \mathbb{S}(d) \mapsto \mathbb{R}$ be a given monotone and sublinear function and let $(B_t)_{t\geq 0}$ be a d-dimensional G-Brownian motion. For each fixed $\mathbf{a} \in \mathbb{R}^d$, we still use $B_t^{\mathbf{a}} := \langle \mathbf{a}, B_t \rangle$. Then $(B_t^{\mathbf{a}})_{t\geq 0}$ is a 1-dimensional $G_{\mathbf{a}}$-Brownian motion with $G_{\mathbf{a}}(\alpha) = \frac{1}{2}(\sigma^2_{\mathbf{aa}^T}\alpha^+ - \sigma^2_{-\mathbf{aa}^T}\alpha^-)$, where $\sigma^2_{\mathbf{aa}^T} = 2G(\mathbf{aa}^T)$ and $\sigma^2_{-\mathbf{aa}^T} = -2G(-\mathbf{aa}^T)$. Similarly to the 1-dimensional case, we can define Itô's integral by

$$I(\eta) := \int_0^T \eta_t dB_t^{\mathbf{a}}, \quad \text{for } \eta \in M_G^2(0,T).$$

We still have, for each $\eta \in M_G^2(0,T)$,

$$\hat{\mathbb{E}}\left[\int_0^T \eta_t dB_t^{\mathbf{a}}\right] = 0,$$

$$\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t^{\mathbf{a}}\right)^2\right] \leq \sigma^2_{\mathbf{aa}^T}\hat{\mathbb{E}}\left[\int_0^T \eta_t^2 dt\right].$$

Furthermore, Proposition 3.3.6 still holds for the integral with respect to $B_t^{\mathbf{a}}$.

3.4 Quadratic Variation Process of *G*-Brownian Motion

We first consider the quadratic variation process of 1-dimensional G-Brownian motion $(B_t)_{t\geq 0}$ with $B_1 \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$. Let π_t^N, $N = 1, 2, \cdots$, be a sequence of partitions of $[0,t]$. We start with the obvious relations

$$\begin{aligned} B_t^2 &= \sum_{j=0}^{N-1}(B_{t_{j+1}^N}^2 - B_{t_j^N}^2) \\ &= \sum_{j=0}^{N-1} 2B_{t_j^N}(B_{t_{j+1}^N} - B_{t_j^N}) + \sum_{j=0}^{N-1}(B_{t_{j+1}^N} - B_{t_j^N})^2. \end{aligned}$$

As $\mu(\pi_t^N) \to 0$, the first term of the right side converges to $2\int_0^t B_s dB_s$ in $L_G^2(\Omega)$. The second term must be convergent. We denote its limit by $\langle B\rangle_t$, i.e.,

$$\langle B\rangle_t := \lim_{\mu(\pi_t^N)\to 0} \sum_{j=0}^{N-1}(B_{t_{j+1}^N} - B_{t_j^N})^2 = B_t^2 - 2\int_0^t B_s dB_s. \tag{3.4.1}$$

By the above construction, $(\langle B\rangle_t)_{t\geq 0}$ is an increasing process with $\langle B\rangle_0 = 0$. We call it the **quadratic variation process** of the G-Brownian motion B. It characterizes the part of statistic uncertainty of G-Brownian motion. It is important to keep in mind that $\langle B\rangle_t$ is not a deterministic process unless $\underline{\sigma} = \overline{\sigma}$, i.e., when $(B_t)_{t\geq 0}$ is a classical Brownian motion. In fact, the following lemma is true.

Lemma 3.4.1 *For each* $0 \leq s \leq t < \infty$, *we have*

$$\hat{\mathbb{E}}[\langle B\rangle_t - \langle B\rangle_s|\Omega_s] = \overline{\sigma}^2(t-s), \tag{3.4.2}$$

$$\hat{\mathbb{E}}[-(\langle B\rangle_t - \langle B\rangle_s)|\Omega_s] = -\underline{\sigma}^2(t-s). \tag{3.4.3}$$

Proof By the definition of $\langle B\rangle$ and Proposition 3.3.6 (iii),

$$\begin{aligned} \hat{\mathbb{E}}[\langle B\rangle_t - \langle B\rangle_s|\Omega_s] &= \hat{\mathbb{E}}\left[B_t^2 - B_s^2 - 2\int_s^t B_u dB_u|\Omega_s\right] \\ &= \hat{\mathbb{E}}[B_t^2 - B_s^2|\Omega_s] = \overline{\sigma}^2(t-s). \end{aligned}$$

The last step follows from Example 3.2.11. We then have (3.4.2). The equality (3.4.3) can be proved analogously in view of the the relation $\hat{\mathbb{E}}[-(B_t^2 - B_s^2)|\Omega_s] = -\underline{\sigma}^2(t-s)$. □

Here is a very interesting property of the quadratic variation process $\langle B\rangle$, just like for the G–Brownian motion B itself: the increment $\langle B\rangle_{s+t} - \langle B\rangle_s$ is independent from Ω_s and identically distributed with $\langle B\rangle_t$. We formulate this as a lemma.

Lemma 3.4.2 *For each fixed* $s, t \geq 0$, $\langle B\rangle_{s+t} - \langle B\rangle_s$ *is identically distributed with* $\langle B\rangle_t$ *and is independent from* Ω_s, *for any* $s \geq 0$.

Proof The claims follow directly from

$$\begin{aligned}\langle B\rangle_{s+t}-\langle B\rangle_s &= B_{s+t}^2-2\int_0^{s+t}B_r dB_r-\left(B_s^2-2\int_0^s B_r dB_r\right)\\ &=(B_{s+t}-B_s)^2-2\int_s^{s+t}(B_r-B_s)d(B_r-B_s)\\ &=\langle B^s\rangle_t,\end{aligned}$$

where $\langle B^s\rangle$ is the quadratic variation process of the G-Brownian motion $B_t^s=B_{s+t}-B_s$, $t\geq 0$. □

We now define the integral of a process $\eta\in M_G^{1,0}(0,T)$ with respect to $\langle B\rangle$. We start with the mapping:

$$Q_{0,T}(\eta)=\int_0^T\eta_t d\langle B\rangle_t:=\sum_{j=0}^{N-1}\xi_j(\langle B\rangle_{t_{j+1}}-\langle B\rangle_{t_j}):M_G^{1,0}(0,T)\mapsto L_G^1(\Omega_T).$$

Lemma 3.4.3 *For each $\eta\in M_G^{1,0}(0,T)$,*

$$\hat{\mathbb{E}}[|Q_{0,T}(\eta)|]\leq\overline{\sigma}^2\hat{\mathbb{E}}\left[\int_0^T|\eta_t|dt\right]. \tag{3.4.4}$$

Thus $Q_{0,T}:M_G^{1,0}(0,T)\mapsto L_G^1(\Omega_T)$ is a continuous linear mapping. Consequently, $Q_{0,T}$ can be uniquely extended to $M_G^1(0,T)$. We still denote this mapping by

$$\int_0^T\eta_t d\langle B\rangle_t:=Q_{0,T}(\eta)\ \text{for}\ \eta\in M_G^1(0,T).$$

As before, the following relation holds:

$$\hat{\mathbb{E}}\left[\left|\int_0^T\eta_t d\langle B\rangle_t\right|\right]\leq\overline{\sigma}^2\hat{\mathbb{E}}\left[\int_0^T|\eta_t|dt\right]\ \text{for}\ \eta\in M_G^1(0,T). \tag{3.4.5}$$

Proof First, for each $j=1,\cdots,N-1$, we have

$$\begin{aligned}&\hat{\mathbb{E}}[|\xi_j|(\langle B\rangle_{t_{j+1}}-\langle B\rangle_{t_j})-\overline{\sigma}^2|\xi_j|(t_{j+1}-t_j)]\\ =&\hat{\mathbb{E}}[\hat{\mathbb{E}}[|\xi_j|(\langle B\rangle_{t_{j+1}}-\langle B\rangle_{t_j})|\Omega_{t_j}]-\overline{\sigma}^2|\xi_j|(t_{j+1}-t_j)]\\ =&\hat{\mathbb{E}}[|\xi_j|\overline{\sigma}^2(t_{j+1}-t_j)-\overline{\sigma}^2|\xi_j|(t_{j+1}-t_j)]=0.\end{aligned}$$

Then (3.4.4) can be checked as follows:

$$
\begin{aligned}
&\hat{\mathbb{E}}\left[\left|\sum_{j=0}^{N-1}\xi_j(\langle B\rangle_{t_{j+1}}-\langle B\rangle_{t_j})\right|\right] \le \hat{\mathbb{E}}\left[\sum_{j=0}^{N-1}|\xi_j|(\langle B\rangle_{t_{j+1}}-\langle B\rangle_{t_j})\right] \\
\le&\hat{\mathbb{E}}\left[\sum_{j=0}^{N-1}|\xi_j|[(\langle B\rangle_{t_{j+1}}-\langle B\rangle_{t_j})-\overline{\sigma}^2(t_{j+1}-t_j)]\right]+\hat{\mathbb{E}}\left[\overline{\sigma}^2\sum_{j=0}^{N-1}|\xi_j|(t_{j+1}-t_j)\right] \\
=&\hat{\mathbb{E}}\left[\overline{\sigma}^2\sum_{j=0}^{N-1}|\xi_j|(t_{j+1}-t_j)\right]=\overline{\sigma}^2\hat{\mathbb{E}}\left[\int_0^T|\eta_t|dt\right].
\end{aligned}
$$

□

Proposition 3.4.4 *Let* $0\le s\le t$, $\xi\in L_G^2(\Omega_s)$, $X\in L_G^1(\Omega)$. *Then*

$$
\begin{aligned}
\hat{\mathbb{E}}[X+\xi(B_t^2-B_s^2)]&=\hat{\mathbb{E}}[X+\xi(B_t-B_s)^2] \\
&=\hat{\mathbb{E}}[X+\xi(\langle B\rangle_t-\langle B\rangle_s)].
\end{aligned}
$$

Proof By (3.4.1) and Proposition 3.3.6 (iii), we have

$$
\begin{aligned}
\hat{\mathbb{E}}[X+\xi(B_t^2-B_s^2)]&=\hat{\mathbb{E}}\left[X+\xi(\langle B\rangle_t-\langle B\rangle_s+2\int_s^t B_u dB_u)\right] \\
&=\hat{\mathbb{E}}[X+\xi(\langle B\rangle_t-\langle B\rangle_s)].
\end{aligned}
$$

We also have

$$
\begin{aligned}
\hat{\mathbb{E}}[X+\xi(B_t^2-B_s^2)]&=\hat{\mathbb{E}}[X+\xi((B_t-B_s)^2+2(B_t-B_s)B_s)] \\
&=\hat{\mathbb{E}}[X+\xi(B_t-B_s)^2].
\end{aligned}
$$

□

We have the following isometry.

Proposition 3.4.5 *Let* $\eta\in M_G^2(0,T)$. *Then*

$$
\hat{\mathbb{E}}\left[\left(\int_0^T\eta_t dB_t\right)^2\right]=\hat{\mathbb{E}}\left[\int_0^T\eta_t^2 d\langle B\rangle_t\right]. \tag{3.4.6}
$$

Proof For any process $\eta\in M_G^{2,0}(0,T)$ of the form

$$
\eta_t(\omega)=\sum_{j=0}^{N-1}\xi_j(\omega)\mathbf{1}_{[t_j,t_{j+1})}(t)
$$

we have $\int_0^T \eta_t dB_t = \sum_{j=0}^{N-1} \xi_j (B_{t_{j+1}} - B_{t_j})$. From Proposition 3.3.6, we get

$$\hat{\mathbb{E}}[X + 2\xi_j (B_{t_{j+1}} - B_{t_j})\xi_i (B_{t_{i+1}} - B_{t_i})] = \hat{\mathbb{E}}[X], \quad \text{for all } X \in L_G^1(\Omega), i \neq j.$$

Thus

$$\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t\right)^2\right] = \hat{\mathbb{E}}[(\sum_{j=0}^{N-1} \xi_j (B_{t_{j+1}} - B_{t_j}))^2] = \hat{\mathbb{E}}\left[\sum_{j=0}^{N-1} \xi_j^2 (B_{t_{j+1}} - B_{t_j})^2\right].$$

From this and Proposition 3.4.4, it follows that

$$\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t\right)^2\right] = \hat{\mathbb{E}}\left[\sum_{j=0}^{N-1} \xi_j^2 (\langle B\rangle_{t_{j+1}} - \langle B\rangle_{t_j})\right] = \hat{\mathbb{E}}\left[\int_0^T \eta_t^2 d\langle B\rangle_t\right].$$

This shows that (3.4.6) holds for $\eta \in M_G^{2,0}(0, T)$. We can continuously extend the above equality to the case $\eta \in M_G^2(0, T)$ and get (3.4.6). □

We now consider the multi-dimensional case. Let $(B_t)_{t\geq 0}$ be a d-dimensional G-Brownian motion. For each fixed $\mathbf{a} \in \mathbb{R}^d$, $(B_t^{\mathbf{a}})_{t\geq 0}$ is a 1-dimensional $G_{\mathbf{a}}$-Brownian motion. Similar to the 1-dimensional case, we can define

$$\langle B^{\mathbf{a}}\rangle_t := \lim_{\mu(\pi_t^N)\to 0} \sum_{j=0}^{N-1} (B^{\mathbf{a}}_{t_{j+1}^N} - B^{\mathbf{a}}_{t_j^N})^2 = (B_t^{\mathbf{a}})^2 - 2\int_0^t B_s^{\mathbf{a}} dB_s^{\mathbf{a}},$$

where $\langle B^{\mathbf{a}}\rangle$ is called the **quadratic variation process** of $B^{\mathbf{a}}$. The above results, see Lemma 3.4.3 and Proposition 3.4.5, also hold for $\langle B^{\mathbf{a}}\rangle$. In particular,

$$\hat{\mathbb{E}}\left[\left|\int_0^T \eta_t d\langle B^{\mathbf{a}}\rangle_t\right|\right] \leq \sigma^2_{\mathbf{aa}^T} \hat{\mathbb{E}}\left[\int_0^T |\eta_t| dt\right], \quad \text{for all } \eta \in M_G^1(0, T)$$

and

$$\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t^{\mathbf{a}}\right)^2\right] = \hat{\mathbb{E}}\left[\int_0^T \eta_t^2 d\langle B^{\mathbf{a}}\rangle_t\right] \quad \text{for all } \eta \in M_G^2(0, T).$$

Let $\mathbf{a} = (a_1, \cdots, a_d)^T$ and $\bar{\mathbf{a}} = (\bar{a}_1, \cdots, \bar{a}_d)^T$ be two given vectors in $\mathbb{R}^d$. We then have their quadratic variation processes of $\langle B^{\mathbf{a}}\rangle$ and $\langle B^{\bar{\mathbf{a}}}\rangle$. We can define their **mutual variation process** by

$$\begin{aligned}\langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_t &:= \frac{1}{4}[\langle B^{\mathbf{a}} + B^{\bar{\mathbf{a}}}\rangle_t - \langle B^{\mathbf{a}} - B^{\bar{\mathbf{a}}}\rangle_t] \\ &= \frac{1}{4}[\langle B^{\mathbf{a}+\bar{\mathbf{a}}}\rangle_t - \langle B^{\mathbf{a}-\bar{\mathbf{a}}}\rangle_t].\end{aligned}$$

Since $\langle B^{\mathbf{a}-\bar{\mathbf{a}}}\rangle = \langle B^{\bar{\mathbf{a}}-\mathbf{a}}\rangle = \langle -B^{\mathbf{a}-\bar{\mathbf{a}}}\rangle$, we see that $\langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_t = \langle B^{\bar{\mathbf{a}}}, B^{\mathbf{a}}\rangle_t$. In particular, we have $\langle B^{\mathbf{a}}, B^{\mathbf{a}}\rangle = \langle B^{\mathbf{a}}\rangle$. Let π_t^N, $N = 1, 2, \cdots$, be a sequence of partitions of $[0, t]$. We observe that

$$\sum_{k=0}^{N-1}(B^{\mathbf{a}}_{t^N_{k+1}} - B^{\mathbf{a}}_{t^N_k})(B^{\bar{\mathbf{a}}}_{t^N_{k+1}} - B^{\bar{\mathbf{a}}}_{t^N_k}) = \frac{1}{4}\sum_{k=0}^{N-1}[(B^{\mathbf{a}+\bar{\mathbf{a}}}_{t_{k+1}} - B^{\mathbf{a}+\bar{\mathbf{a}}}_{t_k})^2 - (B^{\mathbf{a}-\bar{\mathbf{a}}}_{t_{k+1}} - B^{\mathbf{a}-\bar{\mathbf{a}}}_{t_k})^2].$$

As $\mu(\pi_t^N) \to 0$ we obtain

$$\lim_{N\to\infty}\sum_{k=0}^{N-1}(B^{\mathbf{a}}_{t^N_{k+1}} - B^{\mathbf{a}}_{t^N_k})(B^{\bar{\mathbf{a}}}_{t^N_{k+1}} - B^{\bar{\mathbf{a}}}_{t^N_k}) = \langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_t.$$

We also have

$$\begin{aligned}\langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_t &= \frac{1}{4}[\langle B^{\mathbf{a}+\bar{\mathbf{a}}}\rangle_t - \langle B^{\mathbf{a}-\bar{\mathbf{a}}}\rangle_t] \\ &= \frac{1}{4}\left[(B^{\mathbf{a}+\bar{\mathbf{a}}}_t)^2 - 2\int_0^t B^{\mathbf{a}+\bar{\mathbf{a}}}_s dB^{\mathbf{a}+\bar{\mathbf{a}}}_s - (B^{\mathbf{a}-\bar{\mathbf{a}}}_t)^2 + 2\int_0^t B^{\mathbf{a}-\bar{\mathbf{a}}}_s dB^{\mathbf{a}-\bar{\mathbf{a}}}_s\right] \\ &= B^{\mathbf{a}}_t B^{\bar{\mathbf{a}}}_t - \int_0^t B^{\mathbf{a}}_s dB^{\bar{\mathbf{a}}}_s - \int_0^t B^{\bar{\mathbf{a}}}_s dB^{\mathbf{a}}_s.\end{aligned}$$

Now for each $\eta \in M^1_G(0, T)$, we can consistently define

$$\int_0^T \eta_t d\langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_t := \frac{1}{4}\int_0^T \eta_t d\langle B^{\mathbf{a}+\bar{\mathbf{a}}}\rangle_t - \frac{1}{4}\int_0^T \eta_t d\langle B^{\mathbf{a}-\bar{\mathbf{a}}}\rangle_t.$$

Lemma 3.4.6 *Let $\eta^N \in M^{2,0}_G(0, T)$, $N = 1, 2, \cdots$, be of the form*

$$\eta^N_t(\omega) = \sum_{k=0}^{N-1}\xi^N_k(\omega)\mathbf{1}_{[t^N_k, t^N_{k+1})}(t)$$

with $\mu(\pi^N_T) \to 0$ and $\eta^N \to \eta$ in $M^2_G(0, T)$, as $N \to \infty$. Then we have the following convergence in $L^2_G(\Omega_T)$:

$$\sum_{k=0}^{N-1}\xi^N_k(B^{\mathbf{a}}_{t^N_{k+1}} - B^{\mathbf{a}}_{t^N_k})(B^{\bar{\mathbf{a}}}_{t^N_{k+1}} - B^{\bar{\mathbf{a}}}_{t^N_k}) \to \int_0^T \eta_t d\langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_t.$$

Proof Since

$$\langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_{t_{k+1}^N} - \langle B^{\mathbf{a}}, B^{\bar{\mathbf{a}}}\rangle_{t_k^N} = (B^{\mathbf{a}}_{t_{k+1}^N} - B^{\mathbf{a}}_{t_k^N})(B^{\bar{\mathbf{a}}}_{t_{k+1}^N} - B^{\bar{\mathbf{a}}}_{t_k^N}) \\ - \int_{t_k^N}^{t_{k+1}^N} (B^{\mathbf{a}}_s - B^{\mathbf{a}}_{t_k^N})dB^{\bar{\mathbf{a}}}_s - \int_{t_k^N}^{t_{k+1}^N} (B^{\bar{\mathbf{a}}}_s - B^{\bar{\mathbf{a}}}_{t_k^N})dB^{\mathbf{a}}_s,$$

we only need to show the convergence

$$\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}(\xi_k^N)^2\left(\int_{t_k^N}^{t_{k+1}^N}(B^{\mathbf{a}}_s - B^{\mathbf{a}}_{t_k^N})dB^{\bar{\mathbf{a}}}_s\right)^2\right] \to 0.$$

For each $k = 1, \cdots, N-1$, denoting $c = \overline{\sigma}^2_{\mathbf{a}\mathbf{a}^T}\overline{\sigma}^2_{\bar{\mathbf{a}}\bar{\mathbf{a}}^T}/2$, we have,

$$\begin{aligned}
&\hat{\mathbb{E}}\left[(\xi_k^N)^2\left(\int_{t_k^N}^{t_{k+1}^N}(B^{\mathbf{a}}_s - B^{\mathbf{a}}_{t_k^N})dB^{\bar{\mathbf{a}}}_s\right)^2 - c(\xi_k^N)^2(t_{k+1}^N - t_k^N)^2\right]\\
=&\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[(\xi_k^N)^2\left(\int_{t_k^N}^{t_{k+1}^N}(B^{\mathbf{a}}_s - B^{\mathbf{a}}_{t_k^N})dB^{\bar{\mathbf{a}}}_s\right)^2 |\Omega_{t_k^N}\right] - c(\xi_k^N)^2(t_{k+1}^N - t_k^N)^2\right]\\
\leq&\hat{\mathbb{E}}[c(\xi_k^N)^2(t_{k+1}^N - t_k^N)^2 - c(\xi_k^N)^2(t_{k+1}^N - t_k^N)^2] \to 0, \quad \text{as } N \to \infty.
\end{aligned}$$

Thus

$$\begin{aligned}
&\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}(\xi_k^N)^2\left(\int_{t_k^N}^{t_{k+1}^N}(B^{\mathbf{a}}_s - B^{\mathbf{a}}_{t_k^N})dB^{\bar{\mathbf{a}}}_s\right)^2\right]\\
\leq&\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}(\xi_k^N)^2\left[\left(\int_{t_k^N}^{t_{k+1}^N}(B^{\mathbf{a}}_s - B^{\mathbf{a}}_{t_k^N})dB^{\bar{\mathbf{a}}}_s\right)^2 - c(t_{k+1}^N - t_k^N)^2\right]\right]\\
&+\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}c(\xi_k^N)^2(t_{k+1}^N - t_k^N)^2\right]\\
\leq&\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}c(\xi_k^N)^2(t_{k+1}^N - t_k^N)^2\right] \leq c\mu(\pi_T^N)\hat{\mathbb{E}}\left[\int_0^T |\eta_t^N|^2 dt\right] \to 0, \quad \text{as} \quad N \to \infty.
\end{aligned}$$

□

3.5 Distribution of the Quadratic Variation Process $\langle B \rangle$

In this section, we first consider the 1-dimensional G-Brownian motion $(B_t)_{t\geq 0}$ with $B_1 \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$.

The quadratic variation process $\langle B \rangle$ of the **G**-Brownian motion B is a very interesting process. We have seen that the **G**-Brownian motion B is a typical process with variance uncertainty but without mean-uncertainty. In fact, all distributional uncertainty of the G-Brownian motion B is concentrated in $\langle B \rangle$. Moreover, $\langle B \rangle$ itself is a typical process with mean-uncertainty. This fact will be applied later to measure the mean-uncertainty of risk positions.

Lemma 3.5.1 *We have the following upper bound:*

$$\hat{\mathbb{E}}[\langle B \rangle_t^2] \leq 10\overline{\sigma}^4 t^2. \tag{3.5.1}$$

Proof Indeed,

$$\begin{aligned} \hat{\mathbb{E}}[\langle B \rangle_t^2] &= \hat{\mathbb{E}}\left[\left({B_t}^2 - 2\int_0^t B_u dB_u\right)^2\right] \\ &\leq 2\hat{\mathbb{E}}[B_t^4] + 8\hat{\mathbb{E}}\left[\left(\int_0^t B_u dB_u\right)^2\right] \\ &\leq 6\overline{\sigma}^4 t^2 + 8\overline{\sigma}^2 \hat{\mathbb{E}}\left[\int_0^t {B_u}^2 du\right] \\ &\leq 6\overline{\sigma}^4 t^2 + 8\overline{\sigma}^2 \int_0^t \hat{\mathbb{E}}[{B_u}^2] du \\ &= 10\overline{\sigma}^4 t^2. \end{aligned}$$

□

Proposition 3.5.2 *Let $(b_t)_{t\geq 0}$ be a d-dimensional Brownian motion on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ satisfying:*

(i) $b_0 = 0$*;*
(ii) For each $t, s \geq 0$, $b_{t+s} - b_t$ is identically distributed with b_s and independent from $(b_{t_1}, b_{t_2}, \cdots, b_{t_n})$ for all $0 \leq t_1, \cdots, t_n \leq t$*;*
(iii) $\lim_{t\downarrow 0} t^{-1}\hat{\mathbb{E}}[|b_t|^2] = 0$.

Then b_t is maximally distributed in the sense that:

$$\hat{\mathbb{E}}[\varphi(b_t)] = \max_{v\in\Gamma} \varphi(vt),$$

where Γ is the bounded closed and convex subset in $\mathbb{R}^d$ satisfying

$$\max_{v\in\Gamma}(p,v)=\hat{\mathbb{E}}[\langle p,b_1\rangle],\quad p\in\mathbb{R}^d.$$

In particular, if b *is* 1*-dimensional (*$d=1$*), then* $\Gamma=[\underline{\mu},\overline{\mu}]$*, with* $\overline{\mu}=\hat{\mathbb{E}}[b_1]$ *and* $\underline{\mu}=-\hat{\mathbb{E}}[-b_1]$.

Remark 3.5.3 Observe that for a symmetric G-Brownian motion B defined in Definition 3.1.2, the assumption corresponding to (iii) is: $\lim_{t\downarrow 0}\hat{\mathbb{E}}[|B_t|^3]t^{-1}=0$.

Proof We only give a proof for the case $d=1$ (see the proof of Theorem 3.8.2 for a more general situation). We first show that

$$\hat{\mathbb{E}}[pb_t]=g(p)t,\ \ p\in\mathbb{R}.$$

We set $\varphi(t):=\hat{\mathbb{E}}[b_t]$. Then $\varphi(0)=0$ and $\lim_{t\downarrow 0}\varphi(t)=0$. For each $t,s\geq 0$,

$$\begin{aligned}\varphi(t+s)=\hat{\mathbb{E}}[b_{t+s}]&=\hat{\mathbb{E}}[(b_{t+s}-b_s)+b_s]\\&=\varphi(t)+\varphi(s).\end{aligned}$$

Hence $\varphi(t)$ is linear and uniformly continuous in t, which means that $\hat{\mathbb{E}}[b_t]=\overline{\mu}t$. Similarly we obtain that $-\hat{\mathbb{E}}[-b_t]=\underline{\mu}t$.

We now prove that b_t is $N([\underline{\mu}t,\overline{\mu}t]\times\{0\})$-distributed. By Exercise 2.5.3 in Chap. 2, we just need to show that for each fixed $\varphi\in C_{b.Lip}(\mathbb{R})$, the function

$$u(t,x):=\hat{\mathbb{E}}[\varphi(x+b_t)],\ \ (t,x)\in[0,\infty)\times\mathbb{R}$$

is a viscosity solution of the following parabolic PDE:

$$\partial_t u-g(\partial_x u)=0,\quad u|_{t=0}=\varphi \tag{3.5.2}$$

with $g(a)=\overline{\mu}a^+-\underline{\mu}a^-$.

We first notice that u is Lipschitz in x and $\frac{1}{2}$-Hölder continuous in t. Indeed, for each fixed t, $u(t,\cdot)\in C_{b.Lip}(\mathbb{R})$ since

$$\begin{aligned}|\hat{\mathbb{E}}[\varphi(x+b_t)]-\hat{\mathbb{E}}[\varphi(y+b_t)]|&\leq\hat{\mathbb{E}}[|\varphi(x+b_t)-\varphi(y+b_t)|]\\&\leq C|x-y|.\end{aligned}$$

For each $\delta\in[0,t]$, since b_t-b_δ is independent from b_δ, we have

$$\begin{aligned}u(t,x)&=\hat{\mathbb{E}}[\varphi(x+b_\delta+(b_t-b_\delta)]\\&=\hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(y+(b_t-b_\delta))]_{y=x+b_\delta}],\end{aligned}$$

hence

$$u(t,x)=\hat{\mathbb{E}}[u(t-\delta,x+b_\delta)]. \tag{3.5.3}$$

Thus

$$\begin{aligned}|u(t,x)-u(t-\delta,x)| &= |\hat{\mathbb{E}}[u(t-\delta,x+b_\delta)-u(t-\delta,x)]|\\ &\le \hat{\mathbb{E}}[|u(t-\delta,x+b_\delta)-u(t-\delta,x)|]\\ &\le \hat{\mathbb{E}}[C|b_\delta|]\le C_1\sqrt{\delta}.\end{aligned}$$

To prove that u is a viscosity solution of the PDE (3.5.2), we fix a point $(t,x)\in(0,\infty)\times\mathbb{R}$ and let $v\in C_b^{2,2}([0,\infty)\times\mathbb{R})$ be such that $v\ge u$ and $v(t,x)=u(t,x)$. From (3.5.3), we find that

$$v(t,x)=\hat{\mathbb{E}}[u(t-\delta,x+b_\delta)]\le\hat{\mathbb{E}}[v(t-\delta,x+b_\delta)].$$

Therefore, by Taylor's expansion,

$$\begin{aligned}0&\le\hat{\mathbb{E}}[v(t-\delta,x+b_\delta)-v(t,x)]\\ &=\hat{\mathbb{E}}[v(t-\delta,x+b_\delta)-v(t,x+b_\delta)+(v(t,x+b_\delta)-v(t,x))]\\ &=\hat{\mathbb{E}}[-\partial_t v(t,x)\delta+\partial_x v(t,x)b_\delta+I_\delta]\\ &\le-\partial_t v(t,x)\delta+\hat{\mathbb{E}}[\partial_x v(t,x)b_\delta]+\hat{\mathbb{E}}[I_\delta]\\ &=-\partial_t v(t,x)\delta+g(\partial_x v(t,x))\delta+\hat{\mathbb{E}}[I_\delta],\end{aligned}$$

where

$$\begin{aligned}I_\delta=\delta&\int_0^1[-\partial_t v(t-\beta\delta,x+b_\delta)+\partial_t v(t,x)]d\beta\\ &+b_\delta\int_0^1[\partial_x v(t,x+\beta b_\delta)-\partial_x v(t,x)]d\beta.\end{aligned}$$

From the assumption that $\lim_{t\downarrow 0}t^{-1}\hat{\mathbb{E}}[b_t^2]=0$, we can check that

$$\lim_{\delta\downarrow 0}\delta^{-1}\hat{\mathbb{E}}[|I_\delta|]=0,$$

which implies that $\partial_t v(t,x)-g(\partial_x v(t,x))\le 0$. Hence u is a viscosity subsolution of (3.5.2). We can analogously prove that u is also a viscosity supersolution. It follows that b_t is $N([\underline{\mu}t,\overline{\mu}t]\times\{0\})$-distributed. □

It is clear that $\langle B\rangle$ satisfies all the conditions in Proposition 3.5.2, which leads immediately to another statement.

Theorem 3.5.4 *The process $\langle B\rangle_t$ is $N([\underline{\sigma}^2t,\overline{\sigma}^2t]\times\{0\})$-distributed, i.e.,*

$$\hat{\mathbb{E}}[\varphi(\langle B\rangle_t)]=\sup_{\underline{\sigma}^2\le v\le\overline{\sigma}^2}\varphi(vt),\quad \textit{for each}\ \ \varphi\in C_{l.Lip}(\mathbb{R}).\tag{3.5.4}$$

Corollary 3.5.5 *For each* $0 \le t \le T < \infty$, *we have*

$$\underline{\sigma}^2(T-t) \le \langle B\rangle_T - \langle B\rangle_t \le \overline{\sigma}^2(T-t) \ \ in\ L_G^1(\Omega).$$

Proof It is a direct consequence of the relations

$$\hat{\mathbb{E}}[(\langle B\rangle_T - \langle B\rangle_t - \overline{\sigma}^2(T-t))^+] = \sup_{\underline{\sigma}^2 \le v \le \overline{\sigma}^2} (v - \overline{\sigma}^2)^+(T-t) = 0$$

and

$$\hat{\mathbb{E}}[(\langle B\rangle_T - \langle B\rangle_t - \underline{\sigma}^2(T-t))^-] = \sup_{\underline{\sigma}^2 \le v \le \overline{\sigma}^2} (v - \underline{\sigma}^2)^-(T-t) = 0.$$

□

Corollary 3.5.6 *We have, for each* $t, s \ge 0$, $n \in \mathbb{N}$,

$$\hat{\mathbb{E}}[(\langle B\rangle_{t+s} - \langle B\rangle_s)^n|\Omega_s] = \hat{\mathbb{E}}[\langle B\rangle_t^n] = \overline{\sigma}^{2n}t^n \tag{3.5.5}$$

and

$$\hat{\mathbb{E}}[-(\langle B\rangle_{t+s} - \langle B\rangle_s)^n|\Omega_s] = \hat{\mathbb{E}}[-\langle B\rangle_t^n] = -\underline{\sigma}^{2n}t^n. \tag{3.5.6}$$

We now consider the multi-dimensional case. For notational simplicity, we write by $B^i := B^{\mathbf{e}_i}$ for the i-th coordinate of the G-Brownian motion B, under a given orthonormal basis $(\mathbf{e}_1, \cdots, \mathbf{e}_d)$ in the space $\mathbb{R}^d$. We denote

$$\langle B\rangle_t^{ij} := \langle B^i, B^j\rangle_t, \quad \langle B\rangle_t := (\langle B\rangle_t^{ij})_{i,j=1}^d.$$

Then $\langle B\rangle_t$, $t \ge 0$, is an $\mathbb{S}(d)$-valued process. Since

$$\hat{\mathbb{E}}[\langle AB_t, B_t\rangle] = 2G(A)\cdot t \ \text{ for } A \in \mathbb{S}(d),$$

we have

$$\begin{aligned}
\hat{\mathbb{E}}[(\langle B\rangle_t, A)] &= \hat{\mathbb{E}}\left[\sum_{i,j=1}^d a_{ij}\,\langle B\rangle_t^{ij}\right] \\
&= \hat{\mathbb{E}}\left[\sum_{i,j=1}^d a_{ij}\left(B_t^iB_t^j - \int_0^t B_s^i dB_s^j - \int_0^t B_s^j dB_s^i\right)\right] \\
&= \hat{\mathbb{E}}\left[\sum_{i,j=1}^d a_{ij}B_t^iB_t^j\right] = 2G(A)\cdot t \quad \text{for all} \quad A \in \mathbb{S}(d),
\end{aligned}$$

where $A = (a_{ij})_{i,j=1}^d$.

Now we set, for each $\varphi \in C_{l.Lip}(\mathbb{S}(d))$,

$$v(t,x) := \hat{\mathbb{E}}[\varphi(x+\langle B\rangle_t)], \quad (t,x)\in[0,\infty)\times\mathbb{S}(d).$$

Let $\Gamma \subset \mathbb{S}_+(d)$ be the bounded, convex and closed subset such that

$$G(A) = \frac{1}{2}\sup_{B\in\Gamma}(A,B), \qquad A\in\mathbb{S}(d).$$

Proposition 3.5.7 *The function v solves the following first order PDE:*

$$\partial_t v - 2G(Dv) = 0, \; v|_{t=0} = \varphi,$$

where $Dv = (\partial_{x_{ij}} v)_{i,j=1}^d$. *We also have*

$$v(t,x) = \sup_{\gamma\in\Gamma}\varphi(x+t\gamma).$$

Sketch of the Proof. We start with the relation

$$\begin{aligned} v(t+\delta,x) &= \hat{\mathbb{E}}[\varphi(x+\langle B\rangle_\delta + \langle B\rangle_{t+\delta} - \langle B\rangle_\delta)] \\ &= \hat{\mathbb{E}}[v(t,x+\langle B\rangle_\delta)]. \end{aligned}$$

The rest of the proof is similar to that in the 1-dimensional case. □

Corollary 3.5.8 *The following inclusion is true.*

$$\langle B\rangle_t \in t\Gamma := \{t\times\gamma : \gamma\in\Gamma\}.$$

This is equivalent to $d_{t\Gamma}(\langle B\rangle_t) = 0$, *where* $d_U(x) = \inf\{\sqrt{(x-y,x-y)} : y\in U\}$.

Proof Since

$$\hat{\mathbb{E}}[d_{t\Gamma}(\langle B\rangle_t)] = \sup_{\gamma\in\Gamma} d_{t\Gamma}(t\gamma) = 0,$$

it follows that $d_{t\Gamma}(\langle B\rangle_t) = 0$. □

3.6 Itô's Formula

In Theorem 3.6.5 of this section, we provide Itô's formula for a "G-Itô process" X. Let us begin with considering a sufficiently regular function Φ.

Lemma 3.6.1 *Let* $\Phi\in C^2(\mathbb{R}^n)$ *with* $\partial_{x^\nu}\Phi,\ \partial^2_{x^\mu x^\nu}\Phi \in C_{b.Lip}(\mathbb{R}^n)$ *for* $\mu,\nu = 1,\cdots,n$. *Let* $s\in[0,T]$ *be fixed and let* $X = (X^1,\cdots,X^n)^T$ *be an n–dimensional process on* $[s,T]$ *of the form*

$$X_t^\nu = X_s^\nu + \alpha^\nu (t-s) + \eta^{\nu ij}(\langle B\rangle_t^{ij} - \langle B\rangle_s^{ij}) + \beta^{\nu j}(B_t^j - B_s^j).$$

Here, for $\nu = 1, \cdots, n$, $i, j = 1, \cdots, d$, α^ν, $\eta^{\nu ij}$ and $\beta^{\nu j}$ are bounded elements in $L_G^2(\Omega_s)$ and $X_s = (X_s^1, \cdots, X_s^n)^T$ is a given random vector in $L_G^2(\Omega_s)$. Then we have, in $L_G^2(\Omega_t)$,

$$\Phi(X_t) - \Phi(X_s) = \int_s^t \partial_{x^\nu}\Phi(X_u)\beta^{\nu j} dB_u^j + \int_s^t \partial_{x^\nu}\Phi(X_u)\alpha^\nu du \qquad (3.6.1)$$
$$+ \int_s^t [\partial_{x^\nu}\Phi(X_u)\eta^{\nu ij} + \tfrac{1}{2}\partial^2_{x^\mu x^\nu}\Phi(X_u)\beta^{\mu i}\beta^{\nu j}] d\langle B\rangle_u^{ij}.$$

Here we adopt the Einstein convention, i.e., the above repeated indices μ, ν, i and j mean the summation.

Proof For any positive integer N, we set $\delta = (t-s)/N$ and take the partition

$$\pi_{[s,t]}^N = \{t_0^N, t_1^N, \cdots, t_N^N\} = \{s, s+\delta, \cdots, s+N\delta = t\}.$$

We have

$$\Phi(X_t) - \Phi(X_s) = \sum_{k=0}^{N-1}[\Phi(X_{t_{k+1}^N}) - \Phi(X_{t_k^N})] \qquad (3.6.2)$$
$$= \sum_{k=0}^{N-1}\{\partial_{x^\nu}\Phi(X_{t_k^N})(X_{t_{k+1}^N}^\nu - X_{t_k^N}^\nu)$$
$$+ \frac{1}{2}[\partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N})(X_{t_{k+1}^N}^\mu - X_{t_k^N}^\mu)(X_{t_{k+1}^N}^\nu - X_{t_k^N}^\nu) + \rho_k^N]\},$$

where

$$\rho_k^N = [\partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N} + \theta_k(X_{t_{k+1}^N} - X_{t_k^N})) - \partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N})](X_{t_{k+1}^N}^\mu - X_{t_k^N}^\mu)(X_{t_{k+1}^N}^\nu - X_{t_k^N}^\nu)$$

with $\theta_k \in [0, 1]$. The next is to derive that

$$\hat{\mathbb{E}}[|\rho_k^N|^2] = \hat{\mathbb{E}}[|[\partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N} + \theta_k(X_{t_{k+1}^N} - X_{t_k^N})) - \partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N})]$$
$$\times (X_{t_{k+1}^N}^\mu - X_{t_k^N}^\mu)(X_{t_{k+1}^N}^\nu - X_{t_k^N}^\nu)|^2]$$
$$\leq c\hat{\mathbb{E}}[|X_{t_{k+1}^N} - X_{t_k^N}|^6] \leq C[\delta^6 + \delta^3],$$

where c is the Lipschitz constant of $\{\partial^2_{x^\mu x^\nu}\Phi\}_{\mu,\nu=1}^n$ and C is a constant independent of k. Thus

$$\hat{\mathbb{E}}\left[\left|\sum_{k=0}^{N-1}\rho_k^N\right|^2\right] \le N\sum_{k=0}^{N-1}\hat{\mathbb{E}}\left[\left|\rho_k^N\right|^2\right] \to 0, \quad \text{as } N\to\infty.$$

The remaining terms in the summation in the right hand side of (3.6.2) are $\xi_t^N + \zeta_t^N$ with

$$\xi_t^N = \sum_{k=0}^{N-1}\{\partial_{x^\nu}\Phi(X_{t_k^N})[\alpha^\nu(t_{k+1}^N - t_k^N) + \eta^{\nu ij}(\langle B\rangle_{t_{k+1}^N}^{ij} - \langle B\rangle_{t_k^N}^{ij})$$
$$+ \beta^{\nu j}(B_{t_{k+1}^N}^j - B_{t_k^N}^j)] + \frac{1}{2}\partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N})\beta^{\mu i}\beta^{\nu j}(B_{t_{k+1}^N}^i - B_{t_k^N}^i)(B_{t_{k+1}^N}^j - B_{t_k^N}^j)\}$$

and

$$\zeta_t^N = \frac{1}{2}\sum_{k=0}^{N-1}\partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N})\{[\alpha^\mu(t_{k+1}^N - t_k^N) + \eta^{\mu ij}(\langle B\rangle_{t_{k+1}^N}^{ij} - \langle B\rangle_{t_k^N}^{ij})]$$
$$\times [\alpha^\nu(t_{k+1}^N - t_k^N) + \eta^{\nu lm}(\langle B\rangle_{t_{k+1}^N}^{lm} - \langle B\rangle_{t_k^N}^{lm})]$$
$$+ 2[\alpha^\mu(t_{k+1}^N - t_k^N) + \eta^{\mu ij}(\langle B\rangle_{t_{k+1}^N}^{ij} - \langle B\rangle_{t_k^N}^{ij})]\beta^{\nu l}(B_{t_{k+1}^N}^l - B_{t_k^N}^l)\}.$$

We observe that, for each $u \in [t_k^N, t_{k+1}^N)$,

$$\hat{\mathbb{E}}[|\partial_{x^\nu}\Phi(X_u) - \sum_{k=0}^{N-1}\partial_{x^\nu}\Phi(X_{t_k^N})\mathbf{1}_{[t_k^N, t_{k+1}^N)}(u)|^2]$$
$$= \hat{\mathbb{E}}[|\partial_{x^\nu}\Phi(X_u) - \partial_{x^\nu}\Phi(X_{t_k^N})|^2]$$
$$\le c^2\hat{\mathbb{E}}[|X_u - X_{t_k^N}|^2] \le C[\delta + \delta^2],$$

where c is the Lipschitz constant of $\{\partial_{x^\nu}\Phi\}_{\nu=1}^n$ and C is a constant independent of k. Hence $\sum_{k=0}^{N-1}\partial_{x^\nu}\Phi(X_{t_k^N})\mathbf{1}_{[t_k^N, t_{k+1}^N)}(\cdot)$ converges to $\partial_{x^\nu}\Phi(X_\cdot)$ in $M_G^2(0,T)$. Similarly, as $N\to\infty$,

$$\sum_{k=0}^{N-1}\partial^2_{x^\mu x^\nu}\Phi(X_{t_k^N})\mathbf{1}_{[t_k^N, t_{k+1}^N)}(\cdot) \to \partial^2_{x^\mu x^\nu}\Phi(X_\cdot) \quad \textit{in } M_G^2(0,T).$$

From Lemma 3.4.6 and by the definitions of integration with respect to dt, dB_t and $d\langle B\rangle_t$, the limit of ξ_t^N in $L_G^2(\Omega_t)$ is just the right hand side of (3.6.1). The next remark also leads to $\zeta_t^N \to 0$ in $L_G^2(\Omega_t)$. This completes the proof. □

Remark 3.6.2 To show that $\zeta_t^N \to 0$ in $L_G^2(\Omega_t)$, we use the following estimates: for each $\psi_\cdot^N = \sum_{k=0}^{N-1}\xi_{t_k}^N\mathbf{1}_{[t_k^N, t_{k+1}^N)}(\cdot) \in M_G^{2,0}(0,T)$ with $\pi_T^N = \{t_0^N, \cdots, t_N^N\}$ such that

$$\lim_{N\to\infty}\mu(\pi_T^N)=0 \quad \text{and} \quad \hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_{t_k}^N|^2(t_{k+1}^N-t_k^N)\right]\le C,$$

for all $N=1,2,\cdots$, we have

$$\hat{\mathbb{E}}\left[\left|\sum_{k=0}^{N-1}\xi_k^N(t_{k+1}^N-t_k^N)^2\right|^2\right]\to 0, \quad \text{as} \quad N\to\infty.$$

Moreover, for any fixed $\mathbf{a},\bar{\mathbf{a}}\in\mathbb{R}^d$,

$$\hat{\mathbb{E}}\left[\left|\sum_{k=0}^{N-1}\xi_k^N(\langle B^{\mathbf{a}}\rangle_{t_{k+1}^N}-\langle B^{\mathbf{a}}\rangle_{t_k^N})^2\right|^2\right]\le C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2(\langle B^{\mathbf{a}}\rangle_{t_{k+1}^N}-\langle B^{\mathbf{a}}\rangle_{t_k^N})^3\right]$$
$$\le C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2\sigma_{\mathbf{a}\mathbf{a}^T}^6(t_{k+1}^N-t_k^N)^3\right]\to 0,$$

$$\hat{\mathbb{E}}\left[\left|\sum_{k=0}^{N-1}\xi_k^N(\langle B^{\mathbf{a}}\rangle_{t_{k+1}^N}-\langle B^{\mathbf{a}}\rangle_{t_k^N})(t_{k+1}^N-t_k^N)\right|^2\right]$$
$$\le C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2(t_{k+1}^N-t_k^N)(\langle B^{\mathbf{a}}\rangle_{t_{k+1}^N}-\langle B^{\mathbf{a}}\rangle_{t_k^N})^2\right]$$
$$\le C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2\sigma_{\mathbf{a}\mathbf{a}^T}^4(t_{k+1}^N-t_k^N)^3\right]\to 0,$$

as well as

$$\hat{\mathbb{E}}\left[\left|\sum_{k=0}^{N-1}\xi_k^N(t_{k+1}^N-t_k^N)(B_{t_{k+1}^N}^{\mathbf{a}}-B_{t_k^N}^{\mathbf{a}})\right|^2\right]$$
$$\le C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2(t_{k+1}^N-t_k^N)|B_{t_{k+1}^N}^{\mathbf{a}}-B_{t_k^N}^{\mathbf{a}}|^2\right]$$
$$\le C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2\sigma_{\mathbf{a}\mathbf{a}^T}^2(t_{k+1}^N-t_k^N)^2\right]\to 0$$

and

$$\begin{aligned}
&\hat{\mathbb{E}}\left[\left|\sum_{k=0}^{N-1}\xi_k^N(\langle B^{\mathbf{a}}\rangle_{t_{k+1}^N}-\langle B^{\mathbf{a}}\rangle_{t_k^N})(B^{\bar{\mathbf{a}}}_{t_{k+1}^N}-B^{\bar{\mathbf{a}}}_{t_k^N})\right|^2\right]\\
\leq & C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2(\langle B^{\mathbf{a}}\rangle_{t_{k+1}^N}-\langle B^{\mathbf{a}}\rangle_{t_k^N})|B^{\bar{\mathbf{a}}}_{t_{k+1}^N}-B^{\bar{\mathbf{a}}}_{t_k^N}|^2\right]\\
\leq & C\hat{\mathbb{E}}\left[\sum_{k=0}^{N-1}|\xi_k^N|^2\sigma^2_{\mathbf{a}\mathbf{a}^T}\sigma^2_{\bar{\mathbf{a}}\bar{\mathbf{a}}^T}(t_{k+1}^N-t_k^N)^2\right]\to 0.
\end{aligned}$$

□

We are going now to derive a general form of Itô's formula. We start with

$$X_t^\nu = X_0^\nu + \int_0^t \alpha_s^\nu ds + \int_0^t \eta_s^{\nu ij} d\langle B\rangle_s^{ij} + \int_0^t \beta_s^{\nu j} dB_s^j, \quad \nu = 1,\cdots,n, \quad i,j=1,\cdots,d.$$

Proposition 3.6.3 *Let $\Phi \in C^2(\mathbb{R}^n)$ with $\partial_{x^\nu}\Phi$, $\partial^2_{x^\mu x^\nu}\Phi \in C_{b.Lip}(\mathbb{R}^n)$ for $\mu,\nu = 1,\cdots,n$. Let α^ν, $\beta^{\nu j}$ and $\eta^{\nu ij}$, $\nu = 1,\cdots,n$, $i,j=1,\cdots,d$ be bounded processes in $M_G^2(0,T)$. Then for each $t \geq 0$ we have, in $L_G^2(\Omega_t)$, that*

$$\begin{aligned}
\Phi(X_t)-\Phi(X_s) = &\int_s^t \partial_{x^\nu}\Phi(X_u)\beta_u^{\nu j}dB_u^j + \int_s^t \partial_{x^\nu}\Phi(X_u)\alpha_u^\nu du \qquad (3.6.3)\\
&+\int_s^t [\partial_{x^\nu}\Phi(X_u)\eta_u^{\nu ij} + \tfrac{1}{2}\partial^2_{x^\mu x^\nu}\Phi(X_u)\beta_u^{\mu i}\beta_u^{\nu j}]d\langle B\rangle_u^{ij}.
\end{aligned}$$

Proof We first consider the case of α, η and β being step processes of the form

$$\eta_t(\omega) = \sum_{k=0}^{N-1}\xi_k(\omega)\mathbf{1}_{[t_k,t_{k+1})}(t).$$

From Lemma 3.6.1, it is clear that (3.6.3) holds true. Now let

$$X_t^{\nu,N} = X_0^\nu + \int_0^t \alpha_s^{\nu,N}ds + \int_0^t \eta_s^{\nu ij,N}d\langle B\rangle_s^{ij} + \int_0^t \beta_s^{\nu j,N}dB_s^j,$$

where α^N, η^N and β^N are uniformly bounded step processes that converge to α, η and β in $M_G^2(0,T)$ as $N\to\infty$, respectively. From Lemma 3.6.1,

$$\Phi(X_t^N) - \Phi(X_s^N) = \int_s^t \partial_{x^\nu}\Phi(X_u^N)\beta_u^{\nu j,N} dB_u^j + \int_s^t \partial_{x^\nu}\Phi(X_u^N)\alpha_u^{\nu,N} du \qquad (3.6.4)$$
$$+ \int_s^t [\partial_{x^\nu}\Phi(X_u^N)\eta_u^{\nu ij,N} + \tfrac{1}{2}\partial^2_{x^\mu x^\nu}\Phi(X_u^N)\beta_u^{\mu i,N}\beta_u^{\nu j,N}] d\langle B\rangle_u^{ij}.$$

Since

$$\hat{\mathbb{E}}[|X_t^{\nu,N} - X_t^\nu|^2]$$
$$\leq C\hat{\mathbb{E}}\left[\int_0^T [(\alpha_s^{\nu,N} - \alpha_s^\nu)^2 + |\eta_s^{\nu,N} - \eta_s^\nu|^2 + |\beta_s^{\nu,N} - \beta_s^\nu|^2] ds\right],$$

where C is a constant independent of N. It follows that, in the space $M_G^2(0, T)$,

$$\begin{aligned}
\partial_{x^\nu}\Phi(X_\cdot^N)\eta_\cdot^{\nu ij,N} &\to \partial_{x^\nu}\Phi(X_\cdot)\eta_\cdot^{\nu ij},\\
\partial^2_{x^\mu x^\nu}\Phi(X_\cdot^N)\beta_\cdot^{\mu i,N}\beta_\cdot^{\nu j,N} &\to \partial^2_{x^\mu x^\nu}\Phi(X_\cdot)\beta_\cdot^{\mu i}\beta_\cdot^{\nu j},\\
\partial_{x^\nu}\Phi(X_\cdot^N)\alpha_\cdot^{\nu,N} &\to \partial_{x^\nu}\Phi(X_\cdot)\alpha_\cdot^{\nu},\\
\partial_{x^\nu}\Phi(X_\cdot^N)\beta_\cdot^{\nu j,N} &\to \partial_{x^\nu}\Phi(X_\cdot)\beta_\cdot^{\nu j}.
\end{aligned}$$

Therefore, passing to the limit as $N \to \infty$ in both sides of (3.6.4), we get (3.6.3). □

In order to derive Itô's formula for a general function Φ, we first establish a useful inequality. For the G-expectation $\hat{\mathbb{E}}$, we have the following representation (see Chap. 6):

$$\hat{\mathbb{E}}[X] = \sup_{P\in\mathcal{P}} E_P[X] \quad \text{for } X \in L_G^1(\Omega), \qquad (3.6.5)$$

where $\mathcal{P}$ is a weakly compact family of probability measures on $(\Omega, \mathcal{B}(\Omega))$.

Proposition 3.6.4 *Let $\beta \in M_G^p(0, T)$ with $p \geq 2$ and let $\mathbf{a} \in \mathbb{R}^d$ be fixed. Then we have $\int_0^T \beta_t dB_t^{\mathbf{a}} \in L_G^p(\Omega_T)$ and*

$$\hat{\mathbb{E}}\left[\left|\int_0^T \beta_t dB_t^{\mathbf{a}}\right|^p\right] \leq C_p\hat{\mathbb{E}}\left[\left|\int_0^T \beta_t^2 d\langle B^{\mathbf{a}}\rangle_t\right|^{p/2}\right]. \qquad (3.6.6)$$

Proof It suffices to consider the case where β is a step process of the form

$$\beta_t(\omega) = \sum_{k=0}^{N-1} \xi_k(\omega)\mathbf{1}_{[t_k,t_{k+1})}(t).$$

For each $\xi \in Lip(\Omega_t)$ with $t \in [0, T]$, we have

$$\hat{\mathbb{E}}\left[\xi\int_t^T \beta_s dB_s^{\mathbf{a}}\right] = 0.$$

From this we can easily get $E_P[\xi \int_t^T \beta_s dB_s^{\mathbf{a}}] = 0$ for each $P \in \mathcal{P}$, which implies that $(\int_0^t \beta_s dB_s^{\mathbf{a}})_{t\in 0,T]}$ is a P-martingale. Similarly we can prove that

$$M_t := \left(\int_0^t \beta_s dB_s^{\mathbf{a}}\right)^2 - \int_0^t \beta_s^2 d\langle B^{\mathbf{a}}\rangle_s, \quad t \in [0, T],$$

is a P-martingale for each $P \in \mathcal{P}$. By the Burkholder-Davis-Gundy inequalities, we have

$$E_P\left[\left|\int_0^T \beta_t dB_t^{\mathbf{a}}\right|^p\right] \leq C_p E_P\left[\left|\int_0^T \beta_t^2 d\langle B^{\mathbf{a}}\rangle_t\right|^{p/2}\right] \leq C_p \hat{\mathbb{E}}\left[\left|\int_0^T \beta_t^2 d\langle B^{\mathbf{a}}\rangle_t\right|^{p/2}\right],$$

where C_p is a universal constant independent of P. Thus we get (3.6.6). □

We now give the general G-Itô's formula.

Theorem 3.6.5 *Let Φ be a C^2-function on $\mathbb{R}^n$ such that $\partial^2_{x^\mu x^\nu}\Phi$ satisfies polynomial growth condition for $\mu, \nu = 1, \cdots, n$. Let α^ν, $\beta^{\nu j}$ and $\eta^{\nu ij}$, $\nu = 1, \cdots, n$, $i, j = 1, \cdots, d$ be bounded processes in $M_G^2(0, T)$. Then for each $t \geq 0$ we have in $L_G^2(\Omega_t)$*

$$\begin{aligned}\Phi(X_t) - \Phi(X_s) = &\int_s^t \partial_{x^\nu}\Phi(X_u)\beta_u^{\nu j} dB_u^j + \int_s^t \partial_{x^\nu}\Phi(X_u)\alpha_u^\nu du \\ &+ \int_s^t \left[\partial_{x^\nu}\Phi(X_u)\eta_u^{\nu ij} + \tfrac{1}{2}\partial^2_{x^\mu x^\nu}\Phi(X_u)\beta_u^{\mu i}\beta_u^{\nu j}\right] d\langle B\rangle_u^{ij}.\end{aligned} \tag{3.6.7}$$

Proof By the assumptions on Φ, we can choose a sequence of functions $\Phi_N \in C_0^2(\mathbb{R}^n)$ such that

$$|\Phi_N(x) - \Phi(x)| + |\partial_{x^\nu}\Phi_N(x) - \partial_{x^\nu}\Phi(x)| + |\partial^2_{x^\mu x^\nu}\Phi_N(x) - \partial^2_{x^\mu x^\nu}\Phi(x)| \leq \frac{C_1}{N}(1 + |x|^k),$$

where C_1 and k are positive constants independent of N. Obviously, Φ_N satisfies the conditions in Proposition 3.6.3, therefore,

$$\begin{aligned}\Phi_N(X_t) - \Phi_N(X_s) = &\int_s^t \partial_{x^\nu}\Phi_N(X_u)\beta_u^{\nu j} dB_u^j + \int_s^t \partial_{x^\nu}\Phi_N(X_u)\alpha_u^\nu du \\ &+ \int_s^t \left[\partial_{x^\nu}\Phi_N(X_u)\eta_u^{\nu ij} + \tfrac{1}{2}\partial^2_{x^\mu x^\nu}\Phi_N(X_u)\beta_u^{\mu i}\beta_u^{\nu j}\right] d\langle B\rangle_u^{ij}.\end{aligned} \tag{3.6.8}$$

For each fixed $T > 0$, by Proposition 3.6.4, there exists a constant C_2 such that

$$\hat{\mathbb{E}}[|X_t|^{2k}] \leq C_2 \quad \text{for } t \in [0, T].$$

Thus we can show that $\Phi_N(X_t) \to \Phi(X_t)$ as $N \to \infty$ in $L_G^2(\Omega_t)$ and, in $M_G^2(0, T)$,

$$\begin{aligned}\partial_{x^\nu}\Phi_N(X_\cdot)\eta_\cdot^{\nu ij} &\to \partial_{x^\nu}\Phi(X_\cdot)\eta_\cdot^{\nu ij},\\ \partial^2_{x^\mu x^\nu}\Phi_N(X_\cdot)\beta_\cdot^{\mu i}\beta_\cdot^{\nu j} &\to \partial^2_{x^\mu x^\nu}\Phi(X_\cdot)\beta_\cdot^{\mu i}\beta_\cdot^{\nu j},\\ \partial_{x^\nu}\Phi_N(X_\cdot)\alpha_\cdot^{\nu} &\to \partial_{x^\nu}\Phi(X_\cdot)\alpha_\cdot^{\nu},\\ \partial_{x^\nu}\Phi_N(X_\cdot)\beta_\cdot^{\nu j} &\to \partial_{x^\nu}\Phi(X_\cdot)\beta_\cdot^{\nu j}.\end{aligned}$$

We then can pass to limit as $N \to \infty$ in both sides of (3.6.8) to get (3.6.7). □

Corollary 3.6.6 *Let Φ be a polynomial and $\mathbf{a}, \mathbf{a}^\nu \in \mathbb{R}^d$ be fixed for $\nu = 1, \cdots, n$. Then we have*

$$\Phi(X_t) - \Phi(X_s) = \int_s^t \partial_{x^\nu}\Phi(X_u)dB_u^{\mathbf{a}^\nu} + \tfrac{1}{2}\int_s^t \partial^2_{x^\mu x^\nu}\Phi(X_u)d\left\langle B^{\mathbf{a}^\mu}, B^{\mathbf{a}^\nu}\right\rangle_u,$$

where $X_t = (B_t^{\mathbf{a}^1}, \cdots, B_t^{\mathbf{a}^n})^T$. In particular, we have, for $k = 2, 3, \cdots$,

$$(B_t^{\mathbf{a}})^k = k\int_0^t (B_s^{\mathbf{a}})^{k-1}dB_s^{\mathbf{a}} + \frac{k(k-1)}{2}\int_0^t (B_s^{\mathbf{a}})^{k-2}d\langle B^{\mathbf{a}}\rangle_s.$$

If the sublinear expectation $\hat{\mathbb{E}}$ becomes a linear expectation, then the above G-Itô's formula is the classical one.

3.7 Brownian Motion Without Symmetric Condition

In this section, we consider the Brownian motion B without the symmetric condition $\hat{\mathbb{E}}[B_t] = -\hat{\mathbb{E}}[-B_t]$ on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$. The following theorem gives a characterization of the Brownian motion without the symmetric condition.

Theorem 3.7.1 *Let $(B_t)_{t\geq 0}$ be a given $\mathbb{R}^d$–valued Brownian motion on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$. Then, for each fixed $\varphi \in C_{b.Lip}(\mathbb{R}^d)$, the function defined by,*

$$u(t, x) := \hat{\mathbb{E}}[\varphi(x + B_t)], \quad (t, x) \in [0, \infty) \times \mathbb{R}^d$$

is the unique viscosity solution of the following parabolic PDE:

$$\partial_t u - G(Du, D^2u) = 0, \quad u|_{t=0} = \varphi, \tag{3.7.1}$$

where

$$G(p, A) = \lim_{\delta\downarrow 0}\hat{\mathbb{E}}[\langle p, B_\delta\rangle + \tfrac{1}{2}\langle AB_\delta, B_\delta\rangle]\delta^{-1} \text{ for } (p, A) \in \mathbb{R}^d \times \mathbb{S}(d). \tag{3.7.2}$$

Proof We first prove that $\lim_{\delta\downarrow 0}\hat{\mathbb{E}}[\langle p, B_\delta\rangle + \frac{1}{2}\langle AB_\delta, B_\delta\rangle]\delta^{-1}$ exists. For each fixed $(p, A) \in \mathbb{R}^d \times \mathbb{S}(d)$, we set

$$f(t) := \hat{\mathbb{E}}[\langle p, B_t\rangle + \tfrac{1}{2}\langle AB_t, B_t\rangle].$$

Since

$$|f(t+h) - f(t)| \le \hat{\mathbb{E}}[(|p| + 2|A||B_t|)|B_{t+h} - B_t| + |A||B_{t+h} - B_t|^2] \to 0, \quad \text{as } h \to 0,$$

we get that $f(t)$ is a continuous function. Observe that

$$\hat{\mathbb{E}}[\langle q, B_t\rangle] = \hat{\mathbb{E}}[\langle q, B_1\rangle]t, \quad \text{for } q \in \mathbb{R}^d.$$

Thus for each $t, s > 0$,

$$|f(t+s) - f(t) - f(s)| \le C\hat{\mathbb{E}}[|B_t|]s,$$

where $C = |A|\hat{\mathbb{E}}[|B_1|]$. By (iii), there exists a constant $\delta_0 > 0$ such that $\hat{\mathbb{E}}[|B_t|^3] \le t$ for $t \le \delta_0$. Thus for each fixed $t > 0$ and $N \in \mathbb{N}$ such that $Nt \le \delta_0$, we have

$$|f(Nt) - Nf(t)| \le \frac{3}{4}C(Nt)^{4/3}.$$

From this and the continuity of f, it is easy to show that $\lim_{t\downarrow 0} f(t)t^{-1}$ exists. Thus we can get $G(p, A)$ for each $(p, A) \in \mathbb{R}^d \times \mathbb{S}(d)$. It is also easy to check that G is a continuous sublinear function monotone in $A \in \mathbb{S}(d)$.

Then we prove that u is Lipschitz in x and $\frac{1}{2}$-Hölder continuous in t. In fact, for each fixed t, $u(t, \cdot) \in C_{b.Lip}(\mathbb{R}^d)$ since

$$\begin{aligned}|\hat{\mathbb{E}}[\varphi(x + B_t)] - \hat{\mathbb{E}}[\varphi(y + B_t)]| &\le \hat{\mathbb{E}}[|\varphi(x + B_t) - \varphi(y + B_t)|]\\ &\le C|x - y|.\end{aligned}$$

For each $\delta \in [0, t]$, since $B_t - B_\delta$ is independent from B_δ,

$$\begin{aligned}u(t, x) &= \hat{\mathbb{E}}[\varphi(x + B_\delta + (B_t - B_\delta)]\\ &= \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(y + (B_t - B_\delta))]_{y=x+B_\delta}].\end{aligned}$$

Hence

$$u(t, x) = \hat{\mathbb{E}}[u(t - \delta, x + B_\delta)]. \tag{3.7.3}$$

Thus

$$\begin{aligned}|u(t,x)-u(t-\delta,x)| &= |\hat{\mathbb{E}}[u(t-\delta,x+B_\delta)-u(t-\delta,x)]| \\ &\le \hat{\mathbb{E}}[|u(t-\delta,x+B_\delta)-u(t-\delta,x)|] \\ &\le \hat{\mathbb{E}}[C|B_\delta|] \le C\sqrt{G(0,I)+1}\sqrt{\delta}.\end{aligned}$$

To prove that u is a viscosity solution of (3.7.1), we fix a pair $(t,x)\in(0,\infty)\times\mathbb{R}^d$ and let $v\in C_b^{2,3}([0,\infty)\times\mathbb{R}^d)$ be such that $v\ge u$ and $v(t,x)=u(t,x)$. From (3.7.3), we have

$$v(t,x)=\hat{\mathbb{E}}[u(t-\delta,x+B_\delta)]\le\hat{\mathbb{E}}[v(t-\delta,x+B_\delta)].$$

Therefore, by Taylor's expansion,

$$\begin{aligned}0 &\le \hat{\mathbb{E}}[v(t-\delta,x+B_\delta)-v(t,x)] \\ &= \hat{\mathbb{E}}[v(t-\delta,x+B_\delta)-v(t,x+B_\delta)+(v(t,x+B_\delta)-v(t,x))] \\ &= \hat{\mathbb{E}}[-\partial_t v(t,x)\delta+\langle Dv(t,x),B_\delta\rangle+\tfrac{1}{2}\langle D^2v(t,x)B_\delta,B_\delta\rangle+I_\delta] \\ &\le -\partial_t v(t,x)\delta+\hat{\mathbb{E}}[\langle Dv(t,x),B_\delta\rangle+\tfrac{1}{2}\langle D^2v(t,x)B_\delta,B_\delta\rangle]+\hat{\mathbb{E}}[I_\delta],\end{aligned}$$

where

$$\begin{aligned}I_\delta &= \int_0^1 -[\partial_t v(t-\beta\delta,x+B_\delta)-\partial_t v(t,x)]\delta d\beta \\ &\quad+\int_0^1\int_0^1\langle(D^2v(t,x+\alpha\beta B_\delta)-D^2v(t,x))B_\delta,B_\delta\rangle\alpha d\beta d\alpha.\end{aligned}$$

By condition (iii) in Definition 3.1.2, we can check that $\lim_{\delta\downarrow 0}\hat{\mathbb{E}}[|I_\delta|]\delta^{-1}=0$, which implies that $\partial_t v(t,x)-G(Dv(t,x),D^2v(t,x))\le 0$. Hence u is a viscosity subsolution of (3.7.1). We can analogously show that u is also a viscosity supersolution. Thus u is a viscosity solution. □

In many situations we are interested in a $2d$-dimensional Brownian motion $(B_t,b_t)_{t\ge0}$ such that $\hat{\mathbb{E}}[B_t]=-\hat{\mathbb{E}}[-B_t]=0$ and $\hat{\mathbb{E}}[|b_t|^2]/t\to 0$, as $t\downarrow 0$. In this case B is in fact a symmetric Brownian motion. Moreover, the process $(b_t)_{t\ge0}$ satisfies the properties in the Proposition 3.5.2. We define $u(t,x,y)=\hat{\mathbb{E}}[\varphi(x+B_t,y+b_t)]$. By Theorem 3.7.1 it follows that u is the solution of the PDE

$$\partial_t u=G(D_y u,D^2_{xx}u),\quad u|_{t=0}=\varphi\in C_{b.Lip}(\mathbb{R}^{2d}),$$

where G is a sublinear function of $(p,A)\in\mathbb{R}^d\times\mathbb{S}(d)$, defined by

$$G(p,A):=\hat{\mathbb{E}}[\langle p,b_1\rangle+\langle AB_1,B_1\rangle].$$

3.8 G-Brownian Motion Under (Not Necessarily Sublinear) Nonlinear Expectation

Let $\widetilde{\mathbb{E}}$ be a nonlinear expectation and $\hat{\mathbb{E}}$ be a sublinear expectation defined on $(\Omega, \mathcal{H})$ such that $\widetilde{\mathbb{E}}$ is dominated by $\hat{\mathbb{E}}$, namely

$$\widetilde{\mathbb{E}}[X] - \widetilde{\mathbb{E}}[Y] \leq \hat{\mathbb{E}}[X - Y], \quad X, Y \in \mathcal{H}.$$

We can also define a Brownian motion on the nonlinear expectation space $(\Omega, \mathcal{H}, \widetilde{\mathbb{E}})$. We emphasize that here the nonlinear expectation $\widetilde{\mathbb{E}}$ is not necessarily sublinear.

Definition 3.8.1 A d-dimensional process $(B_t)_{t\geq 0}$ on nonlinear expectation space $(\Omega, \mathcal{H}, \widetilde{\mathbb{E}})$ is called a Brownian motion if the following properties are satisfied:

(i) $B_0(\omega) = 0$;

(ii) For each $t, s \geq 0$, the increment $B_{t+s} - B_t$ is identically distributed with B_s and is independent from $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$, for each $n \in \mathbb{N}$ and $0 \leq t_1 \leq \cdots \leq t_n \leq t$;

(iii) $\lim_{t\downarrow 0} t^{-1}\hat{\mathbb{E}}[|B_t|^3] = 0$.

The following theorem gives a characterization of the nonlinear Brownian motion, and provides us with a new generator $\widetilde{G}$ associated with this more general nonlinear Brownian motion.

Theorem 3.8.2 *Let $(B_t, b_t)_{t\geq 0}$ be a given $\mathbb{R}^{2d}$–valued Brownian motion, both on $(\Omega, \mathcal{H}, \widetilde{\mathbb{E}})$ and $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ such that $\hat{\mathbb{E}}[B_t] = \hat{\mathbb{E}}[-B_t] = 0$ and $\lim_{t\to 0} \hat{\mathbb{E}}[|b_t|^2]/t = 0$. Assume that $\widetilde{\mathbb{E}}$ is dominated by $\hat{\mathbb{E}}$. Then, for each fixed $\varphi \in C_{b.Lip}(\mathbb{R}^{2d})$, the function*

$$\tilde{u}(t, x, y) := \widetilde{\mathbb{E}}[\varphi(x + B_t, y + b_t)], \quad (t, x, y) \in [0, \infty) \times \mathbb{R}^{2d}$$

is a viscosity solution of the following parabolic PDE:

$$\partial_t \tilde{u} - \widetilde{G}(D_y \tilde{u}, D_x^2 \tilde{u}) = 0, \quad \tilde{u}|_{t=0} = \varphi. \tag{3.8.1}$$

where

$$\widetilde{G}(p, A) = \widetilde{\mathbb{E}}[\langle p, b_1\rangle + \tfrac{1}{2}\langle AB_1, B_1\rangle], \ (p, A) \in \mathbb{R}^d \times \mathbb{S}(d). \tag{3.8.2}$$

Remark 3.8.3 Let

$$G(p, A) := \hat{\mathbb{E}}[\langle p, b_1\rangle + \tfrac{1}{2}\langle AB_1, B_1\rangle], \quad (p, A) \in \mathbb{R}^d \times \mathbb{S}(d). \tag{3.8.3}$$

Then the function $\widetilde{G}$ is dominated by the sublinear function G in the following sense:

$$\widetilde{G}(p, A) - \widetilde{G}(p', A') \leq G(p - p', A - A'), \ \ (p, A), \ (p', A') \in \mathbb{R}^d \times \mathbb{S}(d). \tag{3.8.4}$$

Conversely, once we have two functions G and $\widetilde{G}$ defined on $(\mathbb{R}^d, \mathbb{S}(d))$ such that G is a sublinear function and monotone in $A \in \mathbb{S}(d)$, and that $\widetilde{G}$ is dominated by G, we can construct a Brownian motion $(B_t, b_t)_{t\geq 0}$ on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ such that a nonlinear expectation $\widetilde{\mathbb{E}}$ is well-defined on $(\Omega, \mathcal{H})$ and is dominated by $\hat{\mathbb{E}}$. Moreover, under $\widetilde{\mathbb{E}}$, $(B_t, b_t)_{t\geq 0}$ is also a $\mathbb{R}^{2d}$-valued Brownian motion in the sense of Definition 3.8.1 and relations (3.8.2) and (3.8.3) are satisfied.

Proof of Theorem 3.8.2 We set

$$f(t) = f_{A,t}(t) := \widetilde{\mathbb{E}}[\langle p, b_t\rangle + \tfrac{1}{2}\langle AB_t, B_t\rangle],\ t \geq 0.$$

Since

$$\begin{aligned}|f(t+h) - f(t)| \leq &\hat{\mathbb{E}}[|p||b_{t+h} - b_t| + (|p| + 2|A||B_t|)|B_{t+h} - B_t| \\ &+ |A||B_{t+h} - B_t|^2] \to 0, \quad \text{as } h \to 0,\end{aligned}$$

we get that $f(t)$ is a continuous function. Since $\hat{\mathbb{E}}[B_t] = \hat{\mathbb{E}}[-B_t] = 0$, it follows from Proposition 3.8 that $\widetilde{\mathbb{E}}[X + \langle p, B_t\rangle] = \widetilde{\mathbb{E}}[X]$ for each $X \in \mathcal{H}$ and $p \in \mathbb{R}^d$. Thus

$$\begin{aligned}f(t+h) &= \widetilde{\mathbb{E}}[\langle p, b_{t+h} - b_t\rangle + \langle p, b_t\rangle \\ &\quad + \tfrac{1}{2}\langle AB_{t+h} - B_t, B_{t+h} - B_t\rangle + \tfrac{1}{2}\langle AB_t, B_t\rangle] \\ &= \widetilde{\mathbb{E}}[\langle p, b_h\rangle + \tfrac{1}{2}\langle AB_h, B_h\rangle] + \widetilde{\mathbb{E}}[\langle p, b_t\rangle + \tfrac{1}{2}\langle AB_t, B_t\rangle] \\ &= f(t) + f(h).\end{aligned}$$

It then follows that $f(t) = f(1)t = \widetilde{G}(A, p)t$. We now prove that the function $\tilde{u}$ is Lipschitz in x and uniformly continuous in t. Indeed, for each fixed t, $\tilde{u}(t, \cdot) \in C_{b.Lip}(\mathbb{R}^d)$ since

$$\begin{aligned}&|\widetilde{\mathbb{E}}[\varphi(x + B_t, y + b_t)] - \widetilde{\mathbb{E}}[\varphi(x' + B_t, y' + b_t)]| \\ &\leq \hat{\mathbb{E}}[|\varphi(x + B_t, y + b_t) - \varphi(x' + B_t, y' + b_t)|] \leq C(|x - x'| + |y - y'|).\end{aligned}$$

For each $\delta \in [0, t]$, since $(B_t - B_\delta, b_t - b_\delta)$ is independent from (B_δ, b_δ),

$$\begin{aligned}\tilde{u}(t, x, y) &= \widetilde{\mathbb{E}}[\varphi(x + B_\delta + (B_t - B_\delta), y + b_\delta + (b_t - b_\delta)] \\ &= \widetilde{\mathbb{E}}[\widetilde{\mathbb{E}}[\varphi(\bar{x} + (B_t - B_\delta), \bar{y} + (b_t - b_\delta))]_{\bar{x}=x+B_\delta, \bar{y}=y+b_\delta}].\end{aligned}$$

Hence

$$\tilde{u}(t, x, y) = \widetilde{\mathbb{E}}[\tilde{u}(t - \delta, x + B_\delta, y + b_\delta)]. \tag{3.8.5}$$

Thus

$$\begin{aligned}|\tilde{u}(t,x,y)-\tilde{u}(t-\delta,x,y)| &= |\widetilde{\mathbb{E}}[\tilde{u}(t-\delta,x+B_\delta,y+b_\delta)-\tilde{u}(t-\delta,x,y)]| \\ &\le \hat{\mathbb{E}}[|\tilde{u}(t-\delta,x+B_\delta,y+b_\delta)-\tilde{u}(t-\delta,x,y)|] \\ &\le C\hat{\mathbb{E}}[|B_\delta|+|b_\delta|].\end{aligned}$$

It follows from condition (iii) in Definition 3.8.1 that $\tilde{u}(t,x,y)$ is continuous in t uniformly in $(t,x,y)\in[0,\infty)\times\mathbb{R}^{2d}$.

To prove that $\tilde{u}$ is a viscosity solution of (3.8.1), we fix $(t,x,y)\in(0,\infty)\times\mathbb{R}^{2d}$ and let $v\in C_b^{2,3}([0,\infty)\times\mathbb{R}^{2d})$ be such that $v\ge u$ and $v(t,x,y)=\tilde{u}(t,x,y)$. From (3.8.5), we have

$$v(t,x,y)=\widetilde{\mathbb{E}}[\tilde{u}(t-\delta,x+B_\delta,y+b_\delta)]\le\widetilde{\mathbb{E}}[v(t-\delta,x+B_\delta,y+b_\delta)].$$

Therefore, by Taylor's expansion,

$$\begin{aligned}0 &\le \widetilde{\mathbb{E}}[v(t-\delta,x+B_\delta,y+b_\delta)-v(t,x,y)]\\ &= \widetilde{\mathbb{E}}[v(t-\delta,x+B_\delta,y+b_\delta)-v(t,x+B_\delta,y+b_\delta)+v(t,x+B_\delta,y+b_\delta)-v(t,x,y)]\\ &= \widetilde{\mathbb{E}}[-\partial_t v(t,x,y)\delta+\langle D_y v(t,x,y),b_\delta\rangle+\langle\partial_x v(t,x,y),B_\delta\rangle+\tfrac{1}{2}\langle D_{xx}^2 v(t,x,y)B_\delta,B_\delta\rangle+I_\delta]\\ &\le -\partial_t v(t,x,y)\delta+\widetilde{\mathbb{E}}[\langle D_y v(t,x,y),b_\delta\rangle+\tfrac{1}{2}\langle D_{xx}^2 v(t,x,y)B_\delta,B_\delta\rangle]+\hat{\mathbb{E}}[I_\delta].\end{aligned}$$

Here

$$\begin{aligned}I_\delta &= \int_0^1 -[\partial_t v(t-\delta\gamma,x+B_\delta,y+b_\delta)-\partial_t v(t,x,y)]\delta d\gamma\\ &+\int_0^1\langle\partial_y v(t,x+\gamma B_\delta,y+\gamma b_\delta)-\partial_y v(t,x,y),b_\delta\rangle d\gamma\\ &+\int_0^1\langle\partial_x v(t,x,y+\gamma b_\delta)-\partial_x v(t,x,y),B_\delta\rangle d\gamma\\ &+\int_0^1\int_0^1\langle(D_{xx}^2 v(t,x+\alpha\gamma B_\delta,y+\gamma b_\delta)-D_{xx}^2 v(t,x,y))B_\delta,B_\delta\rangle\gamma d\gamma d\alpha.\end{aligned}$$

We use assumption (iii) to check that $\lim_{\delta\downarrow 0}\hat{\mathbb{E}}[|I_\delta|]\delta^{-1}=0$. This implies that $\partial_t v(t,x)-\widetilde{G}(Dv(t,x),D^2v(t,x))\le 0$, hence u is a viscosity subsolution of (3.8.1). We can analogously prove that $\tilde{u}$ is a viscosity supersolution. Thus $\tilde{u}$ is a viscosity solution. □

3.9 Construction of Brownian Motions on a Nonlinear Expectation Space

Let $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a given continuous sublinear function monotone in $A \in \mathbb{S}(d)$. By Theorem 1.2.1 in Chap. 1, there exists a bounded, convex and closed subset $\Sigma \subset \mathbb{R}^d \times \mathbb{S}_+(d)$ such that

$$G(p, A) = \sup_{(q,B)\in\Sigma} [\tfrac{1}{2}\mathrm{tr}[AB] + \langle p, q\rangle] \quad \text{for } (p, A) \in \mathbb{R}^d \times \mathbb{S}(d).$$

By the results in Chap. 2, we know that there exists a pair of d-dimensional random vectors (X, Y) which is G-distributed.

Let $\widetilde{G}(\cdot) : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a given function dominated by G in the sense of (3.8.4). The construction of a $\mathbb{R}^{2d}$-dimensional Brownian motion $(B_t, b_t)_{t\geq 0}$ under a nonlinear expectation $\widetilde{\mathbb{E}}$, dominated by a sublinear expectation $\hat{\mathbb{E}}$ is based on a similar approach introduced in Sect. 3.2. In fact, we will see that by our construction $(B_t, b_t)_{t\geq 0}$ is also a Brownian motion under the sublinear expectation $\hat{\mathbb{E}}$.

We denote by $\Omega = C_0^{2d}(\mathbb{R}^+)$ the space of all $\mathbb{R}^{2d}$–valued continuous paths $(\omega_t)_{t\in\mathbb{R}^+}$. For each fixed $T \in [0, \infty)$, we set $\Omega_T := \{\omega_{\cdot\wedge T} : \omega \in \Omega\}$. We will consider the canonical process $(B_t, b_t)(\omega) = \omega_t, t \in [0, \infty)$, for $\omega \in \Omega$. We also follow Sect. 3.2 to introduce the spaces of random variables $Lip(\Omega_T)$ and $Lip(\Omega)$ so that to define the expectations $\hat{\mathbb{E}}$ and $\widetilde{\mathbb{E}}$ on $(\Omega, Lip(\Omega))$.

For this purpose we first construct a sequence of $2d$-dimensional random vectors $(X_i, \eta_i)_{i=1}^\infty$ on a sublinear expectation space $(\overline{\Omega}, \overline{\mathcal{H}}, \overline{\mathbb{E}})$ such that (X_i, η_i) is G-distributed and $(X_{i+1}, \eta_{\ i+1})$ is independent from $((X_1, \eta_1), \cdots, (X_i, \eta_i))$ for each $i = 1, 2, \cdots$. By the definition of G-distribution, the function

$$u(t, x, y) := \overline{\mathbb{E}}[\varphi(x + \sqrt{t}X_1, y + t\eta_1)], \quad t \geq 0, \quad x, y \in \mathbb{R}^d$$

is the viscosity solution of the following parabolic PDE, which is the same as Eq. (2.2.6) in Chap. 2:

$$\partial_t u - G(D_y u, D^2_{xx}u) = 0, \quad u|_{t=0} = \varphi \in C_{l.Lip}(\mathbb{R}^{2d}).$$

We also consider another PDE (see Theorem C.3.5 of Appendix C for the existence and uniqueness):

$$\partial_t \tilde{u} - \widetilde{G}(D_y \tilde{u}, D^2_{xx}\tilde{u}) = 0, \quad \tilde{u}|_{t=0} = \varphi \in C_{l.Lip}(\mathbb{R}^{2d}),$$

and denote $\widetilde{P}_t[\varphi](x, y) = \tilde{u}(t, x, y)$. Then it follows from Theorem C.3.5 in Appendix C, that, for each $\varphi, \psi \in C_{l.Lip}(\mathbb{R}^{2d})$,

$$\widetilde{P}_t[\varphi](x, y) - \widetilde{P}_t[\psi](x, y) \leq \overline{\mathbb{E}}[(\varphi - \psi)(x + \sqrt{t}X_1, y + t\eta_1)].$$

We now introduce a sublinear expectation $\hat{\mathbb{E}}$ and a nonlinear expectation $\widetilde{\mathbb{E}}$ both defined on $Lip(\Omega)$ via the following procedure: for each $X \in Lip(\Omega)$ with

$$X = \varphi(B_{t_1} - B_{t_0}, b_{t_1} - b_{t_0}, \cdots, B_{t_n} - B_{t_{n-1}}, b_{t_n} - b_{t_{n-1}})$$

for $\varphi \in C_{l.Lip}(\mathbb{R}^{2d\times n})$ and $0 = t_0 < t_1 < \cdots < t_n < \infty$, we define

$$\hat{\mathbb{E}}[\varphi(B_{t_1} - B_{t_0}, b_{t_1} - b_{t_0}, \cdots, B_{t_n} - B_{t_{n-1}}, b_{t_n} - b_{t_{n-1}})]$$

$$:= \overline{\mathbb{E}}[\varphi(\sqrt{t_1 - t_0}X_1, (t_1 - t_0)\eta_1, \cdots, \sqrt{t_n - t_{n-1}}X_n, (t_n - t_{n-1})\eta_n)].$$

Then we define

$$\widetilde{\mathbb{E}}[\varphi(B_{t_1} - B_{t_0}, b_{t_1} - b_{t_0}, \cdots, B_{t_n} - B_{t_{n-1}}, b_{t_n} - b_{t_{n-1}})] = \varphi_n,$$

where φ_n is obtained iteratively as follows:

$$\begin{aligned} \varphi_1(x_1, y_1, \cdots, x_{n-1}, y_{n-1}) &= \widetilde{P}_{t_n - t_{n-1}}[\varphi(x_1, y_1, \cdots, x_{n-1}, y_{n-1}, \cdot)](0, 0), \\ &\vdots \\ \varphi_{n-1}(x_1, y_1) &= \widetilde{P}_{t_2 - t_1}[\varphi_{n-2}(x_1, y_1, \cdot)](0, 0), \\ \varphi_n &= \widetilde{P}_{t_1}[\varphi_{n-1}(\cdot)](0, 0). \end{aligned}$$

The related conditional expectation of $X = \varphi(B_{t_1} - B_{t_0}, b_{t_1} - b_{t_0}, \cdots, B_{t_n} - B_{t_{n-1}}, b_{t_n} - b_{t_{n-1}})$ under Ω_{t_j} is defined by

$$\begin{aligned} \hat{\mathbb{E}}[X|\Omega_{t_j}] &= \hat{\mathbb{E}}[\varphi(B_{t_1} - B_{t_0}, b_{t_1} - b_{t_0}, \cdots, B_{t_n} - B_{t_{n-1}}, b_{t_n} - b_{t_{n-1}})\ |\Omega_{t_j}] \quad (3.9.1) \\ &:= \psi(B_{t_1} - B_{t_0}, b_{t_1} - b_{t_0}, \cdots, B_{t_j} - B_{t_{j-1}}, b_{t_j} - b_{t_{j-1}}), \end{aligned}$$

where

$$\psi(x_1, \cdots, x_j) = \overline{\mathbb{E}}[\varphi(x_1, \cdots, x_j, \sqrt{t_{j+1} - t_j}X_{j+1}, (t_1 - t_0)\eta_{j+1}, \cdots, \sqrt{t_n - t_{n-1}}X_n, (t_1 - t_0)\eta_n)].$$

Similarly,

$$\widetilde{\mathbb{E}}[X|\Omega_{t_j}] = \varphi_{n-j}(B_{t_1} - B_{t_0}, b_{t_1} - b_{t_0}, \cdots, B_{t_j} - B_{t_{j-1}}, b_{t_j} - b_{t_{j-1}}).$$

It is easy to check that $\hat{\mathbb{E}}[\cdot]$ (resp., $\widetilde{\mathbb{E}}$) consistently defines a sublinear (resp. nonlinear) expectation on $(\Omega, Lip(\Omega))$. Moreover $(B_t, b_t)_{t\geq 0}$ is a Brownian motion under both $\hat{\mathbb{E}}$ and $\widetilde{\mathbb{E}}$.

Proposition 3.9.1 *Let us list the properties of* $\widetilde{\mathbb{E}}[\cdot|\Omega_t]$ *that hold for each* $X, Y \in Lip(\Omega)$:

(i) *If* $X \geq Y$, *then* $\widetilde{\mathbb{E}}[X|\Omega_t] \geq \widetilde{\mathbb{E}}[Y|\Omega_t]$.
(ii) $\widetilde{\mathbb{E}}[X + \eta|\Omega_t] = \widetilde{\mathbb{E}}[X|\Omega_t] + \eta$, *for each* $t \geq 0$ *and* $\eta \in Lip(\Omega_t)$.
(iii) $\widetilde{\mathbb{E}}[X|\Omega_t] - \widetilde{\mathbb{E}}[Y|\Omega_t] \leq \hat{\mathbb{E}}[X - Y|\Omega_t]$.
(iv) $\widetilde{\mathbb{E}}[\widetilde{\mathbb{E}}[X|\Omega_t]|\Omega_s] = \widetilde{\mathbb{E}}[X|\Omega_{t\wedge s}]$, *in particular*, $\widetilde{\mathbb{E}}[\widetilde{\mathbb{E}}[X|\Omega_t]] = \widetilde{\mathbb{E}}[X]$.
(v) *For each* $X \in Lip(\Omega^t)$, $\widetilde{\mathbb{E}}[X|\Omega_t] = \widetilde{\mathbb{E}}[X]$, *where* $Lip(\Omega^t)$ *is the linear space of random variables of the form*

$$\varphi(B_{t_2} - B_{t_1}, b_{t_2} - b_{t_1}, \cdots, B_{t_{n+1}} - B_{t_n}, b_{t_{n+1}} - b_{t_n}),$$
$$n = 1, 2, \cdots, \ \varphi \in C_{l.Lip}(\mathbb{R}^{d\times n}), \ t_1, \cdots, t_n, t_{n+1} \in [t, \infty).$$

Since $\hat{\mathbb{E}}$ can be considered as a special nonlinear expectation of $\widetilde{\mathbb{E}}$ which is dominated by itself, it follows that $\hat{\mathbb{E}}[\cdot|\Omega_t]$ also satisfies the above properties (i)–(v).

Proposition 3.9.2 *The conditional sublinear expectation* $\hat{\mathbb{E}}\,[\cdot|\Omega_t]$ *satisfies (i)–(v). Moreover* $\hat{\mathbb{E}}[\cdot|\Omega_t]$ *itself is sublinear, i.e.,*
(vi) $\hat{\mathbb{E}}[X|\Omega_t] - \hat{\mathbb{E}}[Y|\Omega_t] \leq \hat{\mathbb{E}}[X - Y|\Omega_t]$, .
(vii) $\hat{\mathbb{E}}[\eta X|\Omega_t] = \eta^+\hat{\mathbb{E}}[X|\Omega_t] + \eta^-\hat{\mathbb{E}}[-X|\Omega_t]$ *for each* $\eta \in Lip(\Omega_t)$.

We now consider the completion of sublinear expectation space $(\Omega, Lip(\Omega), \hat{\mathbb{E}})$. Denote by $L_G^p(\Omega)$, $p \geq 1$, the completion of $Lip(\Omega)$ under the norm $\|X\|_p := (\hat{\mathbb{E}}[|X|^p])^{1/p}$. Similarly, we can define $L_G^p(\Omega_T)$, $L_G^p(\Omega_T^t)$ and $L_G^p(\Omega^t)$. It is clear that for each $0 \leq t \leq T < \infty$, $L_G^p(\Omega_t) \subseteq L_G^p(\Omega_T) \subseteq L_G^p(\Omega)$.

According to Sect. 1.4 in Chap. 1, the expectation $\hat{\mathbb{E}}[\cdot]$ can be continuously extended to $(\Omega, L_G^1(\Omega))$. Moreover, since the nonlinear expectation $\widetilde{\mathbb{E}}$ is dominated by $\hat{\mathbb{E}}$, it can also be continuously extended to $(\Omega, L_G^1(\Omega))$. $(\Omega, L_G^1(\Omega), \hat{\mathbb{E}})$ is a sublinear expectation space while $(\Omega, L_G^1(\Omega), \widetilde{\mathbb{E}})$ is a nonlinear expectation space. We refer to Definition 1.4.4 in Chap. 1.

The next is to look for the extension of conditional expectation. For each fixed $t \leq T$, the conditional expectation $\widetilde{\mathbb{E}}[\cdot|\Omega_t] : Lip(\Omega_T) \mapsto Lip(\Omega_t)$ is a continuous mapping under $\|\cdot\|$. Indeed, we have

$$\widetilde{\mathbb{E}}[X|\Omega_t] - \widetilde{\mathbb{E}}[Y|\Omega_t] \leq \hat{\mathbb{E}}[X - Y|\Omega_t] \leq \hat{\mathbb{E}}[|X - Y||\Omega_t],$$

then

$$|\widetilde{\mathbb{E}}[X|\Omega_t] - \widetilde{\mathbb{E}}[Y|\Omega_t]| \leq \hat{\mathbb{E}}[|X - Y||\Omega_t].$$

We thus obtain

$$\left\|\widetilde{\mathbb{E}}[X|\Omega_t] - \widetilde{\mathbb{E}}[Y|\Omega_t]\right\| \leq \|X - Y\|.$$

It follows that $\widetilde{\mathbb{E}}[\cdot|\Omega_t]$ can also be extended as a continuous mapping

$$\widetilde{\mathbb{E}}[\cdot|\Omega_t] : L_G^1(\Omega_T) \mapsto L_G^1(\Omega_t).$$

If the parameter T is not fixed, then we can obtain $\widetilde{\mathbb{E}}[\cdot|\Omega_t] : L_G^1(\Omega) \mapsto L_G^1(\Omega_t)$.

Remark 3.9.3 Propositions 3.9.1 and 3.9.2 also hold for $X, Y \in L_G^1(\Omega)$. However, in (iv), $\eta \in L_G^1(\Omega_t)$ should be bounded, since $X, Y \in L_G^1(\Omega)$ does not imply that $X \cdot Y \in L_G^1(\Omega)$.

In particular, we have the following independence:

$$\widetilde{\mathbb{E}}[X|\Omega_t] = \widetilde{\mathbb{E}}[X], \quad \forall X \in L_G^1(\Omega^t).$$

We give the following definition similar to the classical one:

Definition 3.9.4 An n-dimensional random vector $Y \in (L_G^1(\Omega))^n$ is said to be independent from Ω_t for some given t if for each $\varphi \in C_{b.Lip}(\mathbb{R}^n)$ we have

$$\widetilde{\mathbb{E}}[\varphi(Y)|\Omega_t] = \widetilde{\mathbb{E}}[\varphi(Y)].$$

3.10 Exercises

Exercise 3.10.1 Let $(B_t)_{t\geq 0}$ be a 1-dimensional G-Brownian motion, such that its value at $t = 1$ is $B_1 \overset{d}{=} N(\{0\} \times [\underline{\sigma}^2, \overline{\sigma}^2])$. Prove that for each $m \in \mathbb{N}$,

$$\hat{\mathbb{E}}[|B_t|^m] = \begin{cases} 2(m-1)!!\overline{\sigma}^m t^{\frac{m}{2}}/\sqrt{2\pi}, & \text{if } m \text{ is odd,} \\ (m-1)!!\overline{\sigma}^m t^{\frac{m}{2}}, & \text{if } m \text{ is even.} \end{cases}$$

Exercise 3.10.2 Show that if $X \in Lip(\Omega_T)$ and $\hat{\mathbb{E}}[X] = -\hat{\mathbb{E}}[-X]$, then $\hat{\mathbb{E}}[X] = E_P[X]$, where P is a Wiener measure on Ω.

Exercise 3.10.3 For each $s, t \geq 0$, we set $B_t^s := B_{t+s} - B_s$. Let $\eta = (\eta_{ij})_{i,j=1}^d \in L_G^1(\Omega_s; \mathbb{S}(d))$. Prove that

$$\hat{\mathbb{E}}[\langle \eta B_t^s, B_t^s \rangle|\Omega_s] = 2G(\eta)t.$$

Exercise 3.10.4 Suppose that $X \in L_G^p(\Omega_T)$ for $p \geq 1$. Prove that there exists a sequence of bounded random variables $X_n \in Lip(\Omega_T)$, $n = 1, \cdots$, such that

$$\lim_{n\to\infty} \hat{\mathbb{E}}[|X - X_n|^p] = 0.$$

Exercise 3.10.5 Prove that for each $X \in Lip(\Omega_T)$, $\sup_{0\leq t\leq T} \hat{\mathbb{E}}_t[X] \in L_G^1(\Omega_T)$.

Exercise 3.10.6 Prove that $\varphi(B_t) \in L_G^1(\Omega_t)$ for each $\varphi \in C(\mathbb{R}^d)$ with a polynomial growth.

Exercise 3.10.7 Prove that, for a fixed $\eta \in M_G^2(0, T)$,

$$\underline{\sigma}^2 \hat{\mathbb{E}}\left[\int_0^T \eta_t^2 dt\right] \leq \hat{\mathbb{E}}\left[(\int_0^T \eta_t dB_t)^2\right] \leq \overline{\sigma}^2 \hat{\mathbb{E}}\left[\int_0^T \eta_t^2 dt\right],$$

where $\overline{\sigma}^2 = \hat{\mathbb{E}}[B_1^2]$ and $\underline{\sigma}^2 = -\hat{\mathbb{E}}[-B_1^2]$.

Exercise 3.10.8 Let $(B_t)_{t\geq 0}$ be a 1-dimensional G-Brownian motion and φ a bounded and Lipschitz function on $\mathbb{R}$. Show that

$$\lim_{N\to\infty} \hat{\mathbb{E}}\left[\left|\sum_{k=0}^{N-1} \varphi(B_{t_k^N})[(B_{t_{k+1}^N} - B_{t_k^N})^2 - (\langle B\rangle_{t_{k+1}^N} - \langle B\rangle_{t_k^N})]\right|\right] = 0,$$

where $t_k^N = kT/N, k = 0, \ldots, N-1$.

Exercise 3.10.9 Prove that, for a fixed $\eta \in M_G^1(0, T)$,

$$\underline{\sigma}^2 \hat{\mathbb{E}}\left[\int_0^T |\eta_t| dt\right] \leq \hat{\mathbb{E}}\left[\int_0^T |\eta_t| d\langle B\rangle_t\right] \leq \overline{\sigma}^2 \hat{\mathbb{E}}\left[\int_0^T |\eta_t| dt\right],$$

where $\overline{\sigma}^2 = \hat{\mathbb{E}}[B_1^2]$ and $\underline{\sigma}^2 = -\hat{\mathbb{E}}[-B_1^2]$.

Exercise 3.10.10 Complete the proof of Proposition 3.5.7.

Exercise 3.10.11 Let B be a 1-dimensional G-Brownian motion and $\widetilde{\mathbb{E}}$ a nonlinear expectation dominated by a G-expectation. Show that for any $\eta \in M_G^2(0, T)$:

(i) $\widetilde{\mathbb{E}}\left[\int_0^T \eta_s dB_s\right] = 0$;
(ii) $\widetilde{\mathbb{E}}\left[\left(\int_0^T \eta_s dB_s\right)^2\right] = \widetilde{\mathbb{E}}\left[\int_0^T |\eta_s|^2 d\langle B\rangle_s\right]$.

Notes and Comments

Bachelier [7] proposed to use the Brownian motion as a model of the fluctuations of stock markets. Independently, Einstein [56] used the Brownian motion to give experimental confirmation of the atomic theory, and Wiener [173] gave a mathematically rigorous construction of the Brownian motion. Here we follow Kolmogorov's idea [103] to construct G-Brownian motions by introducing finite dimensional cylinder function space and the corresponding family of infinite dimensional sublinear distributions, instead of (linear) probability distributions used in [103].

The notions of G-Brownian motions and the related stochastic calculus of Itô's type were firstly introduced by Peng [138] for the 1-dimensional case and then in

(2008) [141] for the multi-dimensional situation. It is very interesting that Denis and Martini [48] studied super-pricing of contingent claims under model uncertainty of volatility. They have introduced a norm in the space of continuous paths $\Omega = C([0, T])$ which corresponds to the L_G^2-norm and developed a stochastic integral. In that paper there are no notions such as nonlinear expectation and the related nonlinear distribution, G-expectation, conditional G-expectation, the related G-normal distribution and independence. On the other hand, by using powerful tools from capacity theory these authors obtained pathwise results for random variables and stochastic processes through the language of "quasi-surely" (see e.g. Dellacherie [42], Dellacherie and Meyer [43], Feyel and de La Pradelle [65]) in place of "almost surely" in classical probability theory.

One of the main motivations to introduce the notion of G-Brownian motions was the necessity to deal with pricing and risk measures under volatility uncertainty in financial markets (see Avellaneda, Lévy and Paras [6] and Lyons [114]). It was well-known that under volatility uncertainty the corresponding uncertain probability measures are singular with respect to each other. This causes a serious problem in the related path analysis to treat, e.g., when dealing with path-dependent derivatives, under a classical probability space. The notion of G-Brownian motions provides a powerful tool to study such a type of problems. Indeed, Biagini, Mancin and Meyer Brandis studied mean-variance hedging under the G-expectation framework in [18]. Fouque, Pun and Wong investigated the asset allocation problem among a risk-free asset and two risky assets with an ambiguous correlation through the theory of G-Brownian motions in [67]. We also remark that Beissner and Riedel [15] studied equilibria under Knightian price uncertainty through sublinear expectation theory, see also [14, 16].

The new Itô's calculus with respect to G-Brownian motion was inspired by Itô's groundbreaking work of [92] on stochastic integration, stochastic differential equations followed by a huge progress in stochastic calculus. We refer to interesting books cited in Chap. 4. Itô's formula given by Theorem 3.6.5 is from [138, 141]. Gao [72] proved a more general Itô's formula for G-Brownian motion. On this occasion an interesting problem appeared: can we establish an Itô's formula under conditions which correspond to the classical one? This problem will be solved in Chap. 8 with quasi surely analysis approach.

Using nonlinear Markovian semigroups known as Nisio's semigroups (see Nisio [119]), Peng [136] studied the processes with Markovian properties under a nonlinear expectation. Denk, Kupper and Nendel studied the relation between Lévy processes under nonlinear expectations, nonlinear semigroups and fully nonlinear PDEs, see [50].

Chapter 4
G-Martingales and Jensen's Inequality

In this chapter, we introduce the notion of G-martingales and the related Jensen's inequality for a new type of G-convex functions. One essential difference from the classical situation is that here "M is a G-martingale" does not imply that "$-M$ is a G-martingale".

4.1 The Notion of G-Martingales

We now give the notion of G-martingales.

Definition 4.1.1 A process $(M_t)_{t\geq 0}$ is called a ***G-supermartingale*** (respectively, ***G-submartingale***) if for any $t \in [0, \infty)$, $M_t \in L_G^1(\Omega_t)$ and for any $s \in [0, t]$, we have

$$\hat{\mathbb{E}}[M_t|\Omega_s] \leq M_s \quad \text{(respectively, } \geq M_s).$$

$(M_t)_{t\geq 0}$ is called a ***G-martingale*** if it is both G-supermartingale and G-submartingale. If a G-martingale M satisfies also

$$\hat{\mathbb{E}}[-M_t|\Omega_s] = -M_s,$$

then it is called a ***symmetric G–martingale***.

Example 4.1.2 For any fixed $X \in L_G^1(\Omega)$, it is clear that $(\hat{\mathbb{E}}[X|\Omega_t])_{t\geq 0}$ is a G–martingale.

Example 4.1.3 For any fixed $\mathbf{a} \in \mathbb{R}^d$, it is easy to check that $(B_t^{\mathbf{a}})_{t\geq 0}$ and $(-B_t^{\mathbf{a}})_{t\geq 0}$ are G–martingales. The process $(\langle B^{\mathbf{a}}\rangle_t - \sigma^2_{\mathbf{a}\mathbf{a}^T} t)_{t\geq 0}$ is a G-martingale since

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_4

$$\begin{aligned}\hat{\mathbb{E}}[\langle B^{\mathbf{a}}\rangle_t - \sigma^2_{\mathbf{aa}^T} t|\Omega_s] &= \hat{\mathbb{E}}[\langle B^{\mathbf{a}}\rangle_s - \sigma^2_{\mathbf{aa}^T} t + (\langle B^{\mathbf{a}}\rangle_t - \langle B^{\mathbf{a}}\rangle_s)|\Omega_s] \\ &= \langle B^{\mathbf{a}}\rangle_s - \sigma^2_{\mathbf{aa}^T} t + \hat{\mathbb{E}}[\langle B^{\mathbf{a}}\rangle_t - \langle B^{\mathbf{a}}\rangle_s] \\ &= \langle B^{\mathbf{a}}\rangle_s - \sigma^2_{\mathbf{aa}^T} s.\end{aligned}$$

However, the processes $(-(\langle B^{\mathbf{a}}\rangle_t - \sigma^2_{\mathbf{aa}^T} t))_{t\geq 0}$ and $((B^{\mathbf{a}}_t)^2)_{t\geq 0}$ are G-submartingales, as seen from the relations

$$\begin{aligned}\hat{\mathbb{E}}[(B^{\mathbf{a}}_t)^2|\Omega_s] &= \hat{\mathbb{E}}[(B^{\mathbf{a}}_s)^2 + (B^{\mathbf{a}}_t - B^{\mathbf{a}}_s)^2 + 2B^{\mathbf{a}}_s(B^{\mathbf{a}}_t - B^{\mathbf{a}}_s)|\Omega_s] \\ &= (B^{\mathbf{a}}_s)^2 + \hat{\mathbb{E}}[(B^{\mathbf{a}}_t - B^{\mathbf{a}}_s)^2|\Omega_s] \\ &= (B^{\mathbf{a}}_s)^2 + \sigma^2_{\mathbf{aa}^T}(t-s) \geq (B^{\mathbf{a}}_s)^2.\end{aligned}$$

Similar reasoning shows that $((B^{\mathbf{a}}_t)^2 - \sigma^2_{\mathbf{aa}^T} t)_{t\geq 0}$ and $((B^{\mathbf{a}}_t)^2 - \langle B^{\mathbf{a}}\rangle_t)_{t\geq 0}$ are G-martingales.

In general, we have the following important property.

Proposition 4.1.4 *Let $M_0 \in \mathbb{R}$, $\varphi = (\varphi^j)_{j=1}^d \in M_G^2(0,T;\mathbb{R}^d)$ and $\eta = (\eta^{ij})_{i,j=1}^d \in M_G^1(0,T;\mathbb{S}(d))$ be given and let*

$$M_t = M_0 + \int_0^t \varphi_u^j dB_u^j + \int_0^t \eta_u^{ij} d\left\langle B\right\rangle_u^{ij} - \int_0^t 2G(\eta_u)du \ \ \textit{for } t \in [0,T].$$

Then M is a G-martingale. As before, we follow the Einstein convention: the above repeated indices i and j meaning the summation.

Proof Since $\hat{\mathbb{E}}[\int_s^t \varphi_u^j dB_u^j|\Omega_s] = \hat{\mathbb{E}}[-\int_s^t \varphi_u^j dB_u^j|\Omega_s] = 0$, we only need to prove that

$$\bar{M}_t = \int_0^t \eta_u^{ij} d\left\langle B\right\rangle_u^{ij} - \int_0^t 2G(\eta_u)du \ \ \text{for } t \in [0,T]$$

is a G-martingale. It suffices to consider the case of $\eta \in M_G^{1,0}(0,T;\mathbb{S}(d))$, i.e.,

$$\eta_t = \sum_{k=0}^{N-1} \eta_{t_k} I_{[t_k,t_{k+1})}(t), \quad 0 = t_0 < t_1 < \cdots < t_n = T.$$

We have, for $s \in [t_{N-1}, t_N]$,

$$\begin{aligned}\hat{\mathbb{E}}[\bar{M}_t|\Omega_s] &= \bar{M}_s + \hat{\mathbb{E}}[(\eta_{t_{N-1}}, \langle B\rangle_t - \langle B\rangle_s) - 2G(\eta_{t_{N-1}})(t-s)|\Omega_s] \\ &= \bar{M}_s + \hat{\mathbb{E}}[(A, \langle B\rangle_t - \langle B\rangle_s)]_{A=\eta_{t_{N-1}}} - 2G(\eta_{t_{N-1}})(t-s) \\ &= \bar{M}_s.\end{aligned}$$

We can repeat this procedure backwardly thus proving the result for $s \in [0, t_{N-1}]$. □

Corollary 4.1.5 *Let $\eta \in M_G^1(0,T)$. Then for any fixed* $\mathbf{a} \in \mathbb{R}^d$*, we have*

$$\sigma^2_{-\mathbf{a}\mathbf{a}^T}\hat{\mathbb{E}}\left[\int_0^T |\eta_t|dt\right] \le \hat{\mathbb{E}}\left[\int_0^T |\eta_t|d\langle B^{\mathbf{a}}\rangle_t\right] \le \sigma^2_{\mathbf{a}\mathbf{a}^T}\hat{\mathbb{E}}\left[\int_0^T |\eta_t|dt\right]. \tag{4.1.1}$$

Proof Proposition 4.1.4 implies that, for any $\xi \in M_G^1(0,T)$,

$$\hat{\mathbb{E}}\left[\int_0^T \xi_t d\langle B^{\mathbf{a}}\rangle_t - \int_0^T 2G_{\mathbf{a}}(\xi_t)dt\right] = 0,$$

where $G_{\mathbf{a}}(\alpha) = \frac{1}{2}(\sigma^2_{\mathbf{a}\mathbf{a}^T}\alpha^+ - \sigma^2_{-\mathbf{a}\mathbf{a}^T}\alpha^-)$. Letting $\xi = |\eta|$ and $\xi = -|\eta|$, we get

$$\hat{\mathbb{E}}\left[\int_0^T |\eta_t|d\langle B^{\mathbf{a}}\rangle_t - \sigma^2_{\mathbf{a}\mathbf{a}^T}\int_0^T |\eta_t|dt\right] = 0,$$
$$\hat{\mathbb{E}}\left[-\int_0^T |\eta_t|d\langle B^{\mathbf{a}}\rangle_t + \sigma^2_{-\mathbf{a}\mathbf{a}^T}\int_0^T |\eta_t|dt\right] = 0.$$

Thus the result follows from the sub-additivity of G-expectation. □

Remark 4.1.6 If $\varphi \equiv 0$ in Proposition 4.1.4, then $M_t = \int_0^t \eta_u^{ij} d\langle B\rangle_u^{ij} - \int_0^t 2G(\eta_u)du$ is a G-martingale. This is a surprising result because M_t is a continuous and non-increasing process.

Remark 4.1.7 It is worth mentioning that for a G-martingale M, in general, $-M$ is not a G-martingale. Notice however, in Proposition 4.1.4 with $\eta \equiv 0$, the process $-M$ is still a G-martingale.

4.2 Heuristic Explanation of G-Martingale Representation

Proposition 4.1.4 tells us that a G-martingale contains a special additional term which is a decreasing martingale of the form

$$K_t = \int_0^t \eta_s d\langle B\rangle_s - \int_0^t G(\eta_s)ds.$$

In this section, we provide a formal proof to show that a G-martingale can be decomposed into a sum of a symmetric martingale and a decreasing martingale.

Let us consider a generator $G : \mathbb{S}(d) \mapsto \mathbb{R}$ satisfying the uniformly elliptic condition, i.e., there exists $\beta > 0$ such that, for each $A, \bar{A} \in \mathbb{S}(d)$ with $A \ge \bar{A}$,

$$G(A) - G(\bar{A}) \ge \beta \mathrm{tr}[A - \bar{A}].$$

For $\xi = (\xi^j)_{j=1}^d \in M_G^2(0,T;\mathbb{R}^d)$ and $\eta = (\eta^{ij})_{i,j=1}^d \in M_G^1(0,T;\mathbb{S}(d))$, we use the following notations

$$\int_0^T \langle \xi_t, dB_t\rangle := \sum_{j=1}^d \int_0^T \xi_t^j dB_t^j;\ \int_0^T (\eta_t, d\langle B\rangle_t) := \sum_{i,j=1}^d \int_0^T \eta_t^{ij} d\left\langle B\right\rangle_t^{ij}.$$

Let us first consider a G-martingale $(M_t)_{t\in[0,T]}$ with terminal condition $M_T = \xi = \varphi(B_T - B_{t_1})$ for $0 \le t_1 \le T < \infty$.

Lemma 4.2.1 *Let $\xi = \varphi(B_T - B_{t_1})$, $\varphi \in C_{b.Lip}(\mathbb{R}^d)$. Then we have the following representation:*

$$\xi = \hat{\mathbb{E}}[\xi] + \int_{t_1}^T \langle \beta_t, dB_t\rangle + \int_{t_1}^T (\eta_t, d\langle B\rangle_t) - \int_{t_1}^T 2G(\eta_t)dt.$$

Proof We know that $u(t,x) = \hat{\mathbb{E}}[\varphi(x + B_T - B_t)]$ is the solution of the following PDE:

$$\partial_t u + G(D^2u) = 0, \quad (t,x) \in [0,T]\times\mathbb{R}^d, \quad u(T,x) = \varphi(x).$$

For any $\varepsilon > 0$, by the interior regularity of u (see Appendix C), we have

$$\|u\|_{C^{1+\alpha/2,2+\alpha}([0,T-\varepsilon]\times\mathbb{R}^d)} < \infty \text{ for some } \alpha \in (0,1).$$

Applying G-Itô's formula to $u(t, B_t - B_{t_1})$ on $[t_1, T-\varepsilon]$, since $Du(t,x)$ is uniformly bounded, letting $\varepsilon \to 0$, we obtain

$$\begin{aligned}
\xi = \hat{\mathbb{E}}[\xi] &+ \int_{t_1}^T \partial_t u(t, B_t - B_{t_1})dt + \int_{t_1}^T \langle Du(t, B_t - B_{t_1}), dB_t\rangle \\
&+ \tfrac{1}{2}\int_{t_1}^T (D^2u(t, B_t - B_{t_1}), d\langle B\rangle_t) \\
= \hat{\mathbb{E}}[\xi] &+ \int_{t_1}^T \langle Du(t, B_t - B_{t_1}), dB_t\rangle + \tfrac{1}{2}\int_{t_1}^T (D^2u(t, B_t - B_{t_1}), d\langle B\rangle_t) \\
&- \int_{t_1}^T G(D^2u(t, B_t - B_{t_1}))dt.
\end{aligned}$$

□

This method can be applied to treat a more general martingale $(M_t)_{0\le t\le T}$ with terminal condition

$$\begin{aligned}
M_T =&\varphi(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_N} - B_{t_{N-1}}), \\
\varphi \in&C_{b.Lip}(\mathbb{R}^{d\times N}),\ \ 0 \le t_1 < t_2 < \cdots < t_N = T < \infty.
\end{aligned} \tag{4.2.1}$$

Indeed, it suffices to consider the case

$$\xi = \hat{\mathbb{E}}[\xi] + \int_0^T \langle \beta_t, dB_t \rangle + \int_0^T (\eta_t, d\langle B \rangle_t) - \int_0^T 2G(\eta_t)dt.$$

For $\xi = \varphi(B_{t_1}, B_T - B_{t_1})$, we set, for each $(x, y) \in \mathbb{R}^{2d}$,

$$u(t, x, y) = \hat{\mathbb{E}}[\varphi(x, y + B_T - B_t)];\ \varphi_1(x) = \hat{\mathbb{E}}[\varphi(x, B_T - B_{t_1})].$$

For $x \in \mathbb{R}^d$, we denote $\bar{\xi} = \varphi(x, B_T - B_{t_1})$. By Lemma 4.2.1, we have

$$\bar{\xi} = \varphi_1(x) + \int_{t_1}^T \langle D_y u(t, x, B_t - B_{t_1}), dB_t \rangle + \frac{1}{2} \int_{t_1}^T (D_y^2 u(t, x, B_t - B_{t_1}), d\langle B \rangle_t)$$
$$- \int_{t_1}^T G(D_y^2 u(t, x, B_t - B_{t_1}))dt.$$

Intuitively, we can replace x by B_{t_1}, apply Lemma 4.2.1 to $\varphi_1(B_{t_1})$ and conclude that

$$\xi = \varphi_1(B_{t_1}) + \int_{t_1}^T \langle D_y u(t, B_{t_1}, B_t - B_{t_1}), dB_t \rangle$$
$$+ \frac{1}{2} \int_{t_1}^T (D_y^2 u(t, B_{t_1}, B_t - B_{t_1}), d\langle B \rangle_t) - \int_{t_1}^T G(D_y^2 u(t, B_{t_1}, B_t - B_{t_1}))dt.$$

We repeat this procedure and show that the G-martingale $(M_t)_{t\in[0,T]}$ with terminal condition M_T given in (4.2.1) has the following representation:

$$M_t = \hat{\mathbb{E}}[M_T] + \int_0^t \langle \beta_s, dB_s \rangle + K_t$$

with $K_t = \int_0^t (\eta_s, d\langle B \rangle_s) - \int_0^t 2G(\eta_s)ds$ for $0 \le t \le T$.

Remark 4.2.2 Here there is a very interesting and challenging question: can we prove the above new G-martingale representation theorem for a general L_G^p-martingale? The answer of this question is provided in Theorem 7.1.1 of Chap. 7.

4.3 *G*-Convexity and Jensen's Inequality for *G*-Expectations

Here the question of interest is whether the well–known Jensen's inequality still holds for G-expectations.

First, we give a new notion of convexity.

Definition 4.3.1 A continuous function $h : \mathbb{R} \mapsto \mathbb{R}$ is called G ***–convex*** if for any bounded $\xi \in L_G^1(\Omega)$, the following Jensen's inequality holds:

$$\hat{\mathbb{E}}[h(\xi)] \geq h(\hat{\mathbb{E}}[\xi]).$$

In this section, we mainly consider C^2-functions.

Proposition 4.3.2 *Let $h \in C^2(\mathbb{R})$. Then the following statements are equivalent:*
(i) *The function h is G –convex.*
(ii) *For each bounded $\xi \in L_G^1(\Omega)$, the following Jensen's inequality holds:*

$$\hat{\mathbb{E}}[h(\xi)|\Omega_t] \geq h(\hat{\mathbb{E}}[\xi|\Omega_t]) \ \textit{for } t \geq 0.$$

(iii) *For each $\varphi \in C_b^2(\mathbb{R}^d)$, the following Jensen's inequality holds:*

$$\hat{\mathbb{E}}[h(\varphi(B_t))] \geq h(\hat{\mathbb{E}}[\varphi(B_t)]) \ \textit{for } t \geq 0.$$

(iv) *The following condition holds for each $(y, z, A) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d)$:*

$$G(h'(y)A + h''(y)zz^T) - h'(y)G(A) \geq 0. \tag{4.3.1}$$

To prove Proposition 4.3.2, we need the following lemmas.

Lemma 4.3.3 *Let $\Phi : \mathbb{R}^d \mapsto \mathbb{S}(d)$ be a continuous function with polynomial growth. Then*

$$\lim_{\delta \downarrow 0} \hat{\mathbb{E}}\left[\int_t^{t+\delta} (\Phi(B_s), d\langle B\rangle_s)\right]\delta^{-1} = 2\hat{\mathbb{E}}[G(\Phi(B_t))]. \tag{4.3.2}$$

Proof If Φ is a Lipschitz function, it is easy to show that

$$\hat{\mathbb{E}}\left[\left|\int_t^{t+\delta} (\Phi(B_s) - \Phi(B_t), d\langle B\rangle_s)\right|\right] \leq C_1 \delta^{3/2},$$

where C_1 is a constant independent of δ. Thus

$$\begin{aligned}\lim_{\delta \downarrow 0} \delta^{-1}\hat{\mathbb{E}}\left[\int_t^{t+\delta} (\Phi(B_s), d\langle B\rangle_s)\right] &= \lim_{\delta \downarrow 0} \delta^{-1}\hat{\mathbb{E}}[(\Phi(B_t), \langle B\rangle_{t+\delta} - \langle B\rangle_s)] \\ &= 2\hat{\mathbb{E}}[G(\Phi(B_t))].\end{aligned}$$

Otherwise, we can choose a sequence of Lipschitz functions $\Phi_N : \mathbb{R}^d \to \mathbb{S}(d)$ such that

$$|\Phi_N(x) - \Phi(x)| \leq \frac{C_2}{N}(1 + |x|^k),$$

where C_2 and k are positive constants independent of N. It is see to show that

$$\hat{\mathbb{E}}\left[\left|\int_t^{t+\delta}(\Phi(B_s)-\Phi_N(B_s),d\langle B\rangle_s)\right|\right]\le\frac{C}{N}\delta$$

and

$$\hat{\mathbb{E}}[|G(\Phi(B_t))-G(\Phi_N(B_t))|]\le\frac{C}{N},$$

where C is a universal constant. Thus

$$\begin{aligned}&\left|\hat{\mathbb{E}}\left[\int_t^{t+\delta}(\Phi(B_s),d\langle B\rangle_s)\right]\delta^{-1}-2\hat{\mathbb{E}}\left[G(\Phi(B_t))\right]\right|\\ \le&\left|\hat{\mathbb{E}}\left[\int_t^{t+\delta}(\Phi_N(B_s),d\langle B\rangle_s)\right]\delta^{-1}-2\hat{\mathbb{E}}\left[G(\Phi_N(B_t))\right]\right|+\frac{3C}{N}.\end{aligned}$$

Then we have

$$\limsup_{\delta\downarrow 0}\left|\hat{\mathbb{E}}\left[\int_t^{t+\delta}(\Phi(B_s),d\langle B\rangle_s)\right]\delta^{-1}-2\hat{\mathbb{E}}[G(\Phi(B_t))]\right|\le\frac{3C}{N}.$$

Since N can be arbitrarily large, this completes the proof. □

Lemma 4.3.4 *Let Ψ be a C^2-function on $\mathbb{R}^d$ with $D^2\Psi$ satisfying a polynomial growth condition. Then we have*

$$\lim_{\delta\downarrow 0}\delta^{-1}(\hat{\mathbb{E}}[\Psi(B_\delta)]-\Psi(0))=G(D^2\Psi(0)). \tag{4.3.3}$$

Proof Applying G-Itô's formula to $\Psi(B_\delta)$, we get

$$\Psi(B_\delta)=\Psi(0)+\int_0^\delta\langle D\Psi(B_s),dB_s\rangle+\tfrac{1}{2}\int_0^\delta(D^2\Psi(B_s),d\langle B\rangle_s).$$

Therefore

$$\hat{\mathbb{E}}[\Psi(B_\delta)]-\Psi(0)=\tfrac{1}{2}\hat{\mathbb{E}}\left[\int_0^\delta(D^2\Psi(B_s),d\langle B\rangle_s)\right].$$

By Lemma 4.3.3, we obtain the result. □

Lemma 4.3.5 *Let $h\in C^2(\mathbb{R})$ and satisfy (4.3.1). For any $\varphi\in C_{b.Lip}(\mathbb{R}^d)$, let $u(t,x)$ be the solution of the G-heat equation:*

$$\partial_t u-G(D^2u)=0\quad(t,x)\in[0,\infty)\times\mathbb{R}^d,\ u(0,x)=\varphi(x). \tag{4.3.4}$$

Then $\tilde{u}(t,x):=h(u(t,x))$ is a viscosity subsolution of the G-heat Eq. (4.3.4) with initial condition $\tilde{u}(0,x)=h(\varphi(x))$.

Proof For each $\varepsilon > 0$, we denote by u_ε the solution of the following PDE:

$$\partial_t u_\varepsilon - G_\varepsilon(D^2 u_\varepsilon) = 0 \quad (t,x) \in [0,\infty) \times \mathbb{R}^d,\ u_\varepsilon(0,x) = \varphi(x),$$

where $G_\varepsilon(A) := G(A) + \varepsilon \mathrm{tr}[A]$. Since G_ε satisfies the uniformly elliptic condition, by Appendix C, we have $u_\varepsilon \in C^{1,2}((0,\infty) \times \mathbb{R}^d)$. By simple calculation, we have

$$\partial_t h(u_\varepsilon) = h'(u_\varepsilon)\partial_t u_\varepsilon = h'(u_\varepsilon) G_\varepsilon(D^2 u_\varepsilon)$$

and

$$\partial_t h(u_\varepsilon) - G_\varepsilon(D^2 h(u_\varepsilon)) = f_\varepsilon(t,x),\ h(u_\varepsilon(0,x)) = h(\varphi(x)),$$

where

$$f_\varepsilon(t,x) = h'(u_\varepsilon)G(D^2 u_\varepsilon) - G(D^2 h(u_\varepsilon)) - \varepsilon h''(u_\varepsilon)|Du_\varepsilon|^2.$$

Since h satisfies (4.3.1), it follows that $f_\varepsilon \le -\varepsilon h''(u_\varepsilon)|Du_\varepsilon|^2$. We can also deduce that $|Du_\varepsilon|$ is uniformly bounded by the Lipschitz constant of φ. It is easy to show that u_ε uniformly converges to u as $\varepsilon \to 0$. Thus $h(u_\varepsilon)$ uniformly converges to $h(u)$ and $h''(u_\varepsilon)$ is uniformly bounded. Then we get

$$\partial_t h(u_\varepsilon) - G_\varepsilon(D^2 h(u_\varepsilon)) \le C\varepsilon,\ h(u_\varepsilon(0,x)) = h(\varphi(x)),$$

where C is a constant independent of ε. By Appendix C, we conclude that $h(u)$ is a viscosity subsolution. □

Proof of Proposition 4.3.2 Obviously **(ii)** $\Longrightarrow$**(i)**$\Longrightarrow$**(iii)** . We now show **(iii)**$\Longrightarrow$**(ii)**. For $\xi \in L_G^1(\Omega)$ of the form

$$\xi = \varphi(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}),$$

where $\varphi \in C_b^2(\mathbb{R}^{d\times n})$, $0 \le t_1 \le \cdots \le t_n < \infty$, by the definitions of $\hat{\mathbb{E}}[\cdot]$ and $\hat{\mathbb{E}}[\cdot|\Omega_t]$, we have

$$\hat{\mathbb{E}}[h(\xi)|\Omega_t] \ge h(\hat{\mathbb{E}}[\xi|\Omega_t]),\ t \ge 0.$$

This Jensen's inequality can be extended to hold under the norm $||\cdot|| = \hat{\mathbb{E}}[|\cdot|]$, to each $\xi \in L_G^1(\Omega)$ satisfying $h(\xi) \in L_G^1(\Omega)$.

Let us show **(iii)**$\Longrightarrow$**(iv)**: for each $\varphi \in C_b^2(\mathbb{R}^d)$, we have $\hat{\mathbb{E}}[h(\varphi(B_t))] \ge h(\hat{\mathbb{E}}[\varphi(B_t)])$ for $t \ge 0$. By Lemma 4.3.4, we know that

$$\lim_{\delta\downarrow 0}(\hat{\mathbb{E}}[\varphi(B_\delta)] - \varphi(0))\delta^{-1} = G(D^2\varphi(0))$$

and

$$\lim_{\delta\downarrow 0}(\hat{\mathbb{E}}[h(\varphi(B_\delta))] - h(\varphi(0)))\delta^{-1} = G(D^2 h(\varphi)(0)).$$

Thus we obtain

$$G(D^2 h(\varphi)(0)) \geq h'(\varphi(0))G(D^2\varphi(0)).$$

For each $(y, z, A) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d)$, we can choose $\varphi \in C_b^2(\mathbb{R}^d)$ such that

$$(\varphi(0), D\varphi(0), D^2\varphi(0)) = (y, z, A).$$

Thus we obtain **(iv)**.

Finally, **(iv)**$\Longrightarrow$**(iii)**: for each $\varphi \in C_b^2(\mathbb{R}^d)$, $u(t, x) = \hat{\mathbb{E}}[\varphi(x + B_t)]$ (respectively, $\bar{u}(t, x) = \hat{\mathbb{E}}[h(\varphi(x + B_t))]$) solves the G-heat Eq. (4.3.4). By Lemma 4.3.5, $h(u)$ is a viscosity subsolution of the G-heat Eq. (4.3.4). It follows from the maximum principle that $h(u(t, x)) \leq \bar{u}(t, x)$. In particular, **(iii)** holds. □

Remark 4.3.6 In fact, **(i)**$\Longleftrightarrow$**(ii)** $\Longleftrightarrow$**(iii)** still hold without assuming that $h \in C^2(\mathbb{R})$.

Proposition 4.3.7 *Let h be a G-convex function and $X \in L_G^1(\Omega)$ be bounded. Then the process $Y_t = h(\hat{\mathbb{E}}[X|\Omega_t])$, $t \geq 0$, is a G-submartingale.*

Proof For each $s \leq t$,

$$\hat{\mathbb{E}}[Y_t|\Omega_s] = \hat{\mathbb{E}}[h(\hat{\mathbb{E}}[X|\Omega_t])|\Omega_s] \geq h(\hat{\mathbb{E}}[X|\Omega_s]) = Y_s.$$ □

4.4 Exercises

Exercise 4.4.1 (a) Let $(M_t)_{t\geq 0}$ be a G-supermartingale. Show that the process $(-M_t)_{t\geq 0}$ is a G-submartingale.

(b) Find a G-submartingale $(M_t)_{t\geq 0}$ such that $(-M_t)_{t\geq 0}$ is not a G-supermartingale.

Exercise 4.4.2 (a) Assume that $(M_t)_{t\geq 0}$ and $(N_t)_{t\geq 0}$ be two G -supermartingales. Prove that their sum $(M_t + N_t)_{t\geq 0}$ is a G -supermartingale.

(b) Assume that $(M_t)_{t\geq 0}$ and $(-M_t)_{t\geq 0}$ are two G-martingales. For each G-submartingale $(N_t)_{t\geq 0}$, prove that $(M_t + N_t)_{t\geq 0}$ is a G-submartingale.

Exercise 4.4.3 Suppose that G satisfies the uniformly elliptic condition and $h \in C^2(\mathbb{R})$. Show that h is G-convex if and only if h is convex.

Notes and Comments

The material in this chapter is mainly from Peng [140].

Peng [130] introduced a filtration consistent (or time consistent, or dynamic) nonlinear expectation, called g-expectation, via BSDE, developed further in (1999)

[132] for some basic properties of the g-martingale such as nonlinear Doob-Meyer decomposition theorem. See also Briand et al. [20] , Chen et al. [29], Chen and Peng [30, 31], Coquet, Hu, Mémin and Peng [35, 36], Peng [132, 135], Peng and Xu [148], Rosazza [152]. These works lead to a conjecture that all properties obtained for g-martingales must have their counterparts for G-martingale. However this conjecture is still far from being complete.

The problem of G-martingale representation has been proposed by Peng [140]. In Sect. 4.2, we only state a result with very regular random variables. Some very interesting developments to this important problem will be provided in Chap. 7.

Under the framework of g-expectation, Chen, Kulperger and Jiang [29], Hu [86], Jiang and Chen [97] investigate the Jensen's inequality for g-expectation. Jia and Peng [95] introduced the notion of g-convex function and obtained many interesting properties. Certainly, a G-convex function concerns fully nonlinear situations.

Chapter 5
Stochastic Differential Equations

In this chapter, we consider the stochastic differential equations and backward stochastic differential equations driven by G-Brownian motion. The conditions and proofs of existence and uniqueness of a stochastic differential equation is similar to the classical situation. However the corresponding problems for backward stochastic differential equations are not that easy, many are still open. We only give partial results to this direction.

5.1 Stochastic Differential Equations

In this chapter, we denote by $\bar{M}_G^p(0, T; \mathbb{R}^n)$, $p \geq 1$, the completion of $M_G^{p,0}(0, T; \mathbb{R}^n)$ under the norm $(\int_0^T \hat{\mathbb{E}}[|\eta_t|^p]dt)^{1/p}$. It is not hard to prove that $\bar{M}_G^p(0, T; \mathbb{R}^n) \subseteq M_G^p(0, T; \mathbb{R}^n)$. We consider all the problems in the space $\bar{M}_G^p(0, T; \mathbb{R}^n)$. The following lemma is useful in our future discussion.

Lemma 5.1.1 *Suppose that* $\varphi \in M_G^2(0, T)$. *Then for* $\mathbf{a} \in \mathbb{R}^d$, *it holds that*

$$\eta_t := \int_0^t \varphi_s dB_s^{\mathbf{a}} \in \bar{M}_G^2(0, T).$$

Proof Choosing a sequence of processes $\varphi^n \in M_G^{2,0}(0, T)$ such that

$$\lim_{n\to\infty} \hat{\mathbb{E}}\left[\int_0^T |\varphi_s - \varphi_s^n|^2 ds\right] = 0.$$

Then for each integer n, it is easy to check that the process $\eta_t^n = \int_0^t \varphi_s^n dB_s^{\mathbf{a}}$ belongs to the space $\bar{M}_G^2(0, T)$.

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_5

On the other hand, it follows from the property of G-Itô integral that

$$\int_0^T \hat{\mathbb{E}}[|\eta_t - \eta_t^n|^2]dt = \sigma^2_{\mathbf{aa}^T} \int_0^T \hat{\mathbb{E}}\left[\int_0^t |\varphi_s - \varphi_s^n|^2 ds\right] dt \leq \sigma^2_{\mathbf{aa}^T} T \hat{\mathbb{E}}\left[\int_0^T |\varphi_s - \varphi_s^n|^2 ds\right],$$

which implies the desired result. □

Now we consider the following SDE driven by a d-dimensional G-Brownian motion:

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t h_{ij}(s, X_s)d\langle B\rangle_s^{ij} + \int_0^t \sigma_j(s, X_s)dB_s^j, \ t \in [0, T], \tag{5.1.1}$$

where the initial condition $X_0 \in \mathbb{R}^n$ is a given constant, b, h_{ij}, σ_j are given functions satisfying $b(\cdot, x)$, $h_{ij}(\cdot, x)$, $\sigma_j(\cdot, x) \in M_G^2(0, T; \mathbb{R}^n)$ for each $x \in \mathbb{R}^n$ and the Lipschitz condition, i.e., $|\phi(t, x) - \phi(t, x')| \leq K|x - x'|$, for each $t \in [0, T]$, $x, x' \in \mathbb{R}^n$, $\phi = b$, h_{ij} and σ_j, respectively. Here the horizon $[0, T]$ can be arbitrarily large. The solution is a process $(X_t)_{t\in[0,T]} \in \bar{M}_G^2(0, T; \mathbb{R}^n)$ satisfying the SDE (5.1.1).

We first introduce the following mapping on a fixed interval $[0, T]$:

$$\Lambda_\cdot : \bar{M}_G^2(0, T; \mathbb{R}^n) \mapsto \bar{M}_G^2(0, T; \mathbb{R}^n)$$

by setting Λ_t, $t \in [0, T]$, with

$$\Lambda_t(Y) = X_0 + \int_0^t b(s, Y_s)ds + \int_0^t h_{ij}(s, Y_s)d\langle B\rangle_s^{ij} + \int_0^t \sigma_j(s, Y_s)dB_s^j.$$

From Lemma 5.1.1 and Exercise 5.4.2 of this chapter, we see that the mapping Λ is well-defined.

We immediately have the following lemma, whose proof is left to the reader.

Lemma 5.1.2 *For any $Y, Y' \in \bar{M}_G^2(0, T; \mathbb{R}^n)$, we have the following estimate:*

$$\hat{\mathbb{E}}[|\Lambda_t(Y) - \Lambda_t(Y')|^2] \leq C \int_0^t \hat{\mathbb{E}}[|Y_s - Y_s'|^2]ds, \ t \in [0, T], \tag{5.1.2}$$

where the constant C depends only on the Lipschitz constant K.

We now prove that the SDE (5.1.1) has a unique solution. We multiply on both sides of (5.1.2) by e^{-2Ct} and integrate them on $[0, T]$, thus deriving

$$\begin{aligned}\int_0^T \hat{\mathbb{E}}[|\Lambda_t(Y) - \Lambda_t(Y')|^2]e^{-2Ct}dt &\leq C \int_0^T e^{-2Ct} \int_0^t \hat{\mathbb{E}}[|Y_s - Y_s'|^2]dsdt \\ &= C \int_0^T \int_s^T e^{-2Ct}dt\hat{\mathbb{E}}[|Y_s - Y_s'|^2]ds \\ &= \frac{1}{2} \int_0^T (e^{-2Cs} - e^{-2CT})\hat{\mathbb{E}}[|Y_s - Y_s'|^2]ds.\end{aligned}$$

We then have

$$\int_0^T \hat{\mathbb{E}}[|\Lambda_t(Y) - \Lambda_t(Y')|^2]e^{-2Ct}dt \le \frac{1}{2}\int_0^T \hat{\mathbb{E}}[|Y_t - Y_t'|^2]e^{-2Ct}dt. \quad (5.1.3)$$

Note that the following two norms are equivalent in the space $\bar{M}_G^2(0, T; \mathbb{R}^n)$:

$$\left(\int_0^T \hat{\mathbb{E}}[|Y_t|^2]dt\right)^{1/2} \sim \left(\int_0^T \hat{\mathbb{E}}[|Y_t|^2]e^{-2Ct}dt\right)^{1/2}.$$

From (5.1.3) we obtain that $\Lambda(Y)$ is a contraction mapping. Consequently, we have the following theorem.

Theorem 5.1.3 *There exists a unique solution* $(X_t)_{0\le t\le T} \in \bar{M}_G^2(0, T; \mathbb{R}^n)$ *of the stochastic differential equation* (5.1.1).

We now consider a particular but important case of a linear SDE. For simplicity, assume that $d = 1$, $n = 1$. and let

$$X_t = X_0 + \int_0^t (b_s X_s + \tilde{b}_s)ds + \int_0^t (h_s X_s + \tilde{h}_s)d\langle B\rangle_s + \int_0^t (\sigma_s X_s + \tilde{\sigma}_s)dB_s,\ t \in [0, T]. \quad (5.1.4)$$

Here $X_0 \in \mathbb{R}$ is given, $b_., h_., \sigma_.$ are given bounded processes in $M_G^2(0, T; \mathbb{R})$ and $\tilde{b}_., \tilde{h}_., \tilde{\sigma}_.$ are given processes in $M_G^2(0, T; \mathbb{R})$. It follows from Theorem 5.1.3 that the linear SDE (5.1.4) has a unique solution.

Remark 5.1.4 The solution of the linear SDE (5.1.4) is

$$X_t = \Gamma_t^{-1}(X_0 + \int_0^t \tilde{b}_s\Gamma_s ds + \int_0^t (\tilde{h}_s - \sigma_s\tilde{\sigma}_s)\Gamma_s d\langle B\rangle_s + \int_0^t \tilde{\sigma}_s\Gamma_s dB_s),\ t \in [0, T],$$

where $\Gamma_t = \exp(-\int_0^t b_s ds - \int_0^t (h_s - \frac{1}{2}\sigma_s^2)d\langle B\rangle_s - \int_0^t \sigma_s dB_s)$.

In particular, if $b_., h_., \sigma_.$ are constants and $\tilde{b}_., \tilde{h}_., \tilde{\sigma}_.$ are zero, then X is a geometric G-Brownian motion.

Definition 5.1.5 We say that $(X_t)_{t\ge0}$ is a **geometric G-Brownian motion** if

$$X_t = \exp(\alpha t + \beta\langle B\rangle_t + \gamma B_t), \quad (5.1.5)$$

where α, β, γ are constants.

5.2 Backward Stochastic Differential Equations (BSDE)

We consider the following type of BSDE:

$$Y_t = \hat{\mathbb{E}}\left[\xi + \int_t^T f(s, Y_s)ds + \int_t^T h_{ij}(s, Y_s)d\left\langle B\right\rangle_s^{ij} \Big| \Omega_t\right], \quad t \in [0, T], \tag{5.2.1}$$

where $\xi \in L_G^1(\Omega_T; \mathbb{R}^n)$, f, h_{ij} are given functions such that $f(\cdot, y)$, $h_{ij}(\cdot, y) \in M_G^1(0, T; \mathbb{R}^n)$ for each $y \in \mathbb{R}^n$ and these functions satisfy the Lipschitz condition, i.e.,

$$|\phi(t, y) - \phi(t, y')| \leq K|y - y'|, \quad \text{for each } \ t \in [0, T], \quad y, y' \in \mathbb{R}^n, \quad \phi = f \ \text{ and } h_{ij}.$$

The solution is a process $(Y_t)_{0 \leq t \leq T} \in \bar{M}_G^1(0, T; \mathbb{R}^n)$ satisfying the above BSDE.

We first introduce the following mapping on a fixed interval $[0, T]$:

$$\Lambda_\cdot : \bar{M}_G^1(0, T; \mathbb{R}^n) \to \bar{M}_G^1(0, T; \mathbb{R}^n)$$

by setting Λ_t, $t \in [0, T]$ as follows:

$$\Lambda_t(Y) = \hat{\mathbb{E}}\left[\xi + \int_t^T f(s, Y_s)ds + \int_t^T h_{ij}(s, Y_s)d\left\langle B\right\rangle_s^{ij} \Big| \Omega_t\right],$$

which is well-defined by Lemma 5.1.1 and Exercises 5.4.2, 5.4.5.

We immediately derive a useful property of Λ_t.

Lemma 5.2.1 *For any $Y, Y' \in \bar{M}_G^1(0, T; \mathbb{R}^n)$, we have the following estimate:*

$$\hat{\mathbb{E}}[|\Lambda_t(Y) - \Lambda_t(Y')|] \leq C \int_t^T \hat{\mathbb{E}}[|Y_s - Y_s'|]ds, \ t \in [0, T], \tag{5.2.2}$$

where the constant C depends only on the Lipschitz constant K.

Now we are going to prove that the BSDE (5.2.1) has a unique solution. We multiplying on both sides of (5.2.2) by e^{2Ct}, and integrate them on $[0, T]$. We find

$$\begin{aligned}\int_0^T \hat{\mathbb{E}}[|\Lambda_t(Y) - \Lambda_t(Y')|]e^{2Ct}dt &\leq C\int_0^T \int_t^T \hat{\mathbb{E}}[|Y_s - Y_s'|]e^{2Ct}dsdt \\ &= C\int_0^T \hat{\mathbb{E}}[|Y_s - Y_s'|] \int_0^s e^{2Ct}dtds \\ &= \frac{1}{2}\int_0^T \hat{\mathbb{E}}[|Y_s - Y_s'|](e^{2Cs} - 1)ds \\ &\leq \frac{1}{2}\int_0^T \hat{\mathbb{E}}[|Y_s - Y_s'|]e^{2Cs}ds. \end{aligned} \tag{5.2.3}$$

We observe that the following two norms in the space $\bar{M}_G^1(0, T; \mathbb{R}^n)$ are equivalent:

$$\int_0^T \hat{\mathbb{E}}[|Y_t|]dt \sim \int_0^T \hat{\mathbb{E}}[|Y_t|]e^{2Ct}dt.$$

From (5.2.3), we can obtain that $\Lambda(Y)$ is a contraction mapping. Consequently, we have proved the following theorem.

Theorem 5.2.2 *There exists a unique solution $(Y_t)_{t\in[0,T]} \in \bar{M}_G^1(0, T; \mathbb{R}^n)$ of the backward stochastic differential equation* (5.2.1).

Let $Y^{(v)}$, $v = 1, 2$, be the solutions of the following BSDE:

$$Y_t^{(v)} = \hat{\mathbb{E}}\left[\xi^{(v)} + \int_t^T (f(s, Y_s^{(v)}) + \varphi_s^{(v)})ds + \int_t^T (h_{ij}(s, Y_s^{(v)}) + \psi_s^{ij,(v)})d\langle B\rangle_s^{ij} \Big|\Omega_t\right].$$

Then the following estimate holds.

Proposition 5.2.3 *We have*

$$\hat{\mathbb{E}}\left[|Y_t^{(1)} - Y_t^{(2)}|\right] \leq Ce^{C(T-t)}\hat{\mathbb{E}}[|\xi^{(1)} - \xi^{(2)}| + \int_t^T |\varphi_s^{(1)} - \varphi_s^{(2)}| + |\psi_s^{ij,(1)} - \psi_s^{ij,(2)}|ds], \tag{5.2.4}$$

where the constant C depends only on the Lipschitz constant K.

Proof As in the proof of Lemma 5.2.1, we have

$$\hat{\mathbb{E}}[|Y_t^{(1)} - Y_t^{(2)}|] \leq C\left(\int_t^T \hat{\mathbb{E}}[|Y_s^{(1)} - Y_s^{(2)}|]ds + \hat{\mathbb{E}}[|\xi^{(1)} - \xi^{(2)}| \right. \\ \left. + \int_t^T |\varphi_s^{(1)} - \varphi_s^{(2)}| + |\psi_s^{ij,(1)} - \psi_s^{ij,(2)}|ds]\right).$$

By applying the Gronwall inequality (see Exercise 5.4.4), we obtain the statement.

Remark 5.2.4 In particular, if $\xi^{(2)} = 0, \varphi_s^{(2)} = -f(s, 0), \psi_s^{ij,(2)} = -h_{ij}(s, 0), \xi^{(1)} = \xi$, $\varphi_s^{(1)} = 0$, $\psi_s^{ij,(1)} = 0$, we obtain the estimate of the solution of the BSDE. Let Y be the solution of the BSDE (5.2.1). Then

$$\hat{\mathbb{E}}[|Y_t|] \leq Ce^{C(T-t)}\hat{\mathbb{E}}\left[|\xi| + \int_t^T |f(s, 0)| + |h_{ij}(s, 0)|ds\right], \tag{5.2.5}$$

where the constant C depends only on the Lipschitz constant K.

5.3 Nonlinear Feynman-Kac Formula

Consider the following SDE:

$$\begin{cases} dX_s^{t,\xi} = b(X_s^{t,\xi})ds + h_{ij}(X_s^{t,\xi})d\langle B\rangle_s^{ij} + \sigma_j(X_s^{t,\xi})dB_s^j, \ s \in [t,T], \\ X_t^{t,\xi} = \xi, \end{cases} \tag{5.3.1}$$

where $\xi \in L_G^2(\Omega_t;\mathbb{R}^n)$ and $b, h_{ij}, \sigma_j : \mathbb{R}^n \mapsto \mathbb{R}^n$ are given Lipschitz functions, i.e., $|\phi(x)-\phi(x')| \leq K|x-x'|$, for all $x, x' \in \mathbb{R}^n$, $\phi = b$, h_{ij} and σ_j.

We then consider the associated BSDE:

$$Y_s^{t,\xi} = \hat{\mathbb{E}}\left[\Phi(X_T^{t,\xi}) + \int_s^T f(X_r^{t,\xi}, Y_r^{t,\xi})dr + \int_s^T g_{ij}(X_r^{t,\xi}, Y_r^{t,\xi})d\left\langle B^i, B^j\right\rangle_r \Big|\Omega_s\right], \tag{5.3.2}$$

where $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a given Lipschitz function and $f, g_{ij} : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}$ are given Lipschitz functions, i.e., $|\phi(x,y)-\phi(x',y')| \leq K(|x-x'|+|y-y'|)$, for each x, $x' \in \mathbb{R}^n$, $y, y' \in \mathbb{R}$, $\phi = f$ and g_{ij}.

We have the following estimates:

Proposition 5.3.1 *For each $\xi, \xi' \in L_G^2(\Omega_t;\mathbb{R}^n)$, we have, for each $s \in [t,T]$,*

$$\hat{\mathbb{E}}[|X_s^{t,\xi} - X_s^{t,\xi'}|^2|\Omega_t] \leq C|\xi-\xi'|^2 \tag{5.3.3}$$

and

$$\hat{\mathbb{E}}[|X_s^{t,\xi}|^2|\Omega_t] \leq C(1+|\xi|^2), \tag{5.3.4}$$

where the constant C depends only on the Lipschitz constant K.

Proof It is easy to see that

$$\hat{\mathbb{E}}[|X_s^{t,\xi} - X_s^{t,\xi'}|^2\big|\Omega_t] \leq C_1(|\xi-\xi'|^2 + \int_t^s \hat{\mathbb{E}}[|X_r^{t,\xi} - X_r^{t,\xi'}|^2|\Omega_t]dr).$$

By the Gronwall inequality, we obtain (5.3.3), namely

$$\hat{\mathbb{E}}[|X_s^{t,\xi} - X_s^{t,\xi'}|^2|\Omega_t] \leq C_1 e^{C_1 T}|\xi-\xi'|^2.$$

Similarly, we derive (5.3.4). □

Corollary 5.3.2 *For any $\xi \in L_G^2(\Omega_t;\mathbb{R}^n)$, we have*

$$\hat{\mathbb{E}}[|X_{t+\delta}^{t,\xi} - \xi|^2|\Omega_t] \leq C(1+|\xi|^2)\delta \ \text{ for } \delta \in [0, T-t], \tag{5.3.5}$$

where the constant C depends only on the Lipschitz constant K.

Proof It is easy to see that

$$\hat{\mathbb{E}}[|X^{t,\xi}_{t+\delta} - \xi|^2 \big| \Omega_t] \le C_1 \int_t^{t+\delta} \left(1 + \hat{\mathbb{E}}[|X^{t,\xi}_s|^2 \big| \Omega_t]\right) ds.$$

Then the result follows from Proposition 5.3.1. □

Proposition 5.3.3 *For each $\xi, \xi' \in L^2_G(\Omega_t; \mathbb{R}^n)$, we have*

$$|Y^{t,\xi}_t - Y^{t,\xi'}_t| \le C|\xi - \xi'| \tag{5.3.6}$$

and

$$|Y^{t,\xi}_t| \le C(1 + |\xi|), \tag{5.3.7}$$

where the constant C depends only on the Lipschitz constant K.

Proof For each $s \in [0, T]$, it is easy to check that

$$|Y^{t,\xi}_s - Y^{t,\xi'}_s| \le C_1 \hat{\mathbb{E}}\left[|X^{t,\xi}_T - X^{t,\xi'}_T| + \int_s^T (|X^{t,\xi}_r - X^{t,\xi'}_r| + |Y^{t,\xi}_r - Y^{t,\xi'}_r|) dr | \Omega_s\right].$$

Since

$$\hat{\mathbb{E}}[|X^{t,\xi}_s - X^{t,\xi'}_s||\Omega_t] \le \left(\hat{\mathbb{E}}[|X^{t,\xi}_s - X^{t,\xi'}_s|^2|\Omega_t]\right)^{1/2},$$

we have

$$\hat{\mathbb{E}}[|Y^{t,\xi}_s - Y^{t,\xi'}_s||\Omega_t] \le C_2(|\xi - \xi'| + \int_s^T \hat{\mathbb{E}}[|Y^{t,\xi}_r - Y^{t,\xi'}_r||\Omega_t] dr).$$

By the Gronwall inequality, we obtain (5.3.6). Similarly we derive (5.3.7). □

We are more interested in the case when $\xi = x \in \mathbb{R}^n$. Define

$$u(t, x) := Y^{t,x}_t, \quad (t, x) \in [0, T] \times \mathbb{R}^n. \tag{5.3.8}$$

By Proposition 5.3.3, we immediately have the following estimates:

$$|u(t, x) - u(t, x')| \le C|x - x'|, \tag{5.3.9}$$

$$|u(t, x)| \le C(1 + |x|), \tag{5.3.10}$$

where the constant C depends only on the Lipschitz constant K.

Remark 5.3.4 It is important to note that $u(t, x)$ is a deterministic function of (t, x), because $X^{t,x}_s$ and $Y^{t,x}_s$ are independent from Ω_t.

Theorem 5.3.5 *For any $\xi \in L_G^2(\Omega_t;\mathbb{R}^n)$, we have*

$$u(t,\xi) = Y_t^{t,\xi}. \tag{5.3.11}$$

Proof Without loss of generality, suppose that $n = 1$.

First, we assume that $\xi \in Lip(\Omega_T)$ is bounded by some constant ρ. Thus for each integer $N > 0$, we can choose a simple function

$$\eta^N = \sum_{i=-N}^{N} x_i \mathbf{I}_{A_i}(\xi)$$

with $x_i = \frac{i\rho}{N}$, $A_i = [\frac{i\rho}{N}, \frac{(i+1)\rho}{N})$ for $i = -N, \ldots, N-1$ and $x_N = \rho$, $A_N = \{\rho\}$. From the definition of u, we conclude that

$$|Y_t^{t,\xi} - u(t,\eta^N)| = |Y_t^{t,\xi} - \sum_{i=-N}^{N} u(t,x_i)\mathbf{I}_{A_i}(\xi)| = |Y_t^{t,\xi} - \sum_{i=-N}^{N} Y_t^{t,x_i}\mathbf{I}_{A_i}(\xi)|$$
$$= \sum_{i=-N}^{N} |Y_t^{t,\xi} - Y_t^{t,x_i}|\mathbf{I}_{A_i}(\xi).$$

Then it follows from Proposition 5.3.3 that

$$|Y_t^{t,\xi} - u(t,\eta^N)| \le C \sum_{i=-N}^{N} |\xi - x_i|\mathbf{I}_{A_i}(\xi) \le C\frac{\rho}{N}.$$

Noting that

$$|u(t,\xi) - u(t,\eta^N)| \le C|\xi - \eta^N| \le C\frac{\rho}{N},$$

we get $\hat{\mathbb{E}}[|Y_t^{t,\xi} - u(t,\xi)|] \le 2C\frac{\rho}{N}$. Since N can be arbitrarily large, we obtain $Y_t^{t,\xi} = u(t,\xi)$.

In the general case, by Exercise 3.10.4 in Chap. 3, we can find a sequence of bounded random variables $\xi_k \in Lip(\Omega_T)$ such that

$$\lim_{k\to\infty} \hat{\mathbb{E}}[|\xi - \xi_k|^2] = 0.$$

Consequently, applying Proposition 5.3.3 again yields that

$$\lim_{k\to\infty} \hat{\mathbb{E}}[|Y_t^{t,\xi} - Y_t^{t,\xi_k}|^2] \le C \lim_{k\to\infty} \hat{\mathbb{E}}[|\xi - \xi_k|^2] = 0,$$

which together with $Y_t^{t,\xi_k} = u(t,\xi_k)$ imply the desired result. □

Proposition 5.3.6 *We have, for* $\delta \in [0, T-t]$,

$$u(t,x) = \hat{\mathbb{E}}\left[u(t+\delta, X_{t+\delta}^{t,x}) + \int_t^{t+\delta} f(X_r^{t,x}, Y_r^{t,x})dr + \int_t^{t+\delta} g_{ij}(X_r^{t,x}, Y_r^{t,x})d\langle B\rangle_r^{ij}\right]. \tag{5.3.12}$$

Proof Since $X_s^{t,x} = X_s^{t+\delta, X_{t+\delta}^{t,x}}$ for $s \in [t+\delta, T]$, we get $Y_{t+\delta}^{t,x} = Y_{t+\delta}^{t+\delta, X_{t+\delta}^{t,x}}$. By Theorem 5.3.5, we have $Y_{t+\delta}^{t,x} = u(t+\delta, X_{t+\delta}^{t,x})$, which implies the result. □

For any $A \in \mathbb{S}(n)$, $p \in \mathbb{R}^n$, $r \in \mathbb{R}$, we set

$$F(A, p, r, x) := G(B(A, p, r, x)) + \langle p, b(x)\rangle + f(x, r),$$

where $B(A, p, r, x)$ is a $d \times d$ symmetric matrix with

$$B_{ij}(A, p, r, x) := \langle A\sigma_i(x), \sigma_j(x)\rangle + \langle p, h_{ij}(x) + h_{ji}(x)\rangle + g_{ij}(x, r) + g_{ji}(x, r).$$

Theorem 5.3.7 *The function* $u(t, x)$ *is the unique viscosity solution of the following PDE:*

$$\begin{cases} \partial_t u + F(D^2u, Du, u, x) = 0, \\ u(T, x) = \Phi(x). \end{cases} \tag{5.3.13}$$

Proof We first show that u is a continuous function. By (5.3.9) we know that u is a Lipschitz function in x. It follows from (5.2.5) and (5.3.4) that

$$\hat{\mathbb{E}}[|Y_s^{t,x}|] \leq C(1+|x|), \quad \text{for } s \in [t, T].$$

In view of (5.3.5) and (5.3.12), we get $|u(t,x) - u(t+\delta, x)| \leq C(1+|x|)(\delta^{1/2}+\delta)$ for $\delta \in [0, T-t]$. Thus u is $\frac{1}{2}$-Hölder continuous in t, which implies that u is a continuous function. We can also show (see Exercise 5.4.8), that for each $p \geq 2$,

$$\hat{\mathbb{E}}[|X_{t+\delta}^{t,x} - x|^p] \leq C(1+|x|^p)\delta^{p/2}. \tag{5.3.14}$$

Now for fixed $(t, x) \in (0, T) \times \mathbb{R}^n$, let $\psi \in C_{l,Lip}^{2,3}([0, T] \times \mathbb{R}^n)$ be such that $\psi \geq u$ and $\psi(t, x) = u(t, x)$. By (5.3.12), (5.3.14) and Taylor's expansion, it follows that, for $\delta \in (0, T-t)$,

$$\begin{aligned} 0 &\leq \hat{\mathbb{E}}\Bigg[\psi(t+\delta, X_{t+\delta}^{t,x}) - \psi(t,x) + \int_t^{t+\delta} f(X_r^{t,x}, Y_r^{t,x})dr \\ &\quad + \int_t^{t+\delta} g_{ij}(X_r^{t,x}, Y_r^{t,x})d\left\langle B^i, B^j\right\rangle_r\Bigg] \\ &\leq \frac{1}{2}\hat{\mathbb{E}}[(B(D^2\psi(t,x), D\psi(t,x), \psi(t,x), x), \langle B\rangle_{t+\delta} - \langle B\rangle_t)] \\ &\quad + (\partial_t\psi(t,x) + \langle D\psi(t,x), b(x)\rangle + f(x, \psi(t,x)))\delta + C(1+|x|^m)\delta^{3/2} \end{aligned}$$

$$\leq (\partial_t \psi(t,x) + F(D^2\psi(t,x), D\psi(t,x), \psi(t,x), x))\delta + C(1+|x|^m)\delta^{3/2},$$

where m is some constant depending on the function ψ. Consequently, it is easy to check that

$$\partial_t \psi(t,x) + F(D^2\psi(t,x), D\psi(t,x), \psi(t,x), x) \geq 0.$$

This implies that u is a viscosity subsolution of (5.3.13). Similarly we can show that u is also a viscosity supersolution of (5.3.13). The uniqueness is from Theorem C.2.9 (in Appendix C). □

Example 5.3.8 Let $B = (B^1, B^2)$ be a 2-dimensional G-Brownian motion with

$$G(A) = G_1(a_{11}) + G_2(a_{22}),$$

where

$$G_i(a) = \frac{1}{2}(\overline{\sigma}_i^2 a^+ - \underline{\sigma}_i^2 a^-), \quad i = 1, 2.$$

In this case, we consider the following 1-dimensional SDE:

$$dX_s^{t,x} = \mu X_s^{t,x} ds + \nu X_s^{t,x} d\left\langle B^1 \right\rangle_s + \sigma X_s^{t,x} dB_s^2, \quad X_t^{t,x} = x,$$

where μ, ν and σ are constants.

The corresponding function u is defined by

$$u(t,x) := \hat{\mathbb{E}}[\varphi(X_T^{t,x})].$$

Then

$$u(t,x) = \hat{\mathbb{E}}[u(t+\delta, X_{t+\delta}^{t,x})]$$

and u is the viscosity solution of the following PDE:

$$\partial_t u + \mu x \partial_x u + 2G_1(\nu x \partial_x u) + \sigma^2 x^2 G_2(\partial_{xx}^2 u) = 0, \ u(T,x) = \varphi(x).$$

5.4 Exercises

Exercise 5.4.1 Prove that $\bar{M}_G^p(0,T;\mathbb{R}^n) \subseteq M_G^p(0,T;\mathbb{R}^n)$.

Exercise 5.4.2 Show that $b(s, Y_s) \in M_G^p(0,T;\mathbb{R}^n)$ for each $Y \in M_G^p(0,T;\mathbb{R}^n)$, where b is given by Eq. (5.1.1).

Exercise 5.4.3 Complete the proof of Lemma 5.1.2.

Exercise 5.4.4 (The Gronwall inequality) Let $u(t)$ be a Lebesgue integrable function in $[0, T]$ such that

$$u(t) \leq C + A\int_0^t u(s)ds \ \text{ for } 0 \leq t \leq T,$$

where $C > 0$ and $A > 0$ are constants. Prove that $u(t) \leq Ce^{At}$ for $0 \leq t \leq T$.

Exercise 5.4.5 For any $\xi \in L_G^1(\Omega_T; \mathbb{R}^n)$, show that the process $(\hat{\mathbb{E}}[\xi|\Omega_t])_{t\in[0,T]}$ belongs to $\bar{M}_G^1(0, T; \mathbb{R}^n)$.

Exercise 5.4.6 Complete the proof of Lemma 5.2.1.

Exercise 5.4.7 Suppose that ξ, f and h_{ij} are all deterministic functions. Solve the BSDE (5.2.1).

Exercise 5.4.8 For each $\xi \in L_G^p(\Omega_t; \mathbb{R}^n)$ with $p \geq 2$, show that SDE (5.3.1) has a unique solution in $\bar{M}_G^p(t, T; \mathbb{R}^n)$. Further, show that the following estimates hold:

$$\hat{\mathbb{E}}_t[|X_{t+\delta}^{t,\xi} - X_{t+\delta}^{t,\xi'}|^p] \leq C|\xi - \xi'|^p,$$

$$\hat{\mathbb{E}}_t[|X_{t+\delta}^{t,\xi}|^p] \leq C(1 + |\xi|^p),$$

$$\hat{\mathbb{E}}_t[\sup_{s\in[t,t+\delta]} |X_s^{t,\xi} - \xi|^p] \leq C(1 + |\xi|^p)\delta^{p/2},$$

where the constant C depends on L, G, p, n and T.

Exercise 5.4.9 Let $\widetilde{\mathbb{E}}$ be a nonlinear expectation dominated by G-expectation, where $\widetilde{G} : \mathbb{S}(d) \mapsto \mathbb{R}$ is dominated by G and $\widetilde{G}(0) = 0$. Then we replace the G-expectation $\hat{\mathbb{E}}$ by $\widetilde{\mathbb{E}}$ in BSDEs (5.2.1) and (5.3.2). Show that

(i) the BSDE (5.2.1) admits a unique solution $Y \in \bar{M}_G^1(0, T)$.
(ii) u is the unique viscosity solution of the PDE (5.3.13) corresponding to $\widetilde{G}$.

Notes and Comments

The material in this chapter is mainly from Peng [140].

There are many excellent books on Itô's stochastic calculus and stochastic differential equations based by Itô's original paper [92]. The ideas of that notes were further developed to build the nonlinear martingale theory. For the corresponding classical Brownian motion framework under a probability measure space, readers are referred to Chung and Williams [34], Dellacherie and Meyer [43], He, Wang and Yan [74], Itô and McKean [93], Ikeda and Watanabe [90], Kallenberg [100], Karatzas and Shreve [101], Øksendal [122], Protter [150], Revuz and Yor [151] and Yong and Zhou [177].

Linear backward stochastic differential equations (BSDEs) were first introduced by Bismut in [17, 19]. Bensoussan developed this approach in [12, 13]. The existence and uniqueness theorem of a general nonlinear BSDE, was obtained in 1990 in Pardoux and Peng [124]. Here we present a version of a proof based on El Karoui, Peng and Quenez [58], which is an excellent survey paper on BSDE theory and its applications, especially in finance. Comparison theorem of BSDEs was obtained in Peng [128] for the case when g is a C^1-function and then in [58] when g is Lipschitz. Nonlinear Feynman-Kac formula for BSDE was introduced by Peng [127, 129]. Here we obtain the corresponding Feynman-Kac formula for a fully nonlinear PDE, within the framework of G-expectation. We also refer to Yong and Zhou [177], as well as Peng [131] (in 1997, in Chinese) and [133] and more resent monographs of Crepey [40], Pardoux and Rascanu [125] and Zhang [179] for systematic presentations of BSDE theory and its applications.

In the framework of fully nonlinear expectation, typically G-expectation, a challenging problem is to prove the well-posedness of a BSDE which is general enough to contain the above 'classical' BSDE as a special case. By applying and developing methods of quasi-surely analysis and aggregations, Soner et al. [156–158], introduced a weak formulation and then proved the existence and uniqueness of weak solution 2nd order BSDE (2BSDE). We also refer to Zhang [179] a systematic presentation. Then, by using a totally different approach of G-martingale representation and a type of Galerkin approximation, Hu et al. [79] proved the existence and uniqueness of solution of BSDE driven by G-Brownian motions (G-BSDE). As in the classical situation, G-BSDE is a natural generalization of representation of G-martingale. The assumption for the well-posedness of 2BSDEs is weaker than that of G-BSDE, whereas the solution (Y, Z, K) obtained by GBSDE is quasi-surely continuous which is in general smoother than that of 2BSDE. A very interesting problem is how to combine the advantages of both methods.

Then Hu and Wang [84] considered ergodic G-BSDEs, see also [77]. In [75], Hu, Lin and Soumana Hima studied G-BSDEs under quadratic assumptions on coefficients. In [111], Li, Peng and Soumana Hima investigated the existence and uniqueness theorem for reflected G-BSDEs. Furthermore, Cao and Tang [25] dealed with reflected Quadratic BSDEs driven by G-Brownian Motions.

Chapter 6
Capacity and Quasi-surely Analysis for G-Brownian Paths

In the last three chapters, we have considered random variables which are elements in a Banach space $L_G^p(\Omega)$. A natural question is whether such elements ξ are still real functions defined on Ω, namely $\xi = \xi(\omega)$, $\omega \in \Omega$. In this chapter we give an affirmative answer: each random variable $\xi \in L_G^p(\Omega)$ is a Borel-measurable function of Ω, and that hidden behind the G-expectation $\hat{\mathbb{E}}$, there exists a family of probability measures $\mathcal{P}$ defined on the measurable space $(\Omega, \mathcal{B}(\Omega))$ such that $\hat{\mathbb{E}}$ is the following type of upper expectation:

$$\hat{\mathbb{E}}[\xi] = \max_{P \in \mathcal{P}} E_P[\xi].$$

In this chapter, we first present a general framework for an upper expectation defined on a metric space $(\Omega, \mathcal{B}(\Omega))$ and the corresponding Choquet capacity to introduce the quasi-surely analysis. The results here are important because they allow us to develop the pathwise analysis for G -Brownian motion. Then we study stochastic process by quasi-surely analysis theory. We prove that, a random variable $\xi \in L_G^p(\Omega)$ is a quasi-continuous function, with respect to $\omega \in \Omega$. This is the generalization of the classical result that a random veritable $\xi \in L^p(\Omega, \mathcal{F}, P)$ is a P-quasi-continuous function in ω. A very important result of this chapter is that, quasi-surely, a G-Brownian motion $(B_t(\omega))_{t \geq t}$ is continuous in t.

6.1 Integration Theory Associated to Upper Probabilities

Let Ω be a complete separable metric space equipped with the distance d, $\mathcal{B}(\Omega)$ the Borel σ-algebra of Ω and $\mathcal{M}$ the collection of all probability measures on $(\Omega, \mathcal{B}(\Omega))$.

- $L^0(\Omega)$: the space of all $\mathcal{B}(\Omega)$ -measurable real functions;
- $B_b(\Omega)$: all bounded functions in $L^0(\Omega)$;
- $C_b(\Omega)$: all continuous functions in $B_b(\Omega)$.

Within this section, we consider a given subset $\mathcal{P} \subseteq \mathcal{M}$.

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_6

Definition 6.1.1 On $(\Omega, \mathcal{B}(\Omega))$, a sequence $\{P_i\}_{i\in\mathbb{N}}$ of probability measures is said to converge weakly to a probability measure P, if $\lim_{i\to\infty}\int_\Omega X(\omega)dP_i = \int_\Omega X(\omega)dP$, for any $X \in C_b(\Omega)$.

We recall the following classical result (see, for example, [151]):

Proposition 6.1.2 *The following conditions are equivalent:*

(i) $\{P_i\}_{i=1}^\infty$ *converges weakly to* P*;*
(ii) $\lim_{i\to\infty} E_{P_i}[X] = E_P[X]$, *for any* $X \in Lip(\Omega)$*;*
(iii) $\limsup_{i\to\infty} P_i(F) \le P(F)$, *for any closed subset* $F \subset \Omega$*;*
(iv) $\liminf_{i\to\infty} P_i(G) \ge P(G)$, *for any open subset* $G \subset \Omega$*.*

6.1.1 Capacity Associated with $\mathcal{P}$

We denote

$$c(A) = c^{\mathcal{P}}(A) := \sup_{P\in\mathcal{P}} P(A), \quad A \in \mathcal{B}(\Omega).$$

Definition 6.1.3 The set function c is called an upper probability associated with $\mathcal{P}$.

One can easily verify the following theorem.

Theorem 6.1.4 *The upper probability* $c(\cdot)$ *is a Choquet capacity, i.e. (see* [33, 42]*):*

1. $0 \le c(A) \le 1, \quad \forall A \subset \Omega$.
2. *If* $A \subset B$*, then* $c(A) \le c(B)$.
3. *If* $(A_n)_{n=1}^\infty$ *is a sequence in* $\mathcal{B}(\Omega)$*, then* $c(\cup_{n=1}^\infty A_n) \le \sum_{n=1}^\infty c(A_n)$.
4. *If* $(A_n)_{n=1}^\infty$ *is an increasing sequence in* $\mathcal{B}(\Omega)$*:* $A_n \uparrow A = \cup_{n=1}^\infty A_n$*, then* $c(\cup_{n=1}^\infty A_n) = \lim_{n\to\infty} c(A_n)$.

Remark 6.1.5 Note that in classical situations, there is another way to define Choquet capacity associated with $\mathcal{P}$, see Exercise 6.5.1. However in general it is different from the upper probability in Definition 6.1.3.

Further, we present a useful result.

Theorem 6.1.6 *For each* $A \in \mathcal{B}(\Omega)$*, we have*

$$c(A) = \sup\{c(K) : \ K \ \textit{is compact and} \ K \subset A\}.$$

Proof It is simply because

$$c(A) = \sup_{P\in\mathcal{P}} \sup_{\substack{K \text{ compact} \\ K\subset A}} P(K) = \sup_{\substack{K \text{ compact} \\ K\subset A}} \sup_{P\in\mathcal{P}} P(K) = \sup_{\substack{K \text{ compact} \\ K\subset A}} c(K).$$

□

Here and in what follows, we use a standard capacity-related terminology:

Definition 6.1.7 A set A is **polar** if $c(A) = 0$ and a property holds "**quasi-surely**" (q.s.) if it holds outside a polar set.

Remark 6.1.8 In other words, $A \in \mathcal{B}(\Omega)$ is polar if and only if $P(A) = 0$ for any $P \in \mathcal{P}$.

We also have in a trivial way a statement like the Borel–Cantelli Lemma.

Lemma 6.1.9 *Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of Borel sets such that*

$$\sum_{n=1}^{\infty} c(A_n) < \infty.$$

Then $\limsup_{n\to\infty} A_n$ *is polar.*

Proof Apply the classical Borel–Cantelli Lemma with respect to each probability $P \in \mathcal{P}$. □

The next result is the well-known Prokhorov's theorem expressed in the language of the capacity:

Theorem 6.1.10 *The set of the probability measures $\mathcal{P}$ is relatively compact if and only if for each $\varepsilon > 0$, there exists a compact set K such that $c(K^c) < \varepsilon$.*

The following two lemmas can be found in [89].

Lemma 6.1.11 *The set of the probability measures $\mathcal{P}$ is relatively compact if and only if for each sequence of closed sets $F_n \downarrow \emptyset$, we have $c(F_n) \downarrow 0$ as $n \to \infty$.*

Proof We outline the proof for readers' convenience.
"$\Longrightarrow$" part: It follows from Theorem 6.1.10 that for any fixed $\varepsilon > 0$, there exists a compact set K such that $c(K^c) < \varepsilon$. Note that $F_n \cap K \downarrow \emptyset$, then there exists an $N > 0$ such that $F_n \cap K = \emptyset$ for $n \geq N$, which implies $\lim_n c(F_n) < \varepsilon$. Since ε can be arbitrarily small, we obtain $c(F_n) \downarrow 0$.
"$\Longleftarrow$" part: For any $\varepsilon > 0$, let $(A_i^k)_{i=1}^{\infty}$ be a sequence of open balls of radius $1/k$ covering Ω. Observe that $(\cup_{i=1}^{n} A_i^k)^c \downarrow \emptyset$, then there exists an n_k such that $c((\cup_{i=1}^{n_k} A_i^k)^c) < \varepsilon 2^{-k}$. For the set $K = \overline{\cap_{k=1}^{\infty} \cup_{i=1}^{n_k} A_i^k}$, it is easy to check that K is compact and $c(K^c) < \varepsilon$. Thus by Theorem 6.1.10, $\mathcal{P}$ is relatively compact. □

Lemma 6.1.12 *Let $\mathcal{P}$ be weakly compact. Then for any sequence of closed sets $F_n \downarrow F$, we have $c(F_n) \downarrow c(F)$.*

Proof Here we also outline the proof. For any fixed $\varepsilon > 0$, by the definition of $c(F_n)$, there exists a $P_n \in \mathcal{P}$ such that $P_n(F_n) \geq c(F_n) - \varepsilon$. Since $\mathcal{P}$ is weakly compact, there exists a subsequence $\{P_{n_k}\}$ and $P \in \mathcal{P}$ such that P_{n_k} converges weakly to P. Thus

$$P(F_m) \geq \limsup_{k\to\infty} P_{n_k}(F_m) \geq \limsup_{k\to\infty} P_{n_k}(F_{n_k}) \geq \lim_{n\to\infty} c(F_n) - \varepsilon.$$

Letting $m \to \infty$, we get $P(F) \geq \lim_{n\to\infty} c(F_n) - \varepsilon$, which yields $c(F_n) \downarrow c(F)$. □

Following [89] (see also [45, 68]) the upper expectation of $\mathcal{P}$ is defined as follows: for any $X \in L^0(\Omega)$ such that $E_P[X]$ exists for each $P \in \mathcal{P}$,

$$\mathbb{E}[X] = \mathbb{E}^{\mathcal{P}}[X] := \sup_{P\in\mathcal{P}} E_P[X].$$

It is easy to verify the following properties:

Theorem 6.1.13 *The upper expectation $\mathbb{E}[\cdot]$ of the family $\mathcal{P}$ is a sublinear expectation on $B_b(\Omega)$ as well as on $C_b(\Omega)$, i.e.,*

1. *for all X, Y in $B_b(\Omega)$, $X \geq Y \Longrightarrow \mathbb{E}[X] \geq \mathbb{E}[Y]$.*
2. *for all X, Y in $B_b(\Omega)$, $\mathbb{E}[X + Y] \leq \mathbb{E}[X] + \mathbb{E}[Y]$.*
3. *for all $\lambda \geq 0$, $X \in B_b(\Omega)$, $\mathbb{E}[\lambda X] = \lambda\mathbb{E}[X]$.*
4. *for all $c \in \mathbb{R}$, $X \in B_b(\Omega)$, $\mathbb{E}[X + c] = \mathbb{E}[X] + c$.*

Moreover, the following properties hold:

Theorem 6.1.14 *We have*

1. *Let $\mathbb{E}[X_n]$ and $\mathbb{E}[\sum_{n=1}^{\infty} X_n]$ be finite. Then $\mathbb{E}[\sum_{n=1}^{\infty} X_n] \leq \sum_{n=1}^{\infty} \mathbb{E}[X_n]$.*
2. *Let $X_n \uparrow X$ and $\mathbb{E}[X_n]$, $\mathbb{E}[X]$ be finite. Then $\mathbb{E}[X_n] \uparrow \mathbb{E}[X]$.*

Definition 6.1.15 The functional $\mathbb{E}[\cdot]$ is said to be **regular** if for any sequence $\{X_n\}_{n=1}^{\infty}$ in $C_b(\Omega)$ such that $X_n \downarrow 0$ on Ω, we have $\mathbb{E}[X_n] \downarrow 0$.

Similar to Lemma 6.1.11 we have:

Theorem 6.1.16 *The sublinear expectation $\mathbb{E}[\cdot]$ is regular if and only if $\mathcal{P}$ is relatively compact.*

Proof "$\Longrightarrow$" part: For any sequence of closed subsets $F_n \downarrow \emptyset$ such that F_n, $n = 1, 2, \cdots$, are non-empty (otherwise the proof is trivial), there exists a sequence of functions $\{g_n\}_{n=1}^{\infty} \subset C_b(\Omega)$ satisfying

$$0 \leq g_n \leq 1, \quad g_n = 1 \text{ on } F_n \text{ and } g_n = 0 \text{ on } \{\omega \in \Omega : d(\omega, F_n) \geq \tfrac{1}{n}\}.$$

If we set $f_n = \wedge_{i=1}^{n} g_i$, it is clear that $f_n \in C_b(\Omega)$ and $\mathbf{1}_{F_n} \leq f_n \downarrow 0$. Since $\mathbb{E}[\cdot]$ is regular, this implies $\mathbb{E}[f_n] \downarrow 0$ and thus $c(F_n) \downarrow 0$. It follows from Lemma 6.1.11 that $\mathcal{P}$ is relatively compact.
"$\Longleftarrow$" part: For any $\{X_n\}_{n=1}^{\infty} \subset C_b(\Omega)$ such that $X_n \downarrow 0$, we have

$$\mathbb{E}[X_n] = \sup_{P\in\mathcal{P}} E_P[X_n] = \sup_{P\in\mathcal{P}} \int_0^{\infty} P(\{X_n \geq t\})dt \leq \int_0^{\infty} c(\{X_n \geq t\})dt.$$

For each fixed $t > 0$, $\{X_n \geq t\}$ is a closed subset and $\{X_n \geq t\} \downarrow \emptyset$ as $n \uparrow \infty$. By Lemma 6.1.11, $c(\{X_n \geq t\}) \downarrow 0$ and thus $\int_0^\infty c(\{X_n \geq t\})dt \downarrow 0$. Consequently $\mathbb{E}[X_n] \downarrow 0$. $\square$

6.1.2 Functional Spaces

We set, for $p > 0$, the following spaces:

- $\mathcal{L}^p := \{X \in L^0(\Omega) : \mathbb{E}[|X|^p] = \sup_{P \in \mathcal{P}} E_P[|X|^p] < \infty\}$;
- $\mathcal{N}^p := \{X \in L^0(\Omega) : \mathbb{E}[|X|^p] = 0\}$;
- $\mathcal{N} := \{X \in L^0(\Omega) : X = 0, c\text{-q.s.}\}$.

It is seen that $\mathcal{L}^p$ and $\mathcal{N}^p$ are linear spaces and $\mathcal{N}^p = \mathcal{N}$, for any $p > 0$.

We denote $\mathbb{L}^p := \mathcal{L}^p/\mathcal{N}$. As usual, we do not care about the distinction between classes and their representatives.

Lemma 6.1.17 *Let $X \in \mathbb{L}^p$. Then for each $\alpha > 0$*

$$c(\{|X| > \alpha\}) \leq \frac{\mathbb{E}[|X|^p]}{\alpha^p}.$$

Proof Just apply the classical Markov inequality with respect to each $P \in \mathcal{P}$. $\square$

Similar to the classical results, we derive now the following proposition which is similar to classical results.

Proposition 6.1.18 *We have*

1. *For any $p \geq 1$, $\mathbb{L}^p$ is a Banach space under the norm $\|X\|_p := (\mathbb{E}[|X|^p])^{1/p}$.*
2. *For any $p < 1$, $\mathbb{L}^p$ is a complete metric space under the distance $d(X, Y) := \mathbb{E}[|X - Y|^p]^{1/p}$.*

Proof For $p \geq 1$, $\|X\|_p = 0$ iff $X = 0$, q.s.. We also have $\|\lambda X\|_p = |\lambda| \cdot \|X\|_p$ for $\lambda \in \mathbb{R}$, and

$$\begin{aligned}
\|X + Y\|_p &= \left(\mathbb{E}\left[|X + Y|^p\right]\right)^{1/p} = \sup_{P \in \mathcal{P}} \left(E_P[|X + Y|^p]\right)^{1/p} \\
&\leq \sup_{P \in \mathcal{P}} \{(E_P\left[|X|^p\right])^{1/p} + \left(E_P[|Y|^p]\right)^{1/p}\} \\
&\leq \sup_{P \in \mathcal{P}} \left(E_P[|X|^p]\right)^{1/p} + \sup_{P \in \mathcal{P}} \left(E_P[|Y|^p]\right)^{1/p} \\
&= \|X\|_p + \|Y\|_p .
\end{aligned}$$

Thus $\|\cdot\|_p$ is a norm in $\mathbb{L}^p$. Now let $\{X_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $\mathbb{L}^p$. We choose a subsequence $\left\{X_{n_i}\right\}_{i=1}^{\infty}$ satisfying $\left\|X_{n_{i+1}}-X_{n_i}\right\|_p \leq 2^{-i}$, $i=1,2,\cdots$. Then

$$\begin{aligned}\left\|\sum_{i=1}^{\infty}|X_{n_{i+1}}-X_{n_i}|\right\|_p &= \sup_{P\in\mathcal{P}}\left(E_P\left[\left|\sum_{i=1}^{\infty}|X_{n_{i+1}}-X_{n_i}|\right|^p\right]\right)^{1/p} \\ &\leq \sup_{P\in\mathcal{P}}\sum_{i=1}^{\infty}\left(E_P\left[\left|X_{n_{i+1}}-X_{n_i}\right|^p\right]\right)^{1/p} \\ &\leq \sum_{i=1}^{\infty}\left\|X_{n_{i+1}}-X_{n_i}\right\|_p \leq 1.\end{aligned}$$

It follows that $\sum_{i=1}^{\infty}|X_{n_{i+1}}-X_{n_i}|<\infty$, q.s.. We now let $X=X_{n_1}+\sum_{i=1}^{\infty}(X_{n_{i+1}}-X_{n_i})$. This function is q.s. defined on Ω. We also have

$$\|X\|_p \leq \left\|X_{n_1}\right\|_p+\sum_{i=1}^{\infty}\left\|X_{n_{i+1}}-X_{n_i}\right\|_p \leq 1+\left\|X_{n_1}\right\|_p<\infty.$$

Hence $X\in\mathbb{L}^p$. On the other hand,

$$\begin{aligned}\left\|X_{n_k}-X\right\|_p &= \sup_{P\in\mathcal{P}}\left(E_P[|\sum_{i=k}^{\infty}(X_{n_{i+1}}-X_{n_i})|^p]\right)^{\frac{1}{p}} \leq \sum_{i=k}^{\infty}\left\|X_{n_{i+1}}-X_{n_i}\right\|_p \\ &\leq 2^{-(k-1)}\to 0,\ \text{as } k\to\infty.\end{aligned}$$

Since $\{X_n\}_{n=1}^{\infty}$ itself is a Cauchy sequence in $\mathbb{L}^p$, we get $\|X_n-X\|_p\to 0$. So $\mathbb{L}^p$ is a Banach space. The proof for $p<1$ is similar. □

We set

$$\begin{aligned}\mathcal{L}^{\infty} &:= \{X\in L^0(\Omega): \exists \text{ a constant } M, \text{ s.t. } |X|\leq M, \quad \text{q.s.}\}; \\ \mathbb{L}^{\infty} &:= \mathcal{L}^{\infty}/\mathcal{N}.\end{aligned}$$

Proposition 6.1.19 *Under the norm*

$$\|X\|_{\infty} := \inf\left\{M\geq 0: |X|\leq M, \quad q.s.\right\},$$

$\mathbb{L}^{\infty}$ *is a Banach space.*

Proof From $\{|X|>\|X\|_{\infty}\}=\cup_{n=1}^{\infty}\left\{|X|\geq\|X\|_{\infty}+\frac{1}{n}\right\}$ we know that $|X|\leq\|X\|_{\infty}$, q.s., then it is easy to check that $\|\cdot\|_{\infty}$ is a norm. The proof of the completeness of $\mathbb{L}^{\infty}$ is similar to the classical result. □

With respect to the distance defined on $\mathbb{L}^p$, $p > 0$, we denote by

- $\mathbb{L}_b^p$ the completion of $B_b(\Omega)$,
- $\mathbb{L}_c^p$ the completion of $C_b(\Omega)$.

By Proposition 6.1.18, we have the inclusions:

$$\mathbb{L}_c^p \subset \mathbb{L}_b^p \subset \mathbb{L}^p, \quad p > 0.$$

The following proposition is obvious and the proof is left to the reader.

Proposition 6.1.20 *We have three statements:*

1. *Let* $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. *Then* $X \in \mathbb{L}^p$ *and* $Y \in \mathbb{L}^q$ *implies*
$$XY \in \mathbb{L}^1 \text{ and } \mathbb{E}[|XY|] \leq \left(\mathbb{E}[|X|^p]\right)^{1/p} \left(\mathbb{E}[|Y|^q]\right)^{1/q};$$
Moreover, $X \in \mathbb{L}_c^p$ *and* $Y \in \mathbb{L}_c^q$ *imply that* $XY \in \mathbb{L}_c^1$*;*
2. $\mathbb{L}^{p_1} \subset \mathbb{L}^{p_2}$, $\mathbb{L}_b^{p_1} \subset \mathbb{L}_b^{p_2}$, $\mathbb{L}_c^{p_1} \subset \mathbb{L}_c^{p_2}$, $0 < p_2 \leq p_1 \leq \infty$*;*
3. $\|X\|_p \uparrow \|X\|_\infty$, *as* $p \to \infty$, *for any* $X \in \mathbb{L}^\infty$.

Proposition 6.1.21 *Let* $p \in (0, \infty]$ *and let* (X_n) *be a sequence in* $\mathbb{L}^p$ *which converges to* X *in* $\mathbb{L}^p$. *Then there exists a subsequence* (X_{n_k}) *which converges to* X *quasi-surely in the sense that it converges to* X *outside a polar set.*

Proof Since convergence in $\mathbb{L}^\infty$ implies convergence in $\mathbb{L}^p$ for all p, we only need to consider the case $p \in (0, \infty)$. We first extract a subsequence (X_{n_k}) such that

$$\mathbb{E}[|X - X_{n_k}|^p] \leq 1/k^{p+2}, \quad k \in \mathbb{N},$$

and set $A_k = \{|X - X_{n_k}| > 1/k\}$ for $k = 1, 2, \cdots$. Then, by the Markov inequality in Lemma 6.1.17, we have

$$c(A_k) \leq k^{-2}.$$

As a consequence of the Borel–Cantelli Lemma 6.1.9, we have $c(\overline{\lim}_{k\to\infty} A_k) = 0$. It follows that outside of this polar set, $X_{n_k}(\omega)$ converges to $X(\omega)$. The proposition is proved. □

We now give a description of $\mathbb{L}_b^p$.

Proposition 6.1.22 *For any* $p > 0$, *the following relations hold:*

$$\mathbb{L}_b^p = \{X \in \mathbb{L}^p : \lim_{n\to\infty} \mathbb{E}[(|X|^p - n)^+] = 0\} \tag{6.1.1}$$

$$= \{X \in \mathbb{L}^p : \lim_{n\to\infty} \mathbb{E}[|X|^p \mathbf{1}_{\{|X|>n\}}] = 0\}. \tag{6.1.2}$$

Proof Since $X \in \mathbb{L}_b^p$ if and only if $|X|^p \in \mathbb{L}_b^1$, it suffices to show that (6.1.1) and (6.1.2) are valid for $p = 1$. We denote $J_1 = \{X \in \mathbb{L}^1 : \lim_{n\to\infty} \mathbb{E}[(|X| - n)^+] = 0\}$. For any $X \in J_1$ let $X_n = (X \wedge n) \vee (-n) \in B_b(\Omega)$. We obtain that

$$\mathbb{E}[|X - X_n|] = \mathbb{E}[(|X| - n)^+] \to 0, \text{ as } n \to \infty.$$

Thus $X \in \mathbb{L}_b^1$.

On the other hand, for any $X \in \mathbb{L}_b^1$, we can find a sequence $\{Y_n\}_{n=1}^{\infty}$ in $B_b(\Omega)$ such that $\mathbb{E}[|X - Y_n|] \to 0$. Let $y_n = n + \sup_{\omega\in\Omega} |Y_n(\omega)|$. Now we have

$$\begin{aligned}\mathbb{E}[(|X| - y_n)^+] &\le \mathbb{E}[(|X| - |Y_n|)^+] + \mathbb{E}[(|Y_n| - y_n)^+] \\ &\le \mathbb{E}[(|X| - |Y_n|)^+] \to 0, \quad \text{as } n \to \infty.\end{aligned}$$

Consequently $X \in J_1$.

Relation (6.1.2) follows from the inequalities

$$(|x|^p - c)^+ \le |x|^p \mathbf{1}_{\{|x|^p > c\}} \le (2|x|^p - c)^+, \quad \forall c > 0, \quad x \in \mathbb{R}.$$

The proof is complete. □

Proposition 6.1.23 *Let $X \in \mathbb{L}_b^1$. Then for any $\varepsilon > 0$, there exists a number $\delta > 0$, such that for all $A \in \mathcal{B}(\Omega)$ with $c(A) \le \delta$, we have $\mathbb{E}[|X|\mathbf{1}_A] \le \varepsilon$.*

Proof For any $\varepsilon > 0$, by Proposition 6.1.22, there exists an $N > 0$ such that $\mathbb{E}[|X|\mathbf{1}_{\{|X|>N\}}] \le \frac{\varepsilon}{2}$. Take $\delta = \frac{\varepsilon}{2N}$. Then for a subset $A \in \mathcal{B}(\Omega)$ with $c(A) \le \delta$, we obtain

$$\begin{aligned}\mathbb{E}[|X|\mathbf{1}_A] &\le \mathbb{E}[|X|\mathbf{1}_A\mathbf{1}_{\{|X|>N\}}] + \mathbb{E}[|X|\mathbf{1}_A\mathbf{1}_{\{|X|\le N\}}] \\ &\le \mathbb{E}[|X|\mathbf{1}_{\{|X|>N\}}] + Nc(A) \le \varepsilon.\end{aligned}$$

□

It is important to note that not every element in $\mathbb{L}^p$ satisfies the condition (6.1.2). We give the following two counterexamples, which show that $\mathbb{L}^1$ and $\mathbb{L}_b^1$ are different spaces even in the case when $\mathcal{P}$ is weakly compact.

Example 6.1.24 Let $\Omega = \mathbb{N}$, $\mathcal{P} = \{P_n : n \in \mathbb{N}\}$ where $P_1(\{1\}) = 1$ and $P_n(\{1\}) = 1 - \frac{1}{n}$, $P_n(\{n\}) = \frac{1}{n}$, for $n = 2, 3, \cdots$. The set $\mathcal{P}$ is weakly compact. We consider a function X on $\mathbb{N}$ defined by $X(n) = n$, $n \in \mathbb{N}$. We have $\mathbb{E}[|X|] = 2$, however $\mathbb{E}[|X|\mathbf{1}_{\{|X|>n\}}] = 1 \nrightarrow 0$. In this case, $X \in \mathbb{L}^1$ and $X \notin \mathbb{L}_b^1$.

Example 6.1.25 Let $\Omega = \mathbb{N}$, $\mathcal{P} = \{P_n : n \in \mathbb{N}\}$ where $P_1(\{1\}) = 1$ and $P_n(\{1\}) = 1 - \frac{1}{n^2}$, $P_n(\{kn\}) = \frac{1}{n^3}$, $k = 1, 2, \ldots, n$, for $n = 2, 3, \cdots$. The set $\mathcal{P}$ is weakly compact. Consider a function X on $\mathbb{N}$ defined by $X(n) = n$, $n \in \mathbb{N}$. We easily find that $\mathbb{E}[|X|] = \frac{25}{16}$ and $n\mathbb{E}[\mathbf{1}_{\{|X|\ge n\}}] = \frac{1}{n} \to 0$, however $\mathbb{E}[|X|\mathbf{1}_{\{|X|\ge n\}}] = \frac{1}{2} + \frac{1}{2n} \nrightarrow 0$. In this case, X is in $\mathbb{L}^1$, continuous and $n\mathbb{E}[\mathbf{1}_{\{|X|\ge n\}}] \to 0$, however it is not in $\mathbb{L}_b^1$.

6.1.3 Properties of Elements of $\mathbb{L}_c^p$

Definition 6.1.26 A mapping X on Ω with values in a topological space is said to be quasi-continuous (q.c.) if

$$\forall \varepsilon > 0, \text{ there exists an open set } O \text{ with } c(O) < \varepsilon \text{ such that } X|_{O^c} \text{ is continuous.}$$

Definition 6.1.27 We say that $X : \Omega \mapsto \mathbb{R}$ has a quasi-continuous version if there exists a quasi-continuous function $Y : \Omega \mapsto \mathbb{R}$ such that $X = Y$ q.s..

Proposition 6.1.28 *Let $p > 0$. Then each element of $\mathbb{L}_c^p$ has a quasi-continuous version.*

Proof Let (X_n) be a Cauchy sequence in $C_b(\Omega)$ with respect to the distance on $\mathbb{L}^p$. Let us choose a subsequence $(X_{n_k})_{k\geq 1}$ such that

$$\mathbb{E}[|X_{n_{k+1}} - X_{n_k}|^p] \leq 2^{-2k}, \quad \forall k \geq 1,$$

and set for all k,

$$A_k = \bigcup_{i=k}^{\infty} \{|X_{n_{i+1}} - X_{n_i}| > 2^{-i/p}\}.$$

Thanks to the subadditivity property and the Markov inequality, we derive that

$$c(A_k) \leq \sum_{i=k}^{\infty} c(|X_{n_{i+1}} - X_{n_i}| > 2^{-i/p}) \leq \sum_{i=k}^{\infty} 2^{-i} = 2^{-k+1}.$$

As a consequence, $\lim_{k\to\infty} c(A_k) = 0$, so the Borel set $A = \bigcap_{k=1}^{\infty} A_k$ is polar.

As each X_{n_k} is continuous, for all $k \geq 1$, A_k is an open set. Moreover, for all k, (X_{n_i}) converges uniformly on A_k^c so that the limit is continuous on each A_k^c. This yields the result. □

The following theorem gives a concrete characterization of the space $\mathbb{L}_c^p$.

Theorem 6.1.29 *For each $p > 0$,*

$$\mathbb{L}_c^p = \{X \in \mathbb{L}^p : X \textit{ has a quasi-continuous version and } \lim_{n\to\infty} \mathbb{E}[(|X|^p - n)^+] = 0\}.$$

Proof We denote

$$J_p = \{X \in \mathbb{L}^p : X \text{ has a quasi-continuous version and } \lim_{n\to\infty} \mathbb{E}[(|X|^p - n)^+] = 0\}.$$

If $X \in \mathbb{L}_c^p$, we know by Proposition 6.1.28 that X has a quasi-continuous version. Since X is also an element of $\mathbb{L}_b^p$, we have by Proposition 6.1.22 that $\lim_{n\to\infty} \mathbb{E}[(|X|^p - n)^+] = 0$. Thus $X \in J_p$.

On the other hand, let $X \in J_p$ be quasi-continuous. Define $Y_n = (X \wedge n) \vee (-n)$ for any $n \in \mathbb{N}$. Since $\lim_{n\to\infty} \mathbb{E}[(|X|^p - n)^+] = 0$, we find that $\mathbb{E}[|X - Y_n|^p] \to 0$.

Moreover, for any $n \in \mathbb{N}$, since Y_n is quasi-continuous, there exists a closed set F_n such that $c(F_n^c) < \frac{1}{n^{p+1}}$ and Y_n is continuous on F_n. It follows from Tietze's extension theorem that there exists $Z_n \in C_b(\Omega)$ such that

$$|Z_n| \le n \text{ and } Z_n = Y_n \text{ on } F_n.$$

We then have

$$\mathbb{E}[|Y_n - Z_n|^p] \le (2n)^p c(F_n^c) \le \frac{(2n)^p}{n^{p+1}}.$$

Hence $\mathbb{E}[|X - Z_n|^p] \le (1 \vee 2^{p-1})(\mathbb{E}[|X - Y_n|^p] + \mathbb{E}[|Y_n - Z_n|^p]) \to 0$, and $X \in \mathbb{L}_c^p$. □

We now provide an example to show that $\mathbb{L}_c^p$ is different from $\mathbb{L}_b^p$ even if the set $\mathcal{P}$ is weakly compact.

Example 6.1.30 In the case $\Omega = [0, 1]$, the family of probabilities $\mathcal{P} = \{\delta_x : x \in [0, 1]\}$ is weakly compact. It is seen that $\mathbb{L}_c^p = C_b(\Omega)$ which is different from $\mathbb{L}_b^p$.

We denote $\mathbb{L}_c^\infty := \{X \in \mathbb{L}^\infty : X \text{ has a quasi-continuous version}\}$.

Proposition 6.1.31 *The space $\mathbb{L}_c^\infty$ is a closed linear subspace of $\mathbb{L}^\infty$.*

Proof For each Cauchy sequence $\{X_n\}_{n=1}^\infty$ in $\mathbb{L}_c^\infty$ under $\|\cdot\|_\infty$, we can find a subsequence $\{X_{n_i}\}_{i=1}^\infty$ such that $\|X_{n_{i+1}} - X_{n_i}\|_\infty \le 2^{-i}$. We may further assume that any X_n is quasi-continuous. Then it is easy to show that for each $\varepsilon > 0$, there exists an open set O such that $c(O) < \varepsilon$ and $|X_{n_{i+1}} - X_{n_i}| \le 2^{-i}$ for all $i \ge 1$ on O^c. This implies that the limit belongs to $\mathbb{L}_c^\infty$. □

As an application of Theorem 6.1.29, we can easily get the following results.

Proposition 6.1.32 *Assume that $X : \Omega \mapsto \mathbb{R}$ has a quasi-continuous version and that there exists a function $f : \mathbb{R}^+ \mapsto \mathbb{R}^+$ satisfying $\lim_{t\to\infty} \frac{f(t)}{t^p} = \infty$ and $\mathbb{E}[f(|X|)] < \infty$. Then $X \in \mathbb{L}_c^p$.*

Proof For any $\varepsilon > 0$, there exists an $N > 0$ such that $\frac{f(t)}{t^p} \ge \frac{1}{\varepsilon}$, for all $t \ge N$. Thus

$$\mathbb{E}[|X|^p \mathbf{1}_{\{|X|>N\}}] \le \varepsilon \mathbb{E}[f(|X|)\mathbf{1}_{\{|X|>N\}}] \le \varepsilon \mathbb{E}[f(|X|)].$$

Hence $\lim_{N\to\infty} \mathbb{E}[|X|^p \mathbf{1}_{\{|X|>N\}}] = 0$. From Theorem 6.1.29 we deduce that $X \in \mathbb{L}_c^p$. □

Lemma 6.1.33 *Let $\{P_n\}_{n=1}^\infty \subset \mathcal{P}$ converges weakly to $P \in \mathcal{P}$ as $n \to \infty$. Then for each $X \in \mathbb{L}_c^1$, we have $E_{P_n}[X] \to E_P[X]$.*

Proof We may assume that X is quasi-continuous, otherwise we can consider its quasi-continuous version which does not change the value E_Q for each $Q \in \mathcal{P}$. For any $\varepsilon > 0$, there exists $Y \in C_b(\Omega)$ such $\mathbb{E}[|X - Y|] \leq \varepsilon$. Obviously, for each $Q \in \mathcal{P}$,

$$|E_Q[X] - E_Q[Y]| \leq E_Q[|X - Y|] \leq \mathbb{E}[|X - Y|] \leq \varepsilon.$$

It then follows that

$$\limsup_{n\to\infty} E_{P_n}[X] \leq \lim_{n\to\infty} E_{P_n}[Y] + \varepsilon = E_P[Y] + \varepsilon \leq E_P[X] + 2\varepsilon.$$

Similarly, we obtain that $\liminf_{n\to\infty} E_{P_n}[X] \geq E_P[X] - 2\varepsilon$. Since ε can be arbitrarily small, we arrive at the required convergence $E_{P_n}[X] \to E_P[X]$. □

Remark 6.1.34 For the case $X \in C_b(\Omega)$, the above Lemma 6.1.33 implies Lemma 3.8.7 in [21].

Now we give an extension of Theorem 6.1.16.

Theorem 6.1.35 *Let $\mathcal{P}$ be weakly compact and let $\{X_n\}_{n=1}^{\infty} \subset \mathbb{L}_c^1$ and $X \in \mathbb{L}^1$ be such that $X_n \downarrow X$, q.s., as $n \to \infty$. Then $\mathbb{E}[X_n] \downarrow \mathbb{E}[X]$.*

Remark 6.1.36 It is important to note that X does not necessarily belong to $\mathbb{L}_c^1$.

Proof In the case $\mathbb{E}[X] > -\infty$, if there exists $\delta > 0$ such that $\mathbb{E}[X_n] > \mathbb{E}[X] + \delta$, $n = 1, 2, \cdots$, we can find a probability measure $P_n \in \mathcal{P}$ such that $E_{P_n}[X_n] > \mathbb{E}[X] + \delta - \frac{1}{n}$, $n = 1, 2, \cdots$. Since $\mathcal{P}$ is weakly compact, there is a subsequence $\{P_{n_i}\}_{i=1}^{\infty}$ that converges weakly to some $P \in \mathcal{P}$. This implies that

$$\begin{aligned} E_P[X_{n_i}] &= \lim_{j\to\infty} E_{P_{n_j}}[X_{n_i}] \geq \limsup_{j\to\infty} E_{P_{n_j}}[X_{n_j}] \\ &\geq \limsup_{j\to\infty} \left\{ \mathbb{E}[X] + \delta - \frac{1}{n_j} \right\} = \mathbb{E}[X] + \delta, i = 1, 2, \cdots . \end{aligned}$$

Thus $E_P[X] \geq \mathbb{E}[X] + \delta$. This contradicts the definition of $\mathbb{E}[\cdot]$. The arguments in the case $\mathbb{E}[X] = -\infty$ are analogous. □

We immediately have the following corollary.

Corollary 6.1.37 *Let $\mathcal{P}$ be weakly compact and let $\{X_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{L}_c^1$ which is decreasing and converging to $X \in \mathbb{L}_c^1$ q.s.. Then $\mathbb{E}[|X_n - X|] \downarrow 0$, as $n \to \infty$.*

6.1.4 Kolmogorov's Criterion

Definition 6.1.38 Let I be a set of indices, $(X_t)_{t\in I}$ and $(Y_t)_{t\in I}$ two processes indexed by I. We say that Y is a quasi-modification of X if for all $t \in I$, $X_t = Y_t$ q.s..

Remark 6.1.39 In Definition 6.1.38, a quasi-modification is also called a modification in some papers.

We now give the Kolmogorov criterion for processes indexed by $\mathbb{R}^d$ with $d \in \mathbb{N}$.

Theorem 6.1.40 *Let $p > 0$ and $(X_t)_{t\in[0,1]^d}$ be a process such that for all $t \in [0, 1]^d$, X_t belongs to $\mathbb{L}^p$. Assume that there exist positive constants c and ε such that*

$$\mathbb{E}[|X_t - X_s|^p] \leq c|t - s|^{d+\varepsilon}.$$

Then X admits a modification $\tilde{X}$ which satisfies the following relation:

$$\mathbb{E}\left[\left(\sup_{s\neq t} \frac{|\tilde{X}_t - \tilde{X}_s|}{|t - s|^\alpha}\right)^p\right] < \infty,$$

for every $\alpha \in [0, \varepsilon/p)$. As a consequence, the trajectories of $\tilde{X}$ are quasi-surely Hölder continuous of order α for every $\alpha < \varepsilon/p$ in the sense that there exists a Borel set N of capacity 0 such that for all $w \in N^c$, the map $t \mapsto \tilde{X}(w)$ is Hölder continuous of order α for every $\alpha < \varepsilon/p$. Moreover, if $X_t \in \mathbb{L}_c^p$ for each t, then we also have $\tilde{X}_t \in \mathbb{L}_c^p$.

Proof Let D be the set of dyadic points in $[0, 1]^d$:

$$D_n = \left\{\left(\frac{i_1}{2^n}, \cdots, \frac{i_d}{2^n}\right); i_1, \cdots, i_d \in \{0, 1, \cdots, 2^n\}\right\}, \quad D = \bigcup_{n=1}^{\infty} D_n.$$

Let $\alpha \in [0, \varepsilon/p)$. We define

$$M = M(\omega) = \sup_{s,t\in D, s\neq t} \frac{|X_t - X_s|}{|t - s|^\alpha}, \quad M_n = \sup_{s,t\in D_n, s\neq t} \frac{|X_t - X_s|}{|t - s|^\alpha}.$$

From the classical Kolmogorov's criterion (see Revuz–Yor [151]), we know that for any $P \in \mathcal{P}$, $E_P[M^p]$ is finite and uniformly bounded with respect to P and such that

$$\mathbb{E}[M^p] = \sup_{P\in\mathcal{P}} E_P[M^p] < \infty.$$

As a consequence, the map $t \mapsto X_t$ is uniformly continuous on D quasi-surely and so we can define process $\tilde{X}$ as follows:

$$\forall t \in [0, 1]^d, \ \tilde{X}_t = \lim_{s\to t, s\in D} X_s.$$

It is now clear that $\tilde{X}$ satisfies the required properties. □

Remark 6.1.41 A particularly interesting example of the above stochastic process X is the 1-dimensional G-Brownian motion path $B_t(\omega) = \omega_t$, $t \in [0, 1]$, for the case $d = 1$. It is easy to prove that M is a convex functional of the Brownian motion paths ω. On the other hand, for any functional of $\xi(\omega)$ of G-Brownian motional path which belongs to $L_G^p(\Omega)$, one can check that $\hat{\mathbb{E}}[\xi] = E_{P_{\overline{\sigma}}}[\xi(\omega)]$, where $P_{\overline{\sigma}}$ is the probability induced by $\overline{\sigma}W$ and W is a classical standard 1-dimensional Brownian motion. A very interesting problem is: can we just use the results of Chap. 3 to prove that $\hat{\mathbb{E}}[M(\omega)] = E_{P_{\overline{\sigma}}}[M(\omega)]$? A positive answer of this question allows us to obtain the corresponding Kolmogorov's criterion for G-Brownian motion in a much simple way.

6.2 G-Expectation as an Upper Expectation

6.2.1 Construction of G-Brownian Motion Through Its Finite Dimensional Distributions

In the following sections of this book, unless otherwise mentioned, we always denote by $\Omega = C_0^d(\mathbb{R}^+)$, the space of all $\mathbb{R}^d-$valued continuous functions $(\omega_t)_{t\in\mathbb{R}^+}$, with $\omega_0 = 0$, equipped with the distance

$$\rho(\omega^1, \omega^2) := \sum_{i=1}^{\infty} 2^{-i}[(\max_{t\in[0,i]} |\omega_t^1 - \omega_t^2|) \wedge 1],$$

and let $\bar{\Omega} = (\mathbb{R}^d)^{[0,\infty)}$ denote the space of all $\mathbb{R}^d-$valued functions $(\bar{\omega}_t)_{t\in\mathbb{R}^+}$.

We also denote by $\mathcal{B}(\Omega)$, the σ-algebra generated by all open sets and let $\mathcal{B}(\bar{\Omega})$ be the σ -algebra generated by all finite dimensional cylinder sets. The corresponding canonical process is $B_t(\omega) = \omega_t$ (respectively, $\bar{B}_t(\bar{\omega}) = \bar{\omega}_t$), $t \in [0, \infty)$ for $\omega \in \Omega$ (respectively, $\bar{\omega} \in \bar{\Omega}$).

In this section we construct a family of probabilities $\mathcal{P}$ on $(\Omega, \mathcal{B}(\Omega))$ such that the G-expectation can be represented as an upper expectation, namely,

$$\hat{\mathbb{E}}[\cdot] = \max_{P\in\mathcal{P}} E_P[\cdot].$$

The spaces of Lipschitz cylinder functions on Ω and $\bar{\Omega}$ are denoted respectively by

$$Lip(\Omega) := \{\varphi(B_{t_1}, B_{t_2}, \cdots, B_{t_n}) : \forall n \geq 1, t_1, \cdots, t_n \in [0, \infty), \forall \varphi \in C_{l.Lip}(\mathbb{R}^{d\times n})\},$$

$$Lip(\bar{\Omega}) := \{\varphi(\bar{B}_{t_1}, \bar{B}_{t_2}, \cdots, \bar{B}_{t_n}) : \forall n \geq 1, t_1, \cdots, t_n \in [0, \infty), \forall \varphi \in C_{l.Lip}(\mathbb{R}^{d\times n})\}.$$

Let $G(\cdot):\mathbb{S}(d)\mapsto\mathbb{R}$ be a given continuous monotone and sublinear function. Following Sect. 6.2 in Chap. 3, we can construct the corresponding G-expectation $\hat{\mathbb{E}}$ on $(\Omega, Lip(\Omega))$. Due to the natural correspondence of $Lip(\bar{\Omega})$ and $Lip(\Omega)$, we also construct a sublinear expectation $\bar{\mathbb{E}}$ on $(\bar{\Omega}, Lip(\bar{\Omega}))$ such that $(\bar{B}_t(\bar{\omega}))_{t\geq 0}$ is a (symmetric) G-Brownian motion.

The main objective in this section is to find a weakly compact family of (σ-additive) probability measures on $(\Omega, \mathcal{B}(\Omega))$ and use them to represent thepg G-expectation $\hat{\mathbb{E}}$. The following lemmas are variations of Lemma 1.3.4 and 1.3.5 in Chap. 1.

Lemma 6.2.1 *Let $0\leq t_1<t_2<\cdots<t_m<\infty$ and $\{\varphi_n\}_{n=1}^{\infty}\subset C_{l.Lip}(\mathbb{R}^{d\times m})$ satisfy $\varphi_n\downarrow 0$ as $n\to\infty$. Then $\bar{\mathbb{E}}[\varphi_n(\bar{B}_{t_1},\bar{B}_{t_2},\cdots,\bar{B}_{t_m})]\downarrow 0$.*

We denote $\mathcal{T}:=\{\underline{t}=(t_1,\ldots,t_m):\forall m\in\mathbb{N}, 0\leq t_1<t_2<\cdots<t_m<\infty\}$.

Lemma 6.2.2 *Let E be a finitely additive linear expectation dominated by $\bar{\mathbb{E}}$ on $Lip(\bar{\Omega})$. Then there exists a unique probability measure Q on $(\bar{\Omega}, \mathcal{B}(\bar{\Omega}))$ such that $E[X]=E_Q[X]$ for each $X\in Lip(\bar{\Omega})$.*

Proof For any fixed $\underline{t}=(t_1,\ldots,t_m)\in\mathcal{T}$, by Lemma 6.2.1, for each sequence $\{\varphi_n\}_{n=1}^{\infty}\subset C_{l.Lip}(\mathbb{R}^{d\times m})$ satisfying $\varphi_n\downarrow 0$, we have $E[\varphi_n(\bar{B}_{t_1},\bar{B}_{t_2},\cdots,\bar{B}_{t_m})]\downarrow 0$. By Daniell-Stone's theorem (see Appendix B), there exists a unique probability measure $Q_{\underline{t}}$ on $(\mathbb{R}^{d\times m},\mathcal{B}(\mathbb{R}^{d\times m}))$ such that $E_{Q_{\underline{t}}}[\varphi]=E[\varphi(\bar{B}_{t_1},\bar{B}_{t_2},\cdots,\bar{B}_{t_m})]$ for each $\varphi\in C_{l.Lip}(\mathbb{R}^{d\times m})$. Thus we get a family of finite dimensional distributions $\{Q_{\underline{t}}:\underline{t}\in\mathcal{T}\}$. It is easy to check that $\{Q_{\underline{t}}:\underline{t}\in\mathcal{T}\}$ is a consistent family. Then by Kolmogorov's consistent extension theorem, there exists a probability measure Q on $(\bar{\Omega},\mathcal{B}(\bar{\Omega}))$ such that $\{Q_{\underline{t}}:\underline{t}\in\mathcal{T}\}$ is the finite dimensional distributions of Q. We now prove the uniqueness. Assume that there exists another probability measure $\bar{Q}$ satisfying the condition. By Daniell-Stone's theorem, Q and $\bar{Q}$ have the same finite-dimensional distributions, hence by the monotone class theorem, $Q=\bar{Q}$. The proof is complete. □

Lemma 6.2.3 *There exists a family of probability measures $\mathcal{P}_e$ on $(\bar{\Omega},\mathcal{B}(\bar{\Omega}))$ such that*

$$\bar{\mathbb{E}}[X]=\max_{Q\in\mathcal{P}_e}E_Q[X],\quad \textit{for } X\in Lip(\bar{\Omega}).$$

Proof By the representation theorem of the sublinear expectation and Lemma 6.2.2, it is easy to get the result. □

For this $\mathcal{P}_e$, we define the associated capacity:

$$\tilde{c}(A):=\sup_{Q\in\mathcal{P}_e}Q(A),\quad A\in\mathcal{B}(\bar{\Omega}),$$

and the upper expectation for each $\mathcal{B}(\bar{\Omega})$-measurable real function X which makes the following definition meaningful:

$$\tilde{\mathbb{E}}[X] := \sup_{Q \in \mathcal{P}_e} E_Q[X].$$

Theorem 6.2.4 *For $(\bar{B}_t)_{t\geq 0}$, there exists a continuous modification $(\tilde{B}_t)_{t\geq 0}$ of $\bar{B}$ in the sense that $\tilde{c}(\{\tilde{B}_t \neq \bar{B}_t\}) = 0$, for each $t \geq 0$ and such that $\tilde{B}_0 = 0$.*

Proof By Lemma 6.2.3, we know that $\bar{\mathbb{E}} = \tilde{\mathbb{E}}$ on $Lip(\bar{\Omega})$. On the other hand, we have

$$\tilde{\mathbb{E}}[|\bar{B}_t - \bar{B}_s|^4] = \bar{\mathbb{E}}[|\bar{B}_t - \bar{B}_s|^4] = d|t-s|^2 \text{ for } s, t \in [0, \infty),$$

where d is a constant depending only on G. By Theorem 6.1.40, there exists a continuous modification $\tilde{B}$ of $\bar{B}$. Since $\tilde{c}(\{\bar{B}_0 \neq 0\}) = 0$, we can set $\tilde{B}_0 = 0$. The proof is complete. □

For any $Q \in \mathcal{P}_e$, let $Q \circ \tilde{B}^{-1}$ denote the probability measure on $(\Omega, \mathcal{B}(\Omega))$ induced by $\tilde{B}$ with respect to Q. We denote $\mathcal{P}_1 = \{Q \circ \tilde{B}^{-1} : Q \in \mathcal{P}_e\}$. By Lemma 6.2.4, we get

$$\tilde{\mathbb{E}}[|\tilde{B}_t - \tilde{B}_s|^4] = \tilde{\mathbb{E}}[|\bar{B}_t - \bar{B}_s|^4] = d|t-s|^2, \forall s, t \in [0, \infty).$$

Applying the well-known criterion for tightness of Kolmogorov–Chentsov's type expressed in terms of moments (see Appendix B), we conclude that $\mathcal{P}_1$ is tight. We denote by $\mathcal{P} = \overline{\mathcal{P}}_1$ the closure of $\mathcal{P}_1$ under the topology of weak convergence, then $\mathcal{P}$ is weakly compact.

Now, we give the representation of the G-expectation.

Theorem 6.2.5 *For each continuous monotone and sublinear function $G : \mathbb{S}(d) \mapsto \mathbb{R}$, let $\hat{\mathbb{E}}$ be the corresponding G-expectation on $(\Omega, Lip(\Omega))$. Then there exists a weakly compact family of probability measures $\mathcal{P}$ on $(\Omega, \mathcal{B}(\Omega))$ such that*

$$\hat{\mathbb{E}}[X] = \max_{P \in \mathcal{P}} E_P[X] \quad \textit{for } X \in Lip(\Omega).$$

Proof By Lemma 6.2.3 and Theorem 6.2.4, we have

$$\hat{\mathbb{E}}[X] = \max_{P \in \mathcal{P}_1} E_P[X] \quad \text{for } X \in Lip(\Omega).$$

For any $X \in Lip(\Omega)$, by Lemma 6.2.1, we get $\hat{\mathbb{E}}[|X - (X \wedge N) \vee (-N)|] \downarrow 0$ as $N \to \infty$. Noting that $\mathcal{P} = \overline{\mathcal{P}}_1$, by the definition of weak convergence, we arrive at the result. □

6.2.2 G-Expectation: A More Explicit Construction

In this subsection we will construct a family $\mathcal{P}$ of probability measures on Ω, for which the upper expectation coincides with the G-expectation $\mathbb{E}[\cdot]$ on $Lip(\Omega)$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(W_t)_{t\geq 0} = (W_t^i)_{i=1,t\geq 0}^d$ a standard d-dimensional Brownian motion under the classical (linear) probability P. The filtration generated by W is denoted by

$$\mathcal{F}_t := \sigma\{W_u, 0 \leq u \leq t\} \vee \mathcal{N}, \quad \mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0},$$

where $\mathcal{N}$ is the collection of all P–null subsets. We also denote, for a fixed $s \geq 0$,

$$\mathcal{F}_t^s := \sigma\{W_{s+u} - W_s, 0 \leq u \leq t\} \vee \mathcal{N}, \quad \mathbb{F}^s := \{\mathcal{F}_t^s\}_{t\geq 0}.$$

Let Θ be a given bounded, closed and convex subset in $\mathbb{R}^{d\times d}$. We denote by $\mathcal{A}_{t,T}^{\Theta}$, the collection of all Θ-valued $\mathbb{F}$-adapted process on an interval $[t, T] \subset [0, \infty)$. For any fixed $\theta \in \mathcal{A}_{t,T}^{\Theta}$ we denote

$$B_T^{t,\theta} := \int_t^T \theta_s dW_s.$$

We will show that, for each $n = 1, 2, \cdots, \varphi \in C_{b.Lip}(\mathbb{R}^{d\times n})$ and $0 \leq t_1, \cdots, t_n < \infty$, the G-expectation defined in [138, 141] can be equivalently defined as follows:

$$\mathbb{E}[\varphi(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})] = \sup_{\theta\in\mathcal{A}_{0,\infty}^{\Theta}} E_P[\varphi(B_{t_1}^{0,\theta}, B_{t_2}^{t_1,\theta}, \cdots, B_{t_n}^{t_{n-1},\theta})].$$

Given $\varphi \in C_{b.Lip}(\mathbb{R}^n \times \mathbb{R}^d)$, $0 \leq t \leq T < \infty$ and $\zeta \in L^2(\Omega, \mathcal{F}_t, P; \mathbb{R}^n)$, we define

$$\Lambda_{t,T}[\zeta] = \operatorname{ess}\sup_{\theta\in\mathcal{A}_{t,T}^{\Theta}} E_P[\varphi(\zeta, B_T^{t,\theta})|\mathcal{F}_t]. \tag{6.2.1}$$

Lemma 6.2.6 *For any θ_1 and θ_2 in $\mathcal{A}_{t,T}^{\Theta}$, there exists $\theta \in \mathcal{A}_{t,T}^{\Theta}$ such that*

$$E_P[\varphi(\zeta, B_T^{t,\theta})|\mathcal{F}_t] = E_P[\varphi(\zeta, B_T^{t,\theta^1})|\mathcal{F}_t] \vee E_P[\varphi(\zeta, B_T^{t,\theta^2})|\mathcal{F}_t]. \tag{6.2.2}$$

Consequently, there exists a sequence $\{\theta_i\}_{i=1}^{\infty}$ in the set $\mathcal{A}_{t,T}^{\Theta}$, such that

$$E_P[\varphi(\zeta, B_T^{t,\theta_i})|\mathcal{F}_t] \nearrow \Lambda_{t,T}[\zeta], \quad P\text{-a.s.}. \tag{6.2.3}$$

We also have, for each $s \leq t$,

$$E_P[\operatorname{ess}\sup_{\theta\in\mathcal{A}_{t,T}^{\Theta}} E_P[\varphi(\zeta, B_T^{t,\theta})|\mathcal{F}_t]|\mathcal{F}_s] = \operatorname{ess}\sup_{\theta\in\mathcal{A}_{t,T}^{\Theta}} E_P[\varphi(\zeta, B_T^{t,\theta})|\mathcal{F}_s]. \tag{6.2.4}$$

Proof We set $A = \left\{\omega : E_P[\varphi(\zeta, B_T^{t,\theta_1})|\mathcal{F}_t](\omega) \geq E_P[\varphi(\zeta, B_T^{t,\theta_2})|\mathcal{F}_t](\omega)\right\}$ and take $\theta_s = I_{[t,T]}(s)(I_A\theta_s^1 + I_{A^C}\theta_s^2)$. Since

$$\varphi(\zeta, B_T^{t,\theta}) = \mathbf{1}_A \varphi(\zeta, B_T^{t,\theta^1}) + \mathbf{1}_{A^c} \varphi(\zeta, B_T^{t,\theta^2}),$$

we derive (6.2.2) and then (6.2.3). Relation (6.2.4) follows from (6.2.2) and Yan's commutation theorem (cf [176] (in Chinese) and Theorem a3 in the Appendix of [134]). □

Lemma 6.2.7 *The mapping* $\Lambda_{t,T}[\cdot] : L^2(\Omega, \mathcal{F}_t, P; \mathbb{R}^n) \mapsto L^2(\Omega, \mathcal{F}_t, P; \mathbb{R})$ *has the following regularity properties which are valid for any* $\zeta, \zeta' \in L^2(\mathcal{F}_t)$*:*
(i) $\Lambda_{t,T}[\zeta] \le C_\varphi$.
(ii) $|\Lambda_{t,T}[\zeta] - \Lambda_{t,T}[\zeta']| \le k_\varphi |\zeta - \zeta'|$.
Here $C_\varphi = \sup_{(x,y)} \varphi(x, y)$ *and* k_φ *is the Lipschitz constant of* φ.

Proof We only need to prove (ii). We have

$$\begin{aligned}\Lambda_{t,T}[\zeta] - \Lambda_{t,T}[\zeta'] &\le \operatorname*{ess\,sup}_{\mathcal{A}^\Theta_{t,T}} E_P\left[\varphi(\zeta, \int_t^T \theta_s dW_s) - \varphi(\zeta', \int_t^T \theta_s dW_s) | \mathcal{F}_t\right] \\ &\le k_\varphi |\zeta - \zeta'|\end{aligned}$$

and, by symmetry, $\Lambda_{t,T}[\zeta'] - \Lambda_{t,T}[\zeta] \le k_\varphi |\zeta - \zeta'|$. Thus (ii) follows. □

Lemma 6.2.8 *For any* $x \in \mathbb{R}^n$, $\Lambda_{t,T}[x]$ *is a deterministic function. Moreover,*

$$\Lambda_{t,T}[x] = \Lambda_{0,T-t}[x]. \tag{6.2.5}$$

Proof Since the collection of processes $(\theta_s)_{s\in[t,T]}$ with

$$\left\{\theta_s = \sum_{j=1}^N I_{A_j} \theta_s^j : \{A_j\}_{j=1}^N \text{ is an } \mathcal{F}_t\text{–partition of } \Omega, \theta^j \in \mathcal{A}^\Theta_{t,T} \text{ is } (\mathbb{F}^t)\text{–adapted}\right\}$$

is dense in $\mathcal{A}^\Theta_{t,T}$, we can take a sequence $\theta_s^i = \sum_{j=1}^{N_i} I_{A_{ij}} \theta_s^{ij}$ of this type of processes such that $E_P[\varphi(x, B_T^{t,\theta^i}) | \mathcal{F}_t] \nearrow \Lambda_{t,T}[x]$. However,

$$\begin{aligned}E_P[\varphi(x, B_T^{t,\theta^i})|\mathcal{F}_t] &= \sum_{j=1}^{N_i} I_{A_{ij}} E_P[\varphi(x, B_T^{t,\theta^{ij}})|\mathcal{F}_t] = \sum_{j=1}^{N_i} I_{A_{ij}} E_P[\varphi(x, B_T^{t,\theta^{ij}})] \\ &\le \max_{1\le j\le N_i} E_P[\varphi(x, B_T^{t,\theta^{ij}})] = E_P[\varphi(x, B_T^{t,\theta^{ij_i}})],\end{aligned}$$

where, for each i, j_i is a maximizer of $\left\{E_P[\varphi(x, B_T^{t,\theta^{ij}})]\right\}_{j=1}^{N_i}$. This implies that

$$\lim_{i\to\infty} E_P[\varphi(x, B_T^{t,\theta^{ij_i}})] = \Lambda_{t,T}[x], \quad \text{a.s.}$$

Hence $\Lambda_{t,T}[x]$ is deterministic. The above reasoning shows that

$$\operatorname{ess}\sup_{\theta\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(x, B_T^{t,\theta})|\mathcal{F}_t] = \operatorname{ess}\sup_{\theta\in\mathcal{A}^{t,\Theta}_{0,T-t}} E_P[\varphi(x, \int_0^{T-t}\theta_s dW_s^t)],$$

where $W_s^t = W_{t+s} - W_t$, $s \geq 0$, and $\mathcal{A}^{t,\Theta}_{0,T-t}$ is the collection of Θ–valued and $\mathbb{F}^t$–adapted processes on $[0, T-t]$. This implies (6.2.5). □

We denote $u_{t,T}(x) := \Lambda_{t,T}[x]$, $t \leq T$. By Lemma 6.2.7, $u_{t,T}(\cdot)$ is a bounded and Lipschitz function.

Lemma 6.2.9 *For any $\zeta \in L^2(\Omega, \mathcal{F}_t, P; \mathbb{R}^n)$, we have*

$$u_{t,T}(\zeta) = \Lambda_{t,T}[\zeta], \quad a.s..$$

Proof By the above regularity properties of $\Lambda_{t,T}[\cdot]$, see Lemma 6.2.7, and $u_{t,T}(\cdot)$ we only need to deal with ζ which is a step function, i.e., $\zeta = \sum_{j=1}^N I_{A_j}x_j$, where $x_j \in \mathbb{R}^n$ and $\{A_j\}_{j=1}^N$ is an $\mathcal{F}_t$–partition of of Ω. For any x_j, let $\{\theta^{ij}\}_{i=1}^\infty$ in $\mathcal{A}^{\Theta}_{t,T}$ be $(\mathcal{F}_s^t)$–adapted processes such that

$$\lim_{i\to\infty} E_P[\varphi(x_j, B_T^{t,\theta^{ij}})|\mathcal{F}_t] = \lim_{i\to\infty} E_P[\varphi(x_j, B_T^{t,\theta^{ij}})] = \Lambda_{t,T}[x_j] = u_{t,T}(x_j), \;\; j = 1, \cdots, N.$$

Setting $\theta^i = \sum_{j=1}^N \theta^{ij} I_{A_j}$, we have

$$\begin{aligned}\Lambda_{t,T}[\zeta] &\geq E_P[\varphi(\zeta, B_T^{t,\theta^i})|\mathcal{F}_t] = E_P\left[\varphi(\sum_{j=1}^N I_{A_j}x_j, B_T^{t,\sum_{j=1}^N I_{A_j}\theta^{ij}})\Big|\mathcal{F}_t\right] \\ &= \sum_{j=1}^N I_{A_j} E_P[\varphi(x_j, B_T^{t,\theta^{ij}})|\mathcal{F}_t] \to \sum_{j=1}^N I_{A_j} u_{t,T}(x_j) = u_{t,T}(\zeta), \text{ as } i \to \infty.\end{aligned}$$

On the other hand, for any given $\theta \in \mathcal{A}^{\Theta}_{t,T}$, we have

$$\begin{aligned}E_P[\varphi(\zeta, B_T^{t,\theta})|\mathcal{F}_t] &= E_P\left[\varphi(\sum_{j=1}^N I_{A_j}x_j, B_T^{t,\theta})\Big|\mathcal{F}_t\right] \\ &= \sum_{j=1}^N I_{A_j} E_P[\varphi(x_j, B_T^{t,\theta})|\mathcal{F}_t] \\ &\leq \sum_{j=1}^N I_{A_j} u_{t,T}(x_j) = u_{t,T}(\zeta).\end{aligned}$$

We thus conclude that $\operatorname{ess\,sup}_{\theta\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(\zeta, B^{t,\theta}_T)|\mathcal{F}_t] \le u_{t,T}(\zeta)$. The proof is complete. □

We present now a result which generalizes the well-known dynamical programming principle:

Theorem 6.2.10 *For any $\varphi \in C_{b.Lip}(\mathbb{R}^n \times \mathbb{R}^{2d})$, $0 \le s \le t \le T$ and $\zeta \in L^2(\Omega, \mathcal{F}_s, P; \mathbb{R}^n)$ we have*

$$\operatorname{ess}\sup_{\theta\in\mathcal{A}^{\Theta}_{s,T}} E_P[\varphi(\zeta, B^{s,\theta}_t, B^{t,\theta}_T)|\mathcal{F}_s] = \operatorname{ess}\sup_{\theta\in\mathcal{A}^{\Theta}_{s,t}} E_P[\psi(\zeta, B^{s,\theta}_t)|\mathcal{F}_s], \tag{6.2.6}$$

where $\psi \in C_{b.Lip}(\mathbb{R}^n \times \mathbb{R}^d)$ is given as follows:

$$\psi(x, y) := ess \sup_{\bar{\theta}\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(x, y, B^{t,\bar{\theta}}_T)|\mathcal{F}_t] = \sup_{\bar{\theta}\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(x, y, B^{t,\bar{\theta}}_T)].$$

Proof It is clear that

$$\operatorname{ess}\sup_{\theta\in\mathcal{A}^{\Theta}_{s,T}} E_P[\varphi(\zeta, B^{s,\theta}_t, B^{t,\theta}_T)|\mathcal{F}_s] = \operatorname{ess}\sup_{\theta\in\mathcal{A}^{\Theta}_{s,t}} \left\{\operatorname{ess}\sup_{\bar{\theta}\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(\zeta, B^{s,\theta}_t, B^{t,\bar{\theta}}_T)|\mathcal{F}_s]\right\}.$$

Relation (6.2.4) and Lemma 6.2.9 imply that

$$\begin{aligned}\operatorname{ess}\sup_{\bar{\theta}\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(\zeta, B^{s,\theta}_t, B^{t,\bar{\theta}}_T)|\mathcal{F}_s] &= E_P[\operatorname{ess}\sup_{\bar{\theta}\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(\zeta, B^{s,\theta}_t, B^{t,\bar{\theta}}_T)|\mathcal{F}_t]|\mathcal{F}_s] \\ &= E_P[\psi(\zeta, B^{s,\theta}_t)|\mathcal{F}_s],\end{aligned}$$

We thus establish (6.2.6). □

For any given $\varphi \in C_{b.Lip}(\mathbb{R}^d)$ and $(t, x) \in [0, T] \times \mathbb{R}^d$, we set

$$v(t, x) := \sup_{\theta\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(x + B^{t,\theta}_T)].$$

Since for each $h \in [0, T - t]$,

$$\begin{aligned}v(t, x) &= \sup_{\theta\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(x + B^{t,\theta}_T)] \\ &= \sup_{\theta\in\mathcal{A}^{\Theta}_{t,T}} E_P[\varphi(x + B^{t,\theta}_{t+h} + B^{t+h,\theta}_T)] \\ &= \sup_{\theta\in\mathcal{A}^{\Theta}_{t,t+h}} E_P[v(t + h, x + B^{t,\theta}_{t+h})].\end{aligned}$$

This gives us the well-known dynamic programming principle:

Proposition 6.2.11 *The function $v(t,x)$ satisfies the following relation:*

$$v(t,x) = \sup_{\theta \in \mathcal{A}^{\Theta}_{t,t+h}} E_P[v(t+h, x + B^{t,\theta}_{t+h})]. \tag{6.2.7}$$

Lemma 6.2.12 *The function v is bounded by* $\sup|\varphi|$. *It is a Lipschitz function in x and $\frac{1}{2}$-Hölder function in t.*

Proof We only need to show the regularity in t. Note that

$$\sup_{\theta \in \mathcal{A}^{\Theta}_{t,t+h}} E_P[v(t+h, x + B^{t,\theta}_{t+h}) - v(t+h, x)] = v(t,x) - v(t+h,x).$$

Since v is a Lipschitz function in x, the absolute value of the left hand side is bounded by the quantity

$$C \sup_{\theta \in \mathcal{A}^{\Theta}_{t,t+h}} E_P[|B^{t,\theta}_{t+h}|] \leq C_1 h^{1/2}.$$

The $\frac{1}{2}$-Hölder of v in t is obtained. □

Theorem 6.2.13 *The function v is a viscosity solution of the G-heat equation:*

$$\frac{\partial v}{\partial t} + G(D^2 v) = 0, \quad on\ (t,x) \in [0,T) \times \mathbb{R}^d,$$
$$v(T,x) = \varphi(x),$$

where the function G is given by

$$G(A) = \frac{1}{2} \max_{\gamma \in \Theta} tr[A\gamma\gamma^T], \quad A \in \mathbb{R}^{d\times d}. \tag{6.2.8}$$

Proof Let $\psi \in C_b^{2,3}((0,T)\times\mathbb{R}^d)$ be such that $\psi \geq v$ and, for a fixed $(t,x) \in (0,T)\times\mathbb{R}^d$, $\psi(t,x) = v(t,x)$. From the dynamic programming principle (6.2.7) it follows that

$$\begin{aligned}
0 &= \sup_{\theta \in \mathcal{A}^{\Theta}_{t,t+h}} E_P[v(t+h, x + B^{t,\theta}_{t+h}) - v(t,x)] \\
&\leq \sup_{\theta \in \mathcal{A}^{\Theta}_{t,t+h}} E_P[\psi(t+h, x + B^{t,\theta}_{t+h}) - \psi(t,x)] \\
&= \sup_{\theta \in \mathcal{A}^{\Theta}_{t,t+h}} E_P\left[\int_t^{t+h} \left(\frac{\partial \psi}{\partial s} + \frac{1}{2}\mathrm{tr}[\theta_s\theta_s^T D^2\psi]\right)(s, x + B^{t,\theta}_s) ds\right].
\end{aligned}$$

Since $(\frac{\partial\psi}{\partial s} + \frac{1}{2}\mathrm{tr}[\theta_s\theta_s^T D^2\psi])(s,y)$ is uniformly Lipschitz in (s,y), we have, for a small $h > 0$, that

$$E_P[\left(\frac{\partial\psi}{\partial s}+\frac{1}{2}\mathrm{tr}[\theta_s\theta_s^T D^2\psi]\right)(s,x+\int_t^s\theta_r dW_r)]$$
$$\leq E_P[\left(\frac{\partial\psi}{\partial s}+\frac{1}{2}\mathrm{tr}[\theta_s\theta_s^T D^2\psi]\right)(t,x)]+Ch^{1/2}.$$

Therefore

$$\sup_{\theta\in\mathcal{A}^{\Theta}_{t,t+h}} E_P\int_t^{t+h}\left(\frac{\partial\psi}{\partial s}+\frac{1}{2}\mathrm{tr}[\theta_s\theta_s^T D^2\psi]\right)(t,x)ds+Ch^{3/2}\geq 0,$$

and since

$$\left(\frac{\partial\psi}{\partial s}+\frac{1}{2}\sup_{\gamma\in\Theta}\mathrm{tr}[\gamma\gamma^T D^2\psi]\right)(t,x)h+Ch^{3/2}\geq 0,$$

we conclude that $[\frac{\partial\psi}{\partial t}+G(D^2\psi)](t,x)\geq 0$. By definition, v is a viscosity subsolution. Similarly we can show that it is also a supersolution. □

We observe that $u(t,x):=v(T-t,x)$, thus u is the viscosity solution of the equation $\frac{\partial u}{\partial t}-G(D^2u)=0$, with Cauchy condition $u(0,x)=\varphi(x)$.

From the uniqueness of the viscosity solution of the G-heat equation and Theorem 6.2.10, we get immediately:

Proposition 6.2.14

$$\mathbb{E}[\varphi(B^0_{t_1},B^{t_1}_{t_2},\cdots,B^{t_{n-1}}_{t_n})]=\sup_{\theta\in\mathcal{A}^{\Theta}_{0,T}} E_P[\varphi(B^{0,\theta}_{t_1},B^{t_1,\theta}_{t_2},\cdots,B^{t_{n-1},\theta}_{t_n})]$$
$$=\sup_{\theta\in\mathcal{A}^{\Theta}_{0,T}} E_{P_\theta}[\varphi(B^0_{t_1},B^{t_1}_{t_2},\cdots,B^{t_{n-1}}_{t_n})],$$

where P_θ is the law of the process $B^{0,\theta}_t=\int_0^t\theta_s dW_s$, $t\geq 0$, for $\theta\in\mathcal{A}^{\Theta}_{0,\infty}$.

We are now going to show that the family $\{P_\theta,\ \theta\in\mathcal{A}^{\Theta}_{0,\infty}\}$ is tight. This property is important for our consideration in the next subsection.

Proposition 6.2.15 *The family of probability measures $\{P_\theta\colon\theta\in\mathcal{A}^{\Theta}_{0,\infty}\}$ on $C^d_0(\mathbb{R}^+)$ is tight.*

Proof Since $B^{s,\theta}_t=\int_s^t\theta_r dW_r$ and $\theta_r\in\Theta$, where Θ is a bounded subset in $\mathbb{R}^{d\times d}$, one can check that, for all $\theta\in\mathcal{A}^{\Theta}_{0,T}$, and $0\leq t\leq s\leq T$,

$$E_P[\left|B^{s,\theta}_t\right|^4]\leq C|s-t|^2,$$

with a constant C depending only on d and the bound $\sup\{|x|:x\in\Theta\}$. Thus

$$\sup_{\theta\in\mathcal{A}^{\Theta}_{0,T}} E_{P_\theta}[|B_s-B_t|^4]\leq C|s-t|^2.$$

We can now apply the well-known moment criterion for tightness of Kolmogorov–Chentsov's type to conclude that $\{P_\theta\colon \theta \in \mathcal{A}_{0,\infty}\}$ is tight. □

6.3 The Capacity of G-Brownian Motion

According to Theorem 6.2.5, we have obtained a weakly compact family of probability measures $\mathcal{P}$ on $(\Omega, \mathcal{B}(\Omega))$ to represent the G-expectation $\hat{\mathbb{E}}[\cdot]$. For this $\mathcal{P}$, we define two quantities, the associated G-**capacity**:

$$\hat{c}(A) := \sup_{P\in\mathcal{P}} P(A), \quad A \in \mathcal{B}(\Omega),$$

and the **upper expectation** for each $X \in L^0(\Omega)$

$$\bar{\mathbb{E}}[X] := \sup_{P\in\mathcal{P}} E_P[X].$$

By Theorem 6.2.5, we know that $\bar{\mathbb{E}} = \hat{\mathbb{E}}$ on $Lip(\Omega)$, thus the $\hat{\mathbb{E}}[|\cdot|]$-completion and the $\bar{\mathbb{E}}[|\cdot|]$-completion of $L_{ip}(\Omega)$ are the same.

For any $T > 0$, we also denote $\Omega_T = C_0([0,T];\mathbb{R}^d)$ equipped with the distance

$$\rho(\omega^1, \omega^2) = \left\|\omega^1 - \omega^2\right\|_{C_0^d([0,T])} := \max_{0\le t\le T} |\omega_t^1 - \omega_t^2|.$$

We now prove that $L_G^1(\Omega) = \mathbb{L}_c^1$, where $\mathbb{L}_c^1$ is defined in Sect. 6.1. First, we need the following classical approximation lemma.

Lemma 6.3.1 *For any $X \in C_b(\Omega)$ and $n = 1, 2, \cdots$, we denote*

$$X^{(n)}(\omega) := \inf_{\omega'\in\Omega}\{X(\omega') + n\left\|\omega - \omega'\right\|_{C_0^d([0,n])}\} \ \ for\ \omega \in \Omega.$$

Then the sequence $\{X^{(n)}\}_{n=1}^\infty$ satisfies:

(i) $-M \le X^{(n)} \le X^{(n+1)} \le \cdots \le X$, $M = \sup_{\omega\in\Omega}|X(\omega)|$;
(ii) $|X^{(n)}(\omega) - X^{(n)}(\omega')| \le n\left\|\omega - \omega'\right\|_{C_0^d([0,n])}, \ \ for\ \omega, \omega' \in \Omega$;
(iii) $X^{(n)}(\omega) \uparrow X(\omega)\ \ for\ \omega \in \Omega$.

Proof Claim (i) is obvious. For Claim (ii), we have the relation

$$\begin{aligned}
&X^{(n)}(\omega) - X^{(n)}(\omega') \\
&\le \sup\nolimits_{\bar{\omega}\in\Omega}\{[X(\bar{\omega}) + n\|\bar{\omega} - \omega\|_{C_0^d([0,n])}] - [X(\bar{\omega}) + n\left\|\bar{\omega} - \omega'\right\|_{C_0^d([0,n])}]\} \\
&\le n\left\|\omega - \omega'\right\|_{C_0^d([0,n])}.
\end{aligned}$$

By symmetry, $X^{(n)}(\omega) - X^{(n)}(\omega') \le n\left\|\omega - \omega'\right\|_{C_0^d([0,n])}$. Thus Claim (ii) follows.

We now prove Claim (iii). For any fixed $\omega \in \Omega$, let $\omega^{(n)} \in \Omega$ be such that

$$X(\omega^{(n)}) + n \left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} \le X^{(n)}(\omega) + \frac{1}{n}.$$

It is clear that $n \left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} \le 2M + 1$ or $\left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} \le \frac{2M+1}{n}$. Since $X \in C_b(\Omega)$, we get $X(\omega^{(n)}) \to X(\omega)$ as $n \to \infty$. We have also that

$$X(\omega) \ge X^{(n)}(\omega) \ge X(\omega^{(n)}) + n \left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} - \frac{1}{n},$$

thus

$$n \left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} \le |X(\omega) - X(\omega^{(n)})| + \frac{1}{n}.$$

We also have that

$$\begin{aligned} X(\omega^{(n)}) - X(\omega) + n \left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} &\ge X^{(n)}(\omega) - X(\omega) \\ &\ge X(\omega^{(n)}) - X(\omega) + n \left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} - \frac{1}{n}. \end{aligned}$$

From the above two relations, we obtain

$$\begin{aligned} |X^{(n)}(\omega) - X(\omega)| &\le |X(\omega^{(n)}) - X(\omega)| + n \left\|\omega - \omega^{(n)}\right\|_{C_0^d([0,n])} + \frac{1}{n} \\ &\le 2(|X(\omega^{(n)}) - X(\omega)| + \frac{1}{n}) \to 0 \text{ as } n \to \infty. \end{aligned}$$

Thus (iii) is obtained. □

Proposition 6.3.2 *For any* $X \in C_b(\Omega)$ *and* $\varepsilon > 0$, *there exists* $Y \in Lip(\Omega)$ *such that*

$$\bar{\mathbb{E}}[|Y - X|] \le \varepsilon. \tag{6.3.1}$$

Consequently, for any $p > 0$, *we have*

$$L_G^p(\Omega) = \mathbb{L}_c^p(\Omega). \tag{6.3.2}$$

Proof We denote $M = \sup_{\omega \in \Omega} |X(\omega)|$. By Theorem 6.1.16 and Lemma 6.3.1, we can find $\mu > 0$, $T > 0$ and $\bar{X} \in C_b(\Omega_T)$ such that $\bar{\mathbb{E}}[|X - \bar{X}|] < \varepsilon/3$, $\sup_{\omega \in \Omega} |\bar{X}(\omega)| \le M$ and

$$|\bar{X}(\omega) - \bar{X}(\omega')| \le \mu \left\|\omega - \omega'\right\|_{C_0^d([0,T])} \quad \textit{for} \quad \omega, \omega' \in \Omega.$$

Now for any positive integer n, we introduce a mapping $\theta_n(\omega) : \Omega \mapsto \Omega$:

$$\theta_n(\omega)(t) = \sum_{k=0}^{n-1} \frac{\mathbf{1}_{[t_k^n, t_{k+1}^n)}(t)}{t_{k+1}^n - t_k^n}[(t_{k+1}^n - t)\omega(t_k^n) + (t - t_k^n)\omega(t_{k+1}^n)] + \mathbf{1}_{[T,\infty)}(t)\omega(t),$$

where $t_k^n = \frac{kT}{n}$, $k = 0, 1, \cdots, n$. We set $\bar{X}^{(n)}(\omega) := \bar{X}(\theta_n(\omega))$, then

$$\begin{aligned}|\bar{X}^{(n)}(\omega) - \bar{X}^{(n)}(\omega')| &\le \mu \sup_{t\in[0,T]} |\theta_n(\omega)(t) - \theta_n(\omega')(t)| \\ &= \mu \sup_{k\in[0,\cdots,n]} |\omega(t_k^n) - \omega'(t_k^n)|.\end{aligned}$$

Let us choose a compact subset $K \subset \Omega$ such that $\bar{\mathbb{E}}[\mathbf{1}_{K^c}] \le \varepsilon/6M$. Since $\sup_{\omega\in K}\sup_{t\in[0,T]} |\omega(t) - \theta_n(\omega)(t)| \to 0$, as $n \to \infty$, we can choose a sufficiently large n_0 such that

$$\begin{aligned}\sup_{\omega\in K} |\bar{X}(\omega) - \bar{X}^{n_0}(\omega)| &= \sup_{\omega\in K} |\bar{X}(\omega) - \bar{X}(\theta_{n_0}(\omega))| \\ &\le \mu \sup_{\omega\in K}\sup_{t\in[0,T]} |\omega(t) - \theta_{n_0}(\omega)(t)| \\ &< \varepsilon/3.\end{aligned}$$

Setting $Y := \bar{X}^{(n_0)}$, it follows that

$$\begin{aligned}\bar{\mathbb{E}}[|X - Y|] &\le \bar{\mathbb{E}}[|X - \bar{X}|] + \bar{\mathbb{E}}[|\bar{X} - \bar{X}^{(n_0)}|] \\ &\le \bar{\mathbb{E}}[|X - \bar{X}|] + \bar{\mathbb{E}}[\mathbf{1}_K|\bar{X} - \bar{X}^{(n_0)}|] + 2M\bar{\mathbb{E}}[\mathbf{1}_{K^c}] \\ &< \varepsilon.\end{aligned}$$

We thus have (6.3.1) which implies (6.3.2). □

By Proposition 6.3.2, we also have a pathwise description of $L_G^p(\Omega)$ for any $p > 0$:

$$L_G^p(\Omega) = \{X \in L^0(\Omega) : X \text{ has a quasi-continuous version and } \lim_{n\to\infty} \bar{\mathbb{E}}[|X|^p I_{\{|X|>n\}}] = 0\}.$$

Furthermore, $\bar{\mathbb{E}}[X] = \hat{\mathbb{E}}[X]$, for any $X \in L_G^1(\Omega)$. Then we can extend the domain of the G-expectation $\hat{\mathbb{E}}$ from $L_G^1(\Omega)$ to the space of random variables $X \in L^0(\Omega)$ which makes the following definition meaningful:

$$\hat{\mathbb{E}}[X] := \bar{\mathbb{E}}[X].$$

Remark 6.3.3 This implies that, all equalities and inequalities established in Chaps. 3–5 which hold true in $L_G^p(\Omega)$ are still true in the sense of $\hat{c}$-quasi-surely.

In the next result, we are going to give some typical Borel measurable functions on Ω which are quasi-continuous.

Theorem 6.3.4 *Let X be a d-dimensional random vector in $L_G^1(\Omega)$. If A is a Borel set of $\mathbb{R}^d$ with $\hat{c}(\{X \in \partial A\}) = 0$, then $\mathbf{1}_{\{X \in A\}} \in L_G^1(\Omega)$.*

Proof For any $\varepsilon > 0$, since $X \in L_G^1(\Omega)$, we can find an open set $O \subset \Omega$ with $\hat{c}(O) \leq \varepsilon/2$ such that $X|_{O^c}$ is continuous. Set $D_i = \{x \in \mathbb{R}^d : d(x, \partial A) \leq 1/i\}$ and $A_i = \{x \in \mathbb{R}^d : d(x, \partial A) < 1/i\}$, it is easy to check that $\{X \in D_i\} \cap O^c$ is closed, $\{X \in A_i\} \subset \{X \in D_i\}$ and $\{X \in D_i\} \cap O^c \downarrow \{X \in \partial A\} \cap O^c$. Then we conclude that

$$\hat{c}(\{X \in D_i\} \cap O^c) \downarrow \hat{c}(\{X \in \partial A\} \cap O^c) = 0.$$

Thus we can find an i_0 such that $\hat{c}(\{X \in A_{i_0}\} \cap O^c) \leq \varepsilon/2$. Setting $O_1 = \{X \in A_{i_0}\} \cup O$, it is easy to verify that $\hat{c}(O_1) \leq \varepsilon$, $O_1^c = \{X \in A_{i_0}^c\} \cap O^c$ is closed and $\mathbf{1}_{\{X \in A\}}$ is continuous on O_1^c. Thus $\mathbf{1}_{\{X \in A\}}$ is quasi-continuous, which implies $\mathbf{1}_{\{X \in A\}} \in L_G^1(\Omega)$. □

Proposition 6.3.5 *Suppose G is non-degenerate, i.e., there exist a constant $\underline{\sigma}^2 > 0$ such that $G(A) - G(B) \geq \underline{\sigma}^2 \mathrm{tr}[A - B]$ for any $A \geq B$. Then it holds that $\mathbf{1}_{\{B_t \in [a,b]\}} \in L_G^1(\Omega_t)$ for any $t > 0$, where $a, b \in \mathbb{R}^d$.*

Proof By Exercise 6.5.8 of this chapter, we conclude that

$$\hat{c}\left(\{B_t \in \partial[a, b]\}\right) = 0,$$

which together with the Theorem 6.3.4 yields the desired result. □

The following example tells us that $L_G^p(\Omega)$ is strictly contained in $\mathbb{L}^p$.

Example 6.3.6 Let us consider an 1-dimensional non-degenerate G-Brownian motion $(B_t)_{t\geq 0}$, i.e.,

$$0 < -\hat{\mathbb{E}}[-B_1^2] = \underline{\sigma}^2 < \overline{\sigma}^2 = \hat{\mathbb{E}}[B_1^2],$$

and let $(\langle B \rangle_t)_{t \geq 0}$ be the quadratic process of B. We claim that, for any $\sigma^2 \in [\underline{\sigma}^2, \overline{\sigma}^2]$, the random variable

$$\xi = \mathbf{1}_{\{\langle B \rangle_1 - \sigma^2 = 0\}} \notin L_G^1(\Omega).$$

To show this, let us choose a sequence $\{\varphi_k\}_{k=1}^{\infty}$ of real continuous functions defined on $\mathbb{R}$ and taking values in $[0, 1]$ such that: $\varphi_k(v) = 0$, for $v \in \mathbb{R} \backslash [-2^{-k}, 2^{-k}]$, $\varphi_k(v) = 1$, for $v \in [-2^{-k-1}, 2^{-k-1}]$, and $\varphi_k \geq \varphi_{k+1}$. It is clear that $\xi_k := \varphi_k(|\langle B \rangle_1 - \sigma^2|) \in L_G^1(\Omega)$ and $\xi_k \downarrow \xi$. Hence, by Theorem 6.1.35, $\hat{\mathbb{E}}[\xi_i] \downarrow \hat{\mathbb{E}}[\xi]$. We can also check that

$$\hat{\mathbb{E}}[\xi_k - \xi_{k+1}] = \max_{v \in [\underline{\sigma}^2, \overline{\sigma}^2]} [\varphi_k(|v - \sigma^2|) - \varphi_{k+1}(|v - \sigma^2|)] = 1.$$

In view of Corollary 6.1.37 we conclude that $\xi \notin L_G^1(\Omega)$, i.e., ξ has no quasi-continuous version.

6.4 Quasi-continuous Processes

We have established in Sect. 6.3 that all random variables in $L_G^p(\Omega)$ are quasi continuous, in ω, with respect to the G-capacity $\hat{c}(\cdot)$. In this section, we are going to prove that, similarly, all stochastic processes in $M_G^p(0,T)$ are quasi continuous in (t,ω). We set $\mathcal{F}_t = \mathcal{B}(\Omega_t)$, $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ and use the following distance between two elements (t,ω) and (t',ω'):

$$\rho((t,\omega),(t',\omega')) = |t-t'| + \max_{s\in[0,T]} |\omega_s - \omega'_s|, \quad t,t' \in [0,T], \quad \omega,\omega' \in \Omega_T.$$

Recall that a process $(\eta_t)_{t\in[0,T]}$ is said to be *progressively measurable* if its restriction on $[0,t]\times\Omega$ is $\mathcal{B}([0,t])\otimes\mathcal{F}_t$-measurable for every t. We define, for $p \geq 1$,

$$\mathbb{M}^p(0,T) = \left\{\eta : \text{progressively measurable on } [0,T]\times\Omega_T \text{ and } \hat{\mathbb{E}}\left[\int_0^T |\eta_t|^p dt\right] < \infty\right\}$$

and the corresponding capacity

$$\bar{c}(A) = \frac{1}{T}\hat{\mathbb{E}}\left[\int_0^T \mathbf{1}_A(t,\omega)dt\right], \quad \text{for } A \in \mathcal{B}([0,T])\times\mathcal{F}_T.$$

Remark 6.4.1 Let A be a progressively measurable set in $[0,T]\times\Omega_T$. It is clear that $\mathbf{1}_A = 0$, $\bar{c}$-q.s. if and only if $\int_0^T \mathbf{1}_A(t,\cdot)dt = 0$ $\hat{c}$-q.s..

In what follows, we do not distinguish two progressively measurable processes η and η' if $\bar{c}(\{\eta \neq \eta'\}) = 0$. For any $p \geq 1$, $\mathbb{M}^p(0,T)$ is a Banach space under the norm $||\eta||_{\mathbb{M}^p} := \left(\hat{\mathbb{E}}\left[\int_0^T |\eta_t|^p dt\right]\right)^{1/p}$. Since $M_G^0(0,T) \subset \mathbb{M}^p(0,T)$ for any $p \geq 1$, $M_G^p(0,T)$ is a closed subspace of $\mathbb{M}^p(0,T)$. We also need to introduce the following space:

$$M_c(0,T) = \{\text{all } \mathbb{F}\text{-adapted processes } \eta \text{ in } C_b([0,T]\times\Omega_T)\}.$$

Proposition 6.4.2 *For any $p \geq 1$, the completion of the space $M_c(0,T)$ under the norm $||\eta||_{\mathbb{M}^p} := (\hat{\mathbb{E}}[\int_0^T |\eta_t|^p dt])^{1/p}$ is $M_G^p(0,T)$.*

Proof We first show that the completion of $M_c(0,T)$ under the norm $||\cdot||_{\mathbb{M}^p}$ belongs to $M_G^p(0,T)$. Indeed, for any fixed $\eta \in M_c(0,T)$, we set

$$\eta_t^{(k)}(\omega) = \sum_{i=0}^{k-1} \eta_{(iT)/k}(\omega)\mathbf{1}_{[\frac{iT}{k},\frac{(i+1)T}{k})}(t).$$

By the characterization of the space $L_G^p(\Omega)$ in Proposition 6.3.2, we get that $\eta_{\frac{iT}{k}} \in L_G^p(\Omega_{\frac{iT}{k}})$ and thus $\eta^{(k)} \in M_G^p(0,T)$. For each $\varepsilon > 0$, since $\mathcal{P}$ is weakly compact, there exists a compact set $K \subset \Omega_T$ such that $\hat{\mathbb{E}}[\mathbf{1}_{K^c}] \le \varepsilon$. Thus

$$\hat{\mathbb{E}}\left[\int_0^T |\eta_t - \eta_t^{(k)}|^p dt\right] \le \hat{\mathbb{E}}\left[\mathbf{1}_K \int_0^T |\eta_t - \eta_t^{(k)}|^p dt\right] + \hat{\mathbb{E}}\left[\mathbf{1}_{K^c} \int_0^T |\eta_t - \eta_t^{(k)}|^p dt\right]$$
$$\le \sup_{(t,\omega)\in[0,T]\times K} T|\eta_t(\omega) - \eta_t^{(k)}(\omega)|^p + (2l)^p T\varepsilon,$$

where l is the bound of η, i.e., $l = \sup_{(t,\omega)\in[0,T]\times\Omega} |\eta_t(\omega)|$. Note that $[0,T]\times K$ is compact and $\eta \in C_b([0,T]\times\Omega_T)$, then

$$\limsup_{k\to\infty} \hat{\mathbb{E}}\left[\int_0^T |\eta_t - \eta_t^{(k)}|^p dt\right] \le (2l)^p T\varepsilon.$$

Since ε can be arbitrarily small, we get $\|\eta^{(k)} - \eta\|_{\mathbb{M}^p} \to 0$ as $k \to \infty$. Thus $\eta \in M_G^p(0,T)$, which implies the desired result.

Now we show the converse part. For each given bounded process $\hat{\eta}_t = \sum_{i=1}^N \xi_i \mathbf{1}_{[t_{i-1},t_i)}(t)$ with $\xi_i \in Lip(\Omega_{t_{i-1}})$, one can find a sequence of functions $\{\phi_i^k\} \subset C([0,\infty)), i \le N, k \ge 1$ such that $\text{supp}(\phi_i^k) \subset (t_{i-1}, t_i)$ and $\int_0^T |\phi_i^k(t) - \mathbf{1}_{[t_{i-1},t_i)}(t)|^p dt \to 0$ as $k \to \infty$. Set $\hat{\eta}_t^{(k)} = \sum_{i=0}^{N-1} \xi_i \phi_i^k(t)$, it is easy to check that $\hat{\eta}^{(k)} \in M_c(0,T)$ and $||\hat{\eta}^{(k)} - \hat{\eta}||_{\mathbb{M}^p} \to 0$ as $k \to \infty$. Thus $M_G^p(0,T)$ belongs to the completion of $M_c(0,T)$ under the norm $||\cdot||_{\mathbb{M}^p}$, which completes the proof. □

Definition 6.4.3 A progressively measurable process $\eta : [0,T]\times\Omega_T \mapsto \mathbb{R}$ is called quasi-continuous (q.c.) if for any $\varepsilon > 0$, there exists a progressively measurable and open set O in $[0,T]\times\Omega_T$ such that $\bar{c}(O) < \varepsilon$ and η is continuous in O^c.

Definition 6.4.4 We say that a progressively measurable process $\eta : [0,T]\times\Omega_T \mapsto \mathbb{R}$ has a quasi-continuous version if there exists a quasi-continuous process η' such that $\bar{c}(\{\eta \ne \eta'\}) = 0$.

Theorem 6.4.5 *For any $p \ge 1$,*

$$M_G^p(0,T) = \left\{\eta \in \mathbb{M}^p(0,T) : \lim_{N\to\infty} \hat{\mathbb{E}}\left[\int_0^T |\eta_t|^p \mathbf{1}_{\{|\eta_t|\ge N\}} dt\right] = 0 \text{ and } \eta \text{ has a quasi-continuous version}\right\}.$$

Proof We denote

$$J_p = \left\{\eta \in \mathbb{M}^p(0,T) : \lim_{N\to\infty} \hat{\mathbb{E}}\left[\int_0^T |\eta_t|^p \mathbf{1}_{\{|\eta_t|\ge N\}} dt\right] = 0 \text{ and } \eta \text{ has a quasi-continuous version}\right\}.$$

Observe that the completion of $M_c(0,T)$ under the norm $||\cdot||_{\mathbb{M}^p}$ is $M_G^p(0,T)$. By a similar analysis as in Proposition 6.1.18, we can prove that $M_G^p(0,T)\subset J_p$.

It remains to show that $\eta\in J_p$ implies $\eta\in M_G^p(0,T)$. For each $n>0$, we introduce $\eta^{(n)}=(\eta\wedge n)\vee(-n)$, and easily see that

$$\hat{\mathbb{E}}\left[\int_0^T|\eta_t-\eta_t^{(n)}|^p dt\right]\leq 2\hat{\mathbb{E}}\left[\int_0^T(|\eta_t|^p-n^p/2)^+dt\right]\to 0,\quad as\quad n\to\infty.$$

Thus, we only need to prove that $\eta^{(n)}\in M_G^p(0,T)$ for each fixed $n>0$. For each $\varepsilon>0$, there exists a progressively measurable open set $O_\varepsilon\subset[0,T]\times\Omega_T$ with $\bar{c}(O_\varepsilon)<\varepsilon$ such that $\eta^{(n)}$ is continuous on O_ε^c.

We could directly apply Tietze's extention theorem (see Appendix A3) to extend the continuous function $\eta^{(n)}$ from O_ε^c to the whole domain of $[0,T]\times\Omega_T$. However, in order to keep the progressive measurability property, we need to justify the extension, step by step, as follows:

For any fixed $k=1,2,\cdots$, we set $t_i^k=iT/k$ for $i=0,\ldots,k$, and

$$F^{i,k}:=\{[0,t_{i-1}^k]\times\Omega_{t_i^k}\}\cup\{O_\varepsilon^c\cap\{(t_{i-1}^k,t_i^k]\times\Omega_{t_i^k}\},\ i=1,\ldots,k.$$

Since O_ε^c is progressively measurable and closed, we conclude that $F^{i,k}$ is a closed subset belonging to $\mathcal{B}([0,t_i^k])\otimes\mathcal{B}(\Omega_{t_i^k})$.

We apply Tietze's extension theorem and extend $\eta^{(n)}$ to $\bar{\eta}^{n,k}$ on $[0,t_1^k]\times\Omega_{t_1^k}$ such that $\bar{\eta}^{n,k}\in C_b([0,t_1^k]\times\Omega_{t_1^k})$, $\bar{\eta}^{n,k}=\eta^{(n)}$ on the set $O_\varepsilon^c\cap\{[0,t_1^k]\times\Omega_{t_1^k}\}$ and $|\bar{\eta}^{n,k}|\leq n$. It is clear that $\bar{\eta}_t^{n,k}$ is $\mathcal{F}_{t_1^k}$-measurable for each $t\in[0,t_1^k]$ since O_ε^c is progressively measurable. We then extend this $\bar{\eta}^{n,k}$ from $[0,t_1^k]\times\Omega_{t_1^k}$ to $F^{2,k}$ by setting $\bar{\eta}_t^{n,k}(\omega)=\eta_t^{(n)}(\omega)$ for $(t,\omega)\in O_\varepsilon^c\cap(t_1^k,t_2^k]\times\Omega_{t_2^k}$. We apply again Tietze's theorem and extend $\bar{\eta}^{n,k}$ from $F^{2,k}$ to $[0,t_2^k]\times\Omega_{t_2^k}$. Moreover, $\bar{\eta}_t^{n,k}$ is an $\mathcal{F}_{t_2^k}$-measurable continuous function bounded in $[-n,n]$ for each $t\leq t_2^k$. We repeat this procedure from $i=1$ to $i=k$ and thus obtain a function $\bar{\eta}^{n,k}\in C_b([0,T]\times\Omega)$ such that $\bar{\eta}^{n,k}=\eta^{(n)}$ on O_ε^c and $\bar{\eta}_t^{n,k}$ is $\mathcal{F}_{t_i^k}$-measurable for $t\in[t_{i-1}^k,t_i^k]$. Then we make a time shift on $\bar{\eta}^{n,k}$ to get the following progressively measurable process:

$$\hat{\eta}_t^{n,k}:=\bar{\eta}_{t-t_1^k}^{n,k}f_k(t),\quad t\in[0,T].$$

Here $f_k(t)$ is a continuous and non-decreasing function on $[0,T]$ with $f_k(t)=0$ for $t\leq t_1^k$ and $f_k(t)=1$ for $t\geq t_2^k$.

It is clear that $\hat{\eta}^{n,k}\in C_b([0,T]\times\Omega)$ and it converges to $\eta^{(n)}$ for each $(t,\omega)\in O_\varepsilon^c$. On the other hand, by Theorem 6.1.6 there exists a compact subset $K_\varepsilon\subset\Omega$ such that $\hat{\mathbb{E}}[\mathbf{1}_{K_\varepsilon^c}]\leq\varepsilon$. It follows, as $k\to\infty$, that $\hat{\eta}_t^{n,k}(\omega)$ converges to $\eta_t^{(n)}(\omega)$ uniformly on $(t,\omega)\in([0,T]\times K_\varepsilon)\cap O_\varepsilon^c$. Consequently

$$\begin{aligned}\hat{\mathbb{E}}\left[\int_0^T |\eta_t^{(n)} - \hat{\eta}_t^{n,k}|^p dt\right]^{1/p} &\le \hat{\mathbb{E}}\left[\int_0^T (\mathbf{1}_{([0,T]\times K_\varepsilon)\cap O_\varepsilon^c} + \mathbf{1}_{([0,T]\times K_\varepsilon)^c\cup O_\varepsilon})|\eta_t^{(n)} - \hat{\eta}_t^{n,k}|^p dt\right]^{1/p} \\ &\le \hat{\mathbb{E}}\left[\int_0^T \mathbf{1}_{([0,T]\times K_\varepsilon)\cap O_\varepsilon^c}|\eta_t^{(n)} - \hat{\eta}_t^{n,k}|^p dt\right]^{1/p} + 2n(T+1)^{1/p}\varepsilon^{1/p} \\ &\to 2n(T+1)^{1/p}\varepsilon^{1/p}, \quad \text{as } k \to \infty.\end{aligned}$$

Since ε can be arbitrarily small, it follows that $\eta^{(n)} \in M_G^p(0,T)$ and thus $\eta \in M_G^p(0,T)$. The proof is complete. □

The next result is a direct consequence of Theorem 6.4.5.

Corollary 6.4.6 *Let $\eta \in M_G^1(0,T)$ and $f \in C_b([0,T]\times\mathbb{R})$. Then $(f(t,\eta_t))_{t\le T} \in M_G^p(0,T)$ for any $p \ge 1$.*

Theorem 6.4.7 *Let $\eta^{(k)} \in M_G^1(0,T)$, $k \ge 1$, be such that $\eta^{(k)} \downarrow \eta$ $\bar{c}$-q.s., as $k \to \infty$. Then $\hat{\mathbb{E}}\left[\int_0^T \eta_t^{(k)} dt\right] \downarrow \hat{\mathbb{E}}\left[\int_0^T \eta_t dt\right]$. Moreover, if $\eta \in M_G^1(0,T)$, then $\hat{\mathbb{E}}\left[\int_0^T |\eta_t^{(k)} - \eta_t| dt\right] \downarrow 0$.*

Proof We choose $\eta^{k,N} \in M_G^0(0,T)$ such that $\hat{\mathbb{E}}\left[\int_0^T |\eta_t^{(k)} - \eta_t^{k,N}|dt\right] \to 0$ as $N \to \infty$. Since $\int_0^T \eta_t^{k,N} dt \in L_G^1(\Omega_T)$ and

$$\hat{\mathbb{E}}[|\int_0^T \eta_t^{k,N} dt - \int_0^T \eta_t^{(k)} dt|] \le \hat{\mathbb{E}}\left[\int_0^T |\eta_t^{(k)} - \eta_t^{k,N}|dt\right].$$

We conclude that $\int_0^T \eta_t^{(k)} dt \in L_G^1(\Omega_T)$ for $k \ge 1$. It then follows from Remark 6.4.1 and Theorem 6.4.5 that $\int_0^T \eta_t^{(k)} dt \downarrow \int_0^T \eta_t dt$ $\hat{c}$-q.s.. Therefore, by Proposition 6.1.31, we obtain that $\hat{\mathbb{E}}\left[\int_0^T \eta_t^{(k)} dt\right] \downarrow \hat{\mathbb{E}}\left[\int_0^T \eta_t dt\right]$. If $\eta \in M_G^1(0,T)$, then $|\eta^{(k)} - \eta| \in M_G^1(0,T)$ and $|\eta^{(k)} - \eta| \downarrow 0$ $\hat{c}$-q.s.. Thus $\hat{\mathbb{E}}\left[\int_0^T |\eta_t^{(k)} - \eta_t|dt\right] \downarrow 0$. □

Here is an example showing that $M_G^p(0,T)$ is strictly contained in $\mathbb{M}^p(0,T)$.

Example 6.4.8 We make the same assumptions as in Example 6.3.6. Then using similar arguments one can show that, for each $\sigma^2 \in [\underline{\sigma}^2, \overline{\sigma}^2]$, the process defined by $(\mathbf{1}_{\{\langle B\rangle_t - \sigma^2 t = 0\}})_{t\in[0,T]}$ is not in $M_G^1(0,T)$.

6.5 Exercises

Exercise 6.5.1 Let $\mathcal{P}$ be a family of weakly compact probabilities on $(\Omega, \mathcal{B}(\Omega))$. Suppose c is the upper probability and $\mathbb{E}$ is the upper expectation associated with $\mathcal{P}$. For any lower semicontinuous function $X \ge 0$ on Ω, we set

$$\mathbb{E}'[X] := \sup\{\mathbb{E}[Y] : \ Y \in C_b(\Omega) \text{ and } 0 \le Y \le X\}.$$

Then we define the outer capacity associated with $\mathcal{P}$ by

$$c'(A) := \inf\{\mathbb{E}'[X] : \ X \text{ is lower semicontinuous and } \mathbf{1}_A \le X\}, \ \forall A \in \mathcal{B}(\Omega).$$

Prove that

$$c'(A) = \inf\{c(O) : \ O \text{is an open set and } O \supset A\}.$$

Exercise 6.5.2 Let $\Omega = \mathbb{R}$ and $\mathcal{P} = \{\frac{1}{2}\delta_x + \frac{1}{2}\delta_{x+2} : x \in [-2, -1]\}$, where δ_x is unit mass measure at x. Suppose c is the upper probability associated with $\mathcal{P}$. Set $A = [-1, 1)$. Show that

(i) $\mathcal{P}$ is weakly compact.
(ii) $c(A) \neq \inf\{c(O) : \ O \text{ is an open set and } O \supset A\}$.

Exercise 6.5.3 Show that, for any $p > 0$,

$$L_G^p(\Omega_T) = \{X \in L^0(\Omega_T) : X \text{ has a quasi-continuous version and } \lim_{n\to\infty} \bar{\mathbb{E}}[|X|^p I_{\{|X|>n\}}] = 0\}.$$

Exercise 6.5.4 Prove that $\hat{\mathbb{E}}[e^{|B_t|}] < \infty$.

Exercise 6.5.5 Let $\mathbb{E}_{\hat{c}}$ be the Choquet expectation given by

$$\mathbb{E}_{\hat{c}}[X] := \int_0^{\infty} \hat{c}(X \ge t)dt + \int_{-\infty}^{0} (\hat{c}(X \ge t) - 1)dt.$$

Prove that $\mathbb{E}_{\hat{c}} = \hat{\mathbb{E}}$ if and only if $\hat{\mathbb{E}}$ is linear.

Exercise 6.5.6 Let η be in $M_G^2(0, T)$. Prove that $\int_0^t \eta_s dB_s$ has a quasi-modification whose paths are continuous.

Exercise 6.5.7 Let η be in $M_G^2(0, T)$. Prove that η is Itô-integrable for every $P \in \mathcal{P}$. Moreover,

$$\int_0^T \eta_s dB_s = \int_0^T \eta_s d_P B_s, \quad P\text{-a.s.},$$

where the right hand side is the classical Itô integral.

Exercise 6.5.8 Suppose that there exist two constants $0 < \underline{\sigma}^2 \le \overline{\sigma}^2 < \infty$ such that

$$\frac{1}{2}\underline{\sigma}^2 \text{tr}[A - B] \le G(A) - G(B) \le \frac{1}{2}\overline{\sigma}^2 \text{tr}[A - B], \text{ for } A \ge B.$$

Let $m \ge 0$, $\alpha = \underline{\sigma}^2(2\overline{\sigma}^2)^{-1}$ and $a \in \mathbb{R}$. Then for any $t \ge 0$, we have

$$\hat{\mathbb{E}}\left[\exp\left(-\frac{m|B_t - a|^2}{2\overline{\sigma}^2}\right)\right] \le (1 + mt)^{-\alpha}.$$

Hint: Show that

$$\tilde{u}_m(t,x) = (1+mt)^{-\alpha} \exp\left(-\frac{m|x-a|^2}{2\overline{\sigma}^2(1+mt)}\right)$$

is a viscosity supersolution of G-heat equation with initial condition $u_m(0,x) = \exp\left(-\frac{m|x-a|^2}{2\overline{\sigma}^2}\right)$.

Exercise 6.5.9 Let $(B_t)_{t\geq 0}$ be a non-degenerate G-Brownian motion. Prove that

(i) $\hat{\mathbb{E}}[-\mathbf{1}_{\{B_t\in[a,b]\}}] < 0$ for any $a < b$ and $t > 0$.
(ii) For any $\varphi, \phi \in C_{l.Lip}(\mathbb{R}^d)$ satisfying $\varphi \leq \phi$ and $\varphi(x) < \phi(x)$ for some x, one holds

$$\hat{\mathbb{E}}[\varphi(B_t)] < \hat{\mathbb{E}}[\phi(B_t)], \ \forall t > 0.$$

Exercise 6.5.10 Prove that the non-degenerate G-Brownian motion path is nowhere differentiable quasi-surely.

Notes and Comments

Choquet capacity was first introduced by Choquet [33], see also Dellacherie [42] and the references therein for more properties. The capacitability of Choquet capacity was first studied by Choquet [33] under 2-alternating case, see Dellacherie and Meyer [43], Huber and Strassen [89] and the references therein for more general case. It seems that the notion of upper expectations was first discussed by Huber [88] in robust statistics. It was rediscovered in mathematical finance, especially in risk measure, see Delbaen [44, 45], Föllmer and Schied [68] etc.

The fundamental framework of quasi-surely stochastic analysis in this chapter is due to Denis and Martini [48]. The results of Sects. 6.1–6.3 for G-Brownian motions were mainly obtained by Denis, Hu and Peng [47]. The upper probability in Sect. 6.1 was firstly introduced in [47] in the framework of G-expectation. Note that the results established in [47] cannot be obtained by using outer capacity introduced by Denis and Martini [48]. In fact the outer capacity may be strictly bigger than the inner capacity which coincides with the upper probability in Definition 6.1.3, see Exercise 6.5.2. An interesting open problem is to prove, or disprove, whether the outer capacity is equal to the upper probability associated with $\mathcal{P}$.

Hu and Peng [80] have introduced an intrinsic and simple approach. This approach can be regarded as a combination and extension of the construction approach of Brownian motion of Kolmogorov (for more general stochastic processes) and a sort of cylinder Lipschitz functions technique already introduced in Chap. 3. Section 6.1 is from [47] and Theorem 6.2.5 was firstly obtained in the same paper, whereas contents of Sects. 6.2 and 6.3 are mainly from [80].

Section 6.4 is mainly based on Hu et al. [85]. Some related discussions can be found in Song [159].

Part III
Stochastic Calculus for General Situations

Chapter 7
G-Martingale Representation Theorem

7.1 G-Martingale Representation Theorem

In Sect. 4.2 of Chap. 4, we presented a new type of G-martingale representation showing that a G-martingale can decomposed into a symmetric one and a decreasing one. Based on this idea, we now provide a complete and rigorous proof of this representation theorem for an L_G^2-martingale.

We consider a non-degenerate G-Brownian motion, i.e., we assume that there exists some constant $\underline{\sigma}^2 > 0$ such that $G(A) - G(B) \geq \underline{\sigma}^2 \text{tr}[A - B]$ for any $A \geq B$. For a fixed $T > 0$, let $S_G^0(0, T) = \{h(t, B_{t_1 \wedge t}, \cdots, B_{t_n \wedge t}) : t_1, \ldots, t_n \in [0, T], h \in C_{b.Lip}(\mathbb{R}^{1+n\times d})\}$. For $\eta \in S_G^0(0, T)$, set $\|\eta\|_{S_G^2} = \{\hat{\mathbb{E}}[\sup_{t\in[0,T]} |\eta_t|^2]\}^{1/p}$ and denote by $S_G^2(0, T)$ the completion of $S_G^0(0, T)$ under the norm $\|\cdot\|_{S_G^2}$. Then the following result holds:

Theorem 7.1.1 (G-Martingale Representation Theorem) *Let ξ be in $L_G^p(\Omega_T)$ for some $p > 2$. Then the martingale $\hat{\mathbb{E}}_t[\xi]$ has a continuous quasi-modification $Y \in S_G^2(0, T)$ given by*

$$Y_t = \hat{\mathbb{E}}[\xi] + \int_0^t Z_s dB_s + K_t, \tag{7.1.1}$$

where $Z \in M_G^2(0, T)$ and K is a non-increasing continuous G-martingale with $K_0 = 0$ and $K_T \in L_G^2(\Omega_T)$. Moreover, the above decomposition is unique.

In order to give the proof of Theorem 7.1.1, we need the following results. First, we state a prior estimate which will be used frequently in the sequel.

Lemma 7.1.2 (Prior estimates) *Let $(Y^{(i)}, Z^{(i)}, K^{(i)}), i = 1, 2$ satisfy equation (7.1.1) corresponding to data $\xi^{(i)}$. Then there exists a constant $C := C(T, \underline{\sigma}, G) > 0$ such that*

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_7

$$\|K_T^{(i)}\|_{L_G^2}^2 + \|Z^{(i)}\|_{M_G^2}^2 \le C\|Y^{(i)}\|_{S_G^2}^2, \quad i = 1, 2, \tag{7.1.2}$$

$$\|Z^{(1)} - Z^{(2)}\|_{M_G^2}^2 \le C\left(\|Y^{(1)} - Y^{(2)}\|_{S_G^2}^2 + \|Y^{(1)} - Y^{(2)}\|_{S_G^2}\left(\left\|Y^{(1)}\right\|_{S_G^2} + \left\|Y^{(2)}\right\|_{S_G^2}\right)\right). \tag{7.1.3}$$

Proof Without loss of generality, we assume that $d = 1$. Set $\hat{Y} := Y^{(1)} - Y^{(2)}$, $\hat{Z} := Z^{(1)} - Z^{(2)}$ and $\hat{K} := K^{(1)} - K^{(2)}$. Then it is easy to check that

$$\hat{Y}_t = \hat{\xi} - \int_t^T \hat{Z}_s dB_s - (\hat{K}_T - \hat{K}_t),\ \forall t \in [0, T].$$

Fix $P \in \mathcal{P}$. Then $\int_0^t \hat{Z}_s dB_s$ is a P-martingale (see Exercise 6.5.7 in Chap. 6). Applying the classical Itô's formula to $\hat{Y}_t$ yields that

$$|\hat{Y}_t|^2 + \int_t^T |\hat{Z}_s|^2 d\langle B\rangle_s = |\hat{\xi}|^2 - 2\int_t^T \hat{Y}_s \hat{Z}_s dB_s - 2\int_t^T \hat{Y}_s d\hat{K}_s,\ P\text{-a.s.} \tag{7.1.4}$$

Since $Y^{(i)} \in S_G^2(0, T)$, we have that

$$E_P[\sup_{s\in[0,T]} |\hat{Y}_s|^2] \le 2(\|Y^{(1)}\|_{S_G^2}^2 + \|Y^{(2)}\|_{S_G^2}^2) < \infty$$

which together with

$$E_P\left[\int_0^T |\hat{Z}|_s^2 ds\right] \le 2(\|Z^{(1)}\|_{M_G^2}^2 + \|Z^{(2)}\|_{M_G^2}^2) < \infty$$

indicates that $\int_0^t \hat{Y}_s \hat{Z}_s dB_s$ is a P-martingale. Note that $\langle B\rangle_t \ge \underline{\sigma}^2 t$. Taking P-expectation to Eq. (7.1.4), we get that

$$\underline{\sigma}^2 E_P\left[\int_0^T |\hat{Z}_s|^2 ds\right] \le E_P[|\hat{\xi}|^2 + \sup_{s\in[0,T]} |\hat{Y}_s| \cdot (|K_T^{(1)}| + |K_T^{(2)}|)]. \tag{7.1.5}$$

Since $Y^{(i)} \in S_G^2(0, T)$, $Z^{(i)} \in M_G^2(0, T)$ and $K_T^{(i)} \in L_G^2(\Omega_T)$, it is easy to check that

$$\int_0^T |\hat{Z}_s|^2 ds \quad \text{and} \quad \sup_{s\in[0,T]} |\hat{Y}_s| \cdot (K_T^{(1)} + K_T^{(2)}) \ \text{ are in } \ L_G^1(\Omega_T).$$

By taking $\sup_{P\in\mathcal{P}} E_P[\cdot] = \hat{\mathbb{E}}[\cdot]$ on both sides of (7.1.5), it follows that

$$\underline{\sigma}^2\hat{\mathbb{E}}\left[\int_0^T |\hat{Z}_s|^2 ds\right] \leq \hat{\mathbb{E}}\left[|\hat{\xi}|^2 + \sup_{s\in[0,T]} |\hat{Y}_s| \cdot (|K_T^{(1)}| + |K_T^{(2)}|)\right]$$
$$\leq \|\hat{Y}\|_{S_G^2}^2 + \|\hat{Y}\|_{S_G^2}(\|K_T^{(1)}\|_{L_G^2} + \|K_T^{(2)}\|_{L_G^2}). \tag{7.1.6}$$

Let us consider two special cases: Case 1: $\xi^{(1)} = 0$, thus $(Y^{(1)}, Z^{(1)}, K^{(1)}) \equiv 0$; Case 2: $\xi^{(2)} = 0$, thus $(Y^{(2)}, Z^{(2)}, K^{(2)}) \equiv 0$. By separately applying the inequality (7.1.6) to these two cases, we obtain

$$\underline{\sigma}^2\hat{\mathbb{E}}\left[\int_0^T |Z_s^{(i)}|^2 ds\right] \leq \|Y^{(i)}\|_{S_G^2}^2 + \|Y^{(i)}\|_{S_G^2}\|K_T^{(i)}\|_{L_G^2}, \quad \text{for } i = 1, 2. \tag{7.1.7}$$

On the other hand, notice that $K_T^{(i)} = \xi^{(i)} - \hat{\mathbb{E}}\left[\xi^{(i)}\right] - \int_0^T Z_s^{(i)} dB_s$. Then there exists some constant C_1 depending on G such that

$$\|K_T^{(i)}\|_{L_G^2}^2 \leq C_1(\|Y^{(i)}\|_{S_G^2}^2 + \|Z^{(i)}\|_{M_G^2}^2).$$

Combining this inequality with the right hand side of (7.1.7), we can find a constant C_2 depending on $\underline{\sigma}^2$ such that

$$\underline{\sigma}^2\|Z^{(i)}\|_{M_G^2}^2 \leq C_2\|Y^{(i)}\|_{S_G^2}^2 + \frac{1}{2}\underline{\sigma}^2\|Z^{(i)}\|_{M_G^2}^2.$$

From here and the inequality (7.1.6) we obtain the estimate (7.1.2), and then (7.1.3). The proof is complete. □

Next we will study the G-martingale representation theorem for cylinder functions.

Lemma 7.1.3 *For any $\xi \in Lip(\Omega_T)$ of the form*

$$\xi = \varphi(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_k} - B_{t_{k-1}}), \quad \varphi \in C_{b.Lip}(\mathbb{R}^{k\times d}), \quad 0 \leq t_1 < \cdots < t_k \leq T, \tag{7.1.8}$$

Eq. (7.1.1) *holds true.*

Proof For simplicity we just treat the case $d = 1$ and leave the case for multi-dimensional G-Brownian motion as an exercise. The proof is divided into the following three steps.

Step 1. We start with a simple case $\xi = \varphi(B_T - B_{t_1})$ for a fixed $t_1 \in (0, T]$ with $\varphi \in C_{b.Lip}(\mathbb{R})$. It is clear that the martingale $\hat{\mathbb{E}}_t[\xi]$ is $Y_t = \hat{\mathbb{E}}_t[\varphi(B_T - B_{t_1})] = u(t, B_t - B_{t_1})$, $t \in [t_1, T]$, where u is the viscosity solution of the G-heat equation

$$\partial_t u(t,x) + G(\partial_{xx}^2 u(t,x)) = 0, \quad u(x,T) = \varphi(x). \tag{7.1.9}$$

By the interior regularity of u (see Appendix C, Theorem C.4.4), for any $\kappa > 0$, there exists a constant $\alpha \in (0, 1)$ such that,

$$||u||_{C^{1+\alpha/2,2+\alpha}([0,T-\kappa]\times\mathbb{R})} < \infty.$$

Applying Itô's formula to $u(t, B_t - B_{t_1})$ on $[t_1, T-\kappa]$, we get

$$\begin{aligned} u(t, B_t - B_{t_1}) =& u(T-\kappa, B_{T-\kappa} - B_{t_1}) \\ & - \int_t^{T-\kappa} \partial_x u(s, B_s - B_{t_1}) dB_s - (K_{T-\kappa} - K_t), \end{aligned} \tag{7.1.10}$$

where $K_t = \frac{1}{2}\int_{t_1}^t \partial_{xx}^2 u(s, B_s - B_{t_1}) d\langle B\rangle_s - \int_{t_1}^t G(\partial_{xx}^2 u(s, B_s - B_{t_1}))ds$ is a non-increasing continuous G-martingale for $t \in [t_1, T-\kappa]$. In fact, we have

$$u(t,x) = \hat{\mathbb{E}}[\varphi(x + B_T - B_t)].$$

Since φ is a Lipschitz function, thus u is also a Lipschitz function with the same Lipschitz constant as φ. Moreover,

$$\begin{aligned} u(t,x) &= \hat{\mathbb{E}}[\varphi(x + B_T - B_t)] = \hat{\mathbb{E}}[\varphi(x + (B_T - B_{t+\delta}) + B_{t+\delta} - B_t)] \\ &= \hat{\mathbb{E}}[u(t+\delta, x + B_\delta)]. \end{aligned}$$

Hence $|u(t,x) - u(t+\delta,x)| \le \hat{\mathbb{E}}[|u(t+\delta, x+B_\delta) - u(t+\delta,x)|] \le C\hat{\mathbb{E}}[|B_\delta|] \le C|\sqrt{\delta}|$. The uniformly Lipschitz continuity of u in x gives us that $|\partial_x u(t,x)| \le L$. Therefore

$$\begin{aligned} & \hat{\mathbb{E}}\left[|Y_{T-\kappa} - Y_T|^2 + \int_{T-\kappa}^T |Z_t|^2 dt\right] \\ &= \hat{\mathbb{E}}\left[|u(T-\kappa, B_{T-\kappa} - B_1) - u(T, B_T - B_1)|^2 + \int_{T-\kappa}^T |\partial_x u(B_t - B_{t_1}, t)|^2 dt\right] \\ &\le C\left(\sqrt{\kappa} + \hat{\mathbb{E}}\left[|B_\kappa|^2 + \int_{T-\kappa}^T |L|^2 dt\right]\right) \to 0, \text{ as } \kappa \to 0, \end{aligned}$$

where $Y_t = u(t, B_t - B_{t_1})$ and $Z_t = \partial_x u(t, B_t - B_{t_1})$. Then by Eq. (7.1.10), we can find a random variable $K_T \in L_G^2(\Omega_T)$ so that $\hat{\mathbb{E}}[(K_{T-\kappa} - K_T)^2] \to 0$ as $\kappa \downarrow 0$. Therefore

$$Y_t = u(t, B_t - B_{t_1})\mathbf{1}_{[t_1,T]}(t) + u(t_1, 0)\mathbf{1}_{[0,t_1)}(t)$$

is a continuous martingale with

$$Z_t = \partial_x u(t, B_t - B_{t_1})\mathbf{1}_{[t_1,T]}(t), \quad Y_t = Y_0 + \int_0^t Z_s dB_s + K_t.$$

Furthermore,

$$\hat{\mathbb{E}}[K_T^2] \leq C\left(|Y_0|^2 + \hat{\mathbb{E}}[|Y_T|^2] + \hat{\mathbb{E}}\left[\int_0^T |Z_s|^2 ds\right]\right) \leq C_1.$$

where C_1 depends only on the bound and the Lipschitz constant of φ. It follows that $Y_t \in S_G^2(0,T)$, for any $t \in [0,T]$, and $Z \in M_G^2(t_1,T)$ and $K_T \in L_G^2(\Omega_T)$.

Step 2. We now consider the case $\xi = \varphi_1(B_T - B_{t_1}, B_{t_1})$ with $\varphi_1 \in C_{b.Lip}(\mathbb{R}^2)$.

For any fixed $y \in \mathbb{R}$, let $u(\cdot,\cdot,y)$ be the solution of the PDE (7.1.9) with terminal condition $\varphi_1(\cdot, y)$. Then it is easy to check that

$$Y_t := \hat{\mathbb{E}}_t[\xi] = u(t, B_t - B_{t_1}, B_{t_1}),\ \forall t \in [t_1, T]$$

By Step 1, we have

$$\begin{aligned} u(t, B_t - B_{t_1}, y) =& u(T, B_T - B_{t_1}, y) \\ & - \int_t^T \partial_x u(s, B_s - B_{t_1}, y) dB_s - (K_T^y - K_t^y). \end{aligned} \tag{7.1.11}$$

For any given $n \in \mathbb{N}$, we take

$$\tilde{h}_i^{(n)}(x) = \mathbf{1}_{[-n+\frac{i}{n}, -n+\frac{i+1}{n})}(x),\ i = 0, \ldots, 2n^2 - 1,$$

and $\tilde{h}_{2n^2}^{(n)} = 1 - \sum_{i=0}^{2n^2-1} h_i^{(n)}$. Through Eq. (7.1.11), we get

$$\tilde{Y}_t^{(n)} = \tilde{Y}_T^{(n)} - \int_t^T \tilde{Z}_s^{(n)} dB_s - (\tilde{K}_T^{(n)} - \tilde{K}_t^{(n)}),$$

where $\tilde{Y}_t^{(n)} = \sum_{i=0}^{2n^2} u(t, B_t - B_{t_1}, -n + i/n)\tilde{h}_i^{(n)}(B_{t_1})$, $\tilde{Z}_t^{(n)} = \sum_{i=0}^{2n^2} \partial_y u(t, B_t - B_{t_1}, -n + i/n)\tilde{h}_i^{(n)}(B_{t_1})$ and $\tilde{K}_t^{(n)} = \sum_{i=0}^{2n^2} K_t^{-n+i/n}\tilde{h}_i^{(n)}(B_{t_1})$. By Proposition 6.3.5 of Chap. 6, $\tilde{h}_i^{(n)}(B_{t_1}) \in L_G^2(\Omega_{t_1})$, we obtain that, for any t, $(\tilde{Y}_t^{(n)}, \tilde{K}_t^{(n)}) \in L_G^2(\Omega_t)$, $\tilde{Z}^{(n)} \in M_G^2(t_1, T)$ and, for any n,

$$\hat{\mathbb{E}}_s[\tilde{K}_r^{(n)}] = \sum_{i=0}^{2n^2} \tilde{h}_i^{(n)}(B_{t_1})\hat{\mathbb{E}}_s[K_r^{-n+i/n}] = \tilde{K}_s^{(n)}, \quad t_1 \leq s \leq r \leq T.$$

On the other hand, for any $t, s \in [0,T]$, $x, \bar{x}, y, \bar{y} \in \mathbb{R}$,

$$\begin{aligned} |u(t,x,y) - u(t,\bar{x},\bar{y})| &\leq |\hat{\mathbb{E}}[\varphi_1(x + B_T - B_t, y)] - \hat{\mathbb{E}}[\varphi_1(\bar{x} + B_T - B_t, \bar{y})]| \\ &\leq C(|x - \bar{x}| + |y - \bar{y}|). \end{aligned}$$

Through the relation $u(t, x, y) = \hat{\mathbb{E}}[u(t+\delta, x+B_\delta, y)]$ we can also get $|u(t, x, y) - u(s, x, y)| \le C\sqrt{|t-s|}$. Thus we deduce that

$$|Y_t - \tilde{Y}_t^{(n)}| \le \sum_{i=1}^{2n^2-1} h_i^{(n)}(B_{t_1})|u(t, -n + \frac{i}{n}, B_t - B_{t_1}) - u(t, B_{t_1}, B_t - B_{t_1})|$$

$$+ (|\tilde{Y}_t^{(n)}| + |Y_t|)\mathbf{1}_{\{|B_{t_1}|>n\}} \le \frac{C}{n} + \frac{2||u||_\infty}{n}|B_{t_1}|,$$

from which we derive that

$$\|Y_t - \tilde{Y}_T^{(n)}\|_{S_G^2}^2 \to 0, \text{ as } n \to \infty.$$

Consequently, by the prior estimate Lemma 7.1.2, there exists a process $Z \in M_G^2(t_1, T)$ so that

$$\lim_{n\to\infty} \|Z^{(n)} - Z\|_{M_G^2} = 0.$$

Denote

$$K_t := Y_t - Y_{t_1} - \int_{t_1}^t Z_s dB_s, \ \forall t \in [t_1, T].$$

Note that for any $t \in [t_1, T]$,

$$(\tilde{K}_t^{(n)})_{n=1}^\infty = (\tilde{Y}_t^{(n)} - \tilde{Y}_{t_1}^{(n)} - \int_{t_1}^t \tilde{Z}_s^{(n)} dB_s)_{n=1}^\infty$$

converges to K_t in $L_G^2(\Omega_t)$. Thus K_t is a non-increasing continuous process such that $\hat{\mathbb{E}}_s[K_t] = K_s$.

For $Y_{t_1} = u(t_1, 0, B_{t_1})$, we can use the same method as in Step 1, now on $[0, t_1]$.

Step 3. The more general case can be iterated similarly as in Steps 2. The proof is complete. □

Finally, we present a generalized Doob's maximal inequalities of G-martingales, which turns to be a very useful tool.

Lemma 7.1.4 *For any $\xi \in Lip(\Omega_T)$ satisfying the condition in Lemma 7.1.3 and for any $1 < p < \bar{p}$ with $p \le 2$, there exist a constant ρ, a process $Z' \in H_G^p(0, T)$, a non-positive random variable $K'_T \in L_G^p(\Omega_T)$ and a constant $C_p > 1$ depending on p such that:*

(i) $\xi = \rho + \int_0^T Z'_t dB_t + K'_T$,
(ii) $\hat{\mathbb{E}}[|K'_T|^p]^{1/p} \le \frac{C_p \bar{p}}{\bar{p}-p} \cdot \left(\hat{\mathbb{E}}[|\xi|]^{1/p} + \hat{\mathbb{E}}[|\xi|^{\bar{p}}]^{1/p}\right)$.

Proof Without loss of generality, assume $d = 1$.

We first consider a special case that $|\xi| \le 1$. By Lemma 7.1.3, we have the following representation:

$$\hat{\mathbb{E}}_t[\xi] = Y_t = \hat{\mathbb{E}}[\xi] + \int_0^t Z_s dB_s + K_t,\ q.s., \tag{7.1.12}$$

where $Y \in S_G^2(0,T)$, $Z \in H_G^2(0,T)$ and K is a non-increasing continuous G-martingale with $K_0 = 0$ and $K_T \in L_G^2(\Omega_T)$. Thus, applying Itô's formula yields that

$$\xi^2 = (\hat{\mathbb{E}}[\xi])^2 + 2\int_0^T Y_t Z_t dB_t + 2\int_0^T Y_t dK_t + \int_0^T |Z_t|^2 d\langle B\rangle_t \tag{7.1.13}$$

Note that $|Y_t| \le 1$ for any $t \in [0,T]$. Then taking G-expectation on both sides of (7.1.13) leads to the inequality

$$\hat{\mathbb{E}}\left[(\int_0^T Z_s dB_s)^2\right] = \hat{\mathbb{E}}\left[\int_0^T |Z_t|^2 d\langle B\rangle_t\right] \le \hat{\mathbb{E}}[|\xi|] + 2\hat{\mathbb{E}}[-K_T].$$

On the other hand, setting $t = T$ in (7.1.12) and taking G-expectation implies that

$$\hat{\mathbb{E}}[-K_T] = \hat{\mathbb{E}}[\xi] + \hat{\mathbb{E}}[-\xi] \le 2\hat{\mathbb{E}}[|\xi|].$$

Therefore, we obtain that $\hat{\mathbb{E}}\left[(\int_0^T Z_s^{(n)} dB_s)^2\right] \le 5\hat{\mathbb{E}}[|\xi|]$ and, once again from (7.1.12),

$$\hat{\mathbb{E}}[(K_T)^2] \le 2\hat{\mathbb{E}}[\xi^2] + 2\hat{\mathbb{E}}\left[\left(\hat{\mathbb{E}}[\xi] + \int_0^T Z_t dB_t\right)^2\right] \le 14\hat{\mathbb{E}}[|\xi|].$$

Consequently, for $1 < p < \bar{p}$ and $p \le 2$, we obtain that

$$\hat{\mathbb{E}}[(-K_T)^p] \le \hat{\mathbb{E}}[-K_T] + \hat{\mathbb{E}}[(K_T)^2],$$

from which it follows that

$$\hat{\mathbb{E}}[(-K_T)^p] \le 16\hat{\mathbb{E}}[|\xi|]. \tag{7.1.14}$$

Now we consider a more general case when $|\xi|$ is bounded by a positive integer M. We set $\xi^{(n)} = (\xi \wedge n) \vee (-n)$ and $\eta^{(n)} = \xi^{(n+1)} - \xi^{(n)}$ for any integer $n = 0, 1, \cdots, M$. Then by Lemma 7.1.3, for any n, we have the following representation:

$$\hat{\mathbb{E}}_t[\eta^{(n)}] = Y_t^{(n)} = \hat{\mathbb{E}}[\eta^{(n)}] + \int_0^t Z_s^{(n)} dB_s + K_t^{(n)}, \quad q.s.. \tag{7.1.15}$$

Also, from (7.1.14), we have

$$\hat{\mathbb{E}}[(-K_T^{(n)})^p] \le 16\hat{\mathbb{E}}[|\eta^n|].$$

Note that $|\eta^{(n)}| \leq \mathbf{1}_{\{|\xi|\geq n\}}$. Thus by Minkowski's inequality, we can derive that

$$\begin{aligned}\hat{\mathbb{E}}[(-\sum_{n=0}^{M-1} K_T^{(n)})^p]^{1/p} \leq & \hat{\mathbb{E}}[(-K_T^0)^p]^{1/p} + \sum_{n=1}^{M-1} \hat{\mathbb{E}}[(-K_T^{(n)})^p]^{1/p} \\ \leq & 2^{4/p}\big(\hat{\mathbb{E}}[|\xi|]^{1/p} + \sum_{n=1}^{M-1} |\hat{c}(|\xi| \geq n)|^{1/p}\big) \\ \leq & 2^{4/p}\big(\hat{\mathbb{E}}[|\xi|]^{1/p} + \sum_{n=1}^{M-1} n^{-\bar{p}/p}\hat{\mathbb{E}}[|\xi|^{\bar{p}}]^{1/p}\big).\end{aligned}$$

Here we have used the Markov inequality (see Lemma 6.1.17 of Chap. 6) in the last step. Since $\sum_{n=1}^{M-1} n^{-\bar{p}/p} \leq \bar{p}/(\bar{p}-p)$, we deduce that

$$\hat{\mathbb{E}}\left[(-\sum_{n=0}^{M-1} K_T^{(n)})^p\right]^{1/p} \leq \frac{2^{4/p}\bar{p}}{\bar{p}-p}\left(\hat{\mathbb{E}}[|\xi|]^{1/p} + \hat{\mathbb{E}}[|\xi|^{\bar{p}}]^{1/p}\right).$$

Note that $\xi = \sum_{n=0}^{M-1} \eta^{(n)}$. Let $K_T' := \sum_{n=0}^{M-1} K_T^{(n)}$, $Z_t' := \sum_{n=0}^{M-1} Z_t^{(n)}$ and $\rho := \sum_{n=0}^{M-1} \hat{\mathbb{E}}[\eta^{(n)}]$. Then we get the desired result. □

Based on the above estimates, we can introduce the following inequality of Doob's type for G-martingales.

Lemma 7.1.5 *Suppose $\alpha \geq 1$ and $\delta > 0$. Then for each $1 < p < \bar{p} := (\alpha+\delta)/\alpha$ with $p \leq 2$ and for all $\xi \in Lip(\Omega_T)$ satisfying the condition in Lemma 7.1.3, there is a constant $C_p > 1$ depending on p such that*

$$\hat{\mathbb{E}}\left[\sup_{t\in[0,T]} \hat{\mathbb{E}}_t[|\xi|^\alpha]\right] \leq \frac{C_p\bar{p}q}{\bar{p}-p}\left((\hat{\mathbb{E}}[|\xi|^{\alpha+\delta}])^{1/(\bar{p}p)} + (\hat{\mathbb{E}}[|\xi|^{\alpha+\delta}])^{1/p}\right), \qquad (7.1.16)$$

Proof For any $1 < p < \bar{p}$ with $p \leq 2$, applying Lemma 7.1.4 to $|\xi|^\alpha \in Lip(\Omega_T)$, we can find a constant ρ, a process $Z' \in H_G^p(0,T)$, a non-positive random variable $K_T' \in L_G^2(\Omega_T)$ such that $|\xi|^\alpha = \rho + \int_0^T Z_t' dB_t + K_T'$. Then we have the relation

$$\sup_{t\in[0,T]} \hat{\mathbb{E}}_t[|\xi|^\alpha] \leq \sup_{t\in[0,T]}\left(\rho + \int_0^t Z_s' dB_s\right).$$

Thus by the classical maximal inequality, we obtain that

$$\begin{aligned}\hat{\mathbb{E}}\left[\sup_{t\in[0,T]} \hat{\mathbb{E}}_t[|\xi|^\alpha]\right] \leq & \hat{\mathbb{E}}\left[\sup_{t\in[0,T]}\left(\rho + \int_0^t Z_s' dB_s\right)\right] \leq \left[\hat{\mathbb{E}}\left(\sup_{t\in[0,T]}(\rho + \int_0^t Z_s' dB_s)^p\right)\right]^{1/p} \\ \leq & q\left\|\rho + \int_0^T Z_t' dB_s\right\|_{L_G^p} \leq q\left(\|\,|\xi|^\alpha\|_{L_G^p} + \|K_T'\|_{L_G^p}\right),\end{aligned}$$

where $q = p/(p-1)$. We combine this with Lemma 7.1.4 to conclude that

$$\hat{\mathbb{E}}\left[\sup_{t\in[0,T]} \hat{\mathbb{E}}_t[|\xi|^\alpha]\right] \leq \frac{C_p \bar{p} q}{\bar{p}-p}\left((\hat{\mathbb{E}}[|\xi|^{\alpha+\delta}])^{1/(\bar{p}p)} + (\hat{\mathbb{E}}[|\xi|^{\alpha+\delta}])^{1/p}\right),$$

which is the desired result. □

Now we are ready to complete the proof of Theorem 7.1.1.

Proof of Theorem 7.1.1 The uniqueness of the decomposition is obvious. For each $\xi \in L_G^\beta(\Omega_T)$, there is a sequences of cylinder random variables $\{\xi_n\}_{n=1}^\infty$ of the form (7.1.8) such that $\lim_{n\to\infty} \hat{\mathbb{E}}[|\xi_n - \xi|^\beta] = 0$. It follows from Lemma 7.1.3 that, for any n, the martingale $\hat{\mathbb{E}}_t[\xi_n]$ has the decomposition

$$\hat{\mathbb{E}}_t[\xi_n] = Y_t^{(n)} = \hat{\mathbb{E}}[\xi_n] + \int_0^t Z_s^{(n)} dB_s + K_t^{(n)}, \tag{7.1.17}$$

where $Y^{(n)} \in S_G^2(0,T)$, $Z^{(n)} \in M_G^2(0,T)$ and $K^{(n)}$ is a non-increasing continuous G-martingale with $K_0^{(n)} = 0$ and $K_T^{(n)} \in L_G^2(\Omega_T)$. Since

$$|\mathbb{E}_t[\xi_n] - \mathbb{E}_t[\xi_m]|^\beta \leq \mathbb{E}_t[|\xi_n - \xi_m|^\beta], \quad \text{for each } m, n = 1, 2, \cdots,$$

It then follows from Lemma 7.1.5 that $\{Y^{(n)}\}_{n=1}^\infty$ is a Cauchy sequence in $S_G^2(0,T)$. Thus we can apply the prior estimates in Lemma 7.1.2 to show that the sequence $\{Z^{(n)}\}_{n=1}^\infty$ also converges in $M_G^2(0,T)$ and consequently $\{K_t^{(n)}\}_{n=1}^\infty$ converges in L_G^2 to K_t. It is easy to check that the triplet (Y, Z, K) satisfies (7.1.1). The proof is complete. □

Notes and Comments

Theorem 7.1.1 is a highly nontrivial generalization of the classical martingale representation problem. It shows that under a nonlinear G-Brownian motion framework, a G-martingale can be decomposed into two essentially different martingales, the first one is an Itô's integral with respect to the G-Brownian motion B, and the second one is a non-increasing G-martingale. The later term vanishes once G is a linear function and thus B becomes to a classical Brownian motion. This means this new type of no-decreasing martingale is completely neglected under the Wiener probability measure and thus we need to use the stronger norm $\|\cdot\|_{L_G^p}$, and the corresponding G-capacity, to investigate this type of very interesting martingales.

Under the condition that ξ is a cylindrical random variable, i.e., $\xi \in Lip(\Omega_T)$, Peng [140, 144] proved Theorem 7.1.1 and thus raised a challenging open problem for the proof of the general situation. An important step towards the solution to this problem was given by Soner, Touzi and Zhang in 2009, although a stronger condition was assumed. More precisely, they defined the norm

$$\|\xi^2\|_{L^2_{\mathcal{E}}} = \hat{\mathbb{E}}[\sup_{t\in[0,T]} \hat{\mathbb{E}}_t[|\xi|^2]]^{1/2},$$

and then denoted by $L^2_{\mathcal{E}}(\Omega_T)$ the completion of $Lip(\Omega_T)$ under this norm. Soner et al. (2009) showed that the martingale representation theorem holds for random variables in $L^2_{\mathcal{E}}(\Omega_T)$. A natural question is: how large this new subspace of $L^2_{\mathcal{E}}(\Omega_T)$? Can we prove that $L^2_{\mathcal{E}}(\Omega_T)$ contains $C_b(\Omega_T)$? or $L^2_{\mathcal{E}}(\Omega_T) = L^2_G(\Omega_T)$?

This challenging problem was finally independently solved by Song [159] for $\xi \in L^p_G(\Omega_T),\ \ p > 1$ and by Soner et al. [155] for $\xi \in L^{p+\epsilon}_G(\Omega_T),\ \ p = 2$.

As a by product, Song [159] established a fundamental inequality of Doob's type for G-martingale theory (see Lemma 7.1.5), which gave a characterization of the relations between $L^p_{\mathcal{E}}(\Omega_T)$ and $L^p_G(\Omega_T)$. The actual proof of Theorem 7.1.1 in this chapter is based on the Doob–Song's inequality (7.1.16). In fact this result was applied in the paper of Hu et al. [78] for the proof of existence and uniqueness theorem of BSDEs driven by G-Brownian motion.

Chapter 8
Some Further Results of Itô's Calculus

In this chapter, we use the quasi-surely analysis theory to develop Itô's integrals without the quasi-continuity condition. This allows us to define Itô's integral on stopping time interval. In particular, this new formulation can be applied to obtain Itô's formula for a general $C^{1,2}$-function, thus extending previously available results.

8.1 A Generalized Itô's Integral

Recall that $B_b(\Omega)$ is the space of all bounded and Borel measurable real functions defined on $\Omega = C_0^d(\mathbb{R}^+)$. We denote by $L_*^p(\Omega)$ the completion of $B_b(\Omega)$ under the natural norm $\|X\|_p := \hat{\mathbb{E}}[|X|^p]^{1/p}$. Similarly, we can define $L_*^p(\Omega_T)$ for any fixed $T \geq 0$. For any fixed $\mathbf{a} \in \mathbb{R}^d$, we still use the notation $B_t^{\mathbf{a}} := \langle \mathbf{a}, B_t \rangle$. Then we introduce the following properties, which are important in our stochastic calculus.

Proposition 8.1.1 *For any* $0 \leq t < T$, $\xi \in L_*^2(\Omega_t)$, *we have*

$$\hat{\mathbb{E}}[\xi(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})] = 0.$$

Proof For a fixed $P \in \mathcal{P}$, $B^{\mathbf{a}}$ is a martingale on $(\Omega, \mathcal{F}_t, P)$. Then we have

$$E_P[\xi(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})] = 0,$$

which completes the proof. □

Proposition 8.1.2 *For any* $0 \leq t \leq T$ *and* $\xi \in B_b(\Omega_t)$, *we have*

$$\hat{\mathbb{E}}[\xi^2(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^2 - \sigma_{\mathbf{a}\mathbf{a}^T}^2 \xi^2 (T-t)] \leq 0. \tag{8.1.1}$$

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7_8

Proof If $\xi \in C_b(\Omega_t)$, then we get that $\hat{\mathbb{E}}[\xi^2(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^2 - \xi^2\sigma_{\mathbf{aa}^T}^2(T-t)] = 0$. Thus (8.1.1) holds for $\xi \in C_b(\Omega_t)$. This implies that, for a fixed $P \in \mathcal{P}$,

$$E_P[\xi^2(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^2 - \xi^2\sigma_{\mathbf{aa}^T}^2(T-t)] \le 0. \tag{8.1.2}$$

If we take $\xi \in B_b(\Omega_t)$, we can find a sequence $\{\xi_n\}_{n=1}^{\infty}$ in $C_b(\Omega_t)$, such that $\xi_n \to \xi$ in $L^p(\Omega, \mathcal{F}_t, P)$, for some $p > 2$. Thus we conclude that

$$E_P[\xi_n^2(B_T^{\mathbf{a}} - B_t^{\mathbf{a}})^2 - \xi_n^2\sigma_{\mathbf{aa}^T}^2(T-t)] \le 0.$$

Then letting $n \to \infty$, we obtain (8.1.2) for $\xi \in B_b(\Omega_t)$. □

In what follows, we use the notation $L_*^p(\Omega)$, instead of $L_G^p(\Omega)$, to generalize Itô's integral on a larger space of stochastic processes $M_*^2(0,T)$ defined as follows. For fixed $p \ge 1$ and $T \in \mathbb{R}_+$, we first consider the following simple type of processes:

$$M_{b,0}(0,T) = \Big\{\eta : \eta_t(\omega) = \sum_{j=0}^{N-1} \xi_j(\omega)\mathbf{1}_{[t_j,t_{j+1})}(t),$$
$$\forall N > 0,\ \ 0 = t_0 < \cdots < t_N = T,\ \ \xi_j(\omega) \in B_b(\Omega_{t_j}),\ j = 0, \cdots, N-1\Big\}.$$

Definition 8.1.3 For an element $\eta \in M_{b,0}(0,T)$ with $\eta_t = \sum_{j=0}^{N-1} \xi_j(\omega)\mathbf{1}_{[t_j,t_j+1)}(t)$, the related Bochner integral is

$$\int_0^T \eta_t(\omega)dt = \sum_{j=0}^{N-1} \xi_j(\omega)(t_{j+1} - t_j).$$

For any $\eta \in M_{b,0}(0,T)$ we set

$$\tilde{\mathbb{E}}_T[\eta] := \frac{1}{T}\hat{\mathbb{E}}\left[\int_0^T \eta_t dt\right] = \frac{1}{T}\hat{\mathbb{E}}\left[\sum_{j=0}^{N-1} \xi_j(\omega)(t_{j+1} - t_j)\right].$$

Then $\tilde{\mathbb{E}} : M_{b,0}(0,T) \mapsto \mathbb{R}$ forms a sublinear expectation. We can introduce a natural norm $\|\eta\|_{M^p(0,T)} = \left\{\hat{\mathbb{E}}\left[\int_0^T |\eta_t|^p dt\right]\right\}^{1/p}$.

Definition 8.1.4 For any $p \ge 1$, we denote by $M_*^p(0,T)$ the completion of $M_{b,0}(0,T)$ under the norm

$$||\eta||_{M^p(0,T)} = \left\{\hat{\mathbb{E}}\left[\int_0^T |\eta_t|^p dt\right]\right\}^{1/p}.$$

We have $M_*^p(0,T) \supset M_*^q(0,T)$, for $p \le q$. The following process

$$\eta_t(\omega) = \sum_{j=0}^{N-1} \xi_j(\omega)\mathbf{1}_{[t_j,t_{j+1})}(t),\ \xi_j \in L_*^p(\Omega_{t_j}),\ j = 1, \cdots, N$$

is also in $M_*^p(0,T)$.

Definition 8.1.5 For any $\eta \in M_{b,0}(0,T)$ of the form

$$\eta_t(\omega) = \sum_{j=0}^{N-1} \xi_j(\omega)\mathbf{1}_{[t_j,t_{j+1})}(t),$$

we define Itô's integral

$$I(\eta) = \int_0^T \eta_s dB_s^{\mathbf{a}} := \sum_{j=0}^{N-1} \xi_j(B_{t_{j+1}}^{\mathbf{a}} - B_{t_j}^{\mathbf{a}}).$$

Lemma 8.1.6 *The mapping $I : M_{b,0}(0,T) \mapsto L_*^2(\Omega_T)$ is a linear continuous mapping and thus can be continuously extended to $I : M_*^2(0,T) \mapsto L_*^2(\Omega_T)$. Moreover, we have*

$$\hat{\mathbb{E}}\left[\int_0^T \eta_s dB_s^{\mathbf{a}}\right] = 0, \tag{8.1.3}$$

$$\hat{\mathbb{E}}\left[(\int_0^T \eta_s dB_s^{\mathbf{a}})^2\right] \le \sigma_{\mathbf{aa}^T}^2 \hat{\mathbb{E}}\left[\int_0^T |\eta_t|^2 dt\right]. \tag{8.1.4}$$

Proof It suffices to prove (8.1.3) and (8.1.4) for any $\eta \in M_{b,0}(0,T)$. From Proposition 8.1.1, for any j,

$$\hat{\mathbb{E}}[\xi_j(B_{t_{j+1}}^{\mathbf{a}} - B_{t_j}^{\mathbf{a}})] = \hat{\mathbb{E}}[-\xi_j(B_{t_{j+1}}^{\mathbf{a}} - B_{t_j}^{\mathbf{a}})] = 0.$$

Thus we obtain (8.1.3):

$$\begin{aligned}\hat{\mathbb{E}}\left[\int_0^T \eta_s dB_s^{\mathbf{a}}\right] &= \hat{\mathbb{E}}\left[\int_0^{t_{N-1}} \eta_s dB_s^{\mathbf{a}} + \xi_{N-1}(B_{t_N}^{\mathbf{a}} - B_{t_{N-1}}^{\mathbf{a}})\right] \\ &= \hat{\mathbb{E}}\left[\int_0^{t_{N-1}} \eta_s dB_s^{\mathbf{a}}\right] = \cdots = \hat{\mathbb{E}}[\xi_0(B_{t_1}^{\mathbf{a}} - B_{t_0}^{\mathbf{a}})] = 0.\end{aligned}$$

We now prove (8.1.4). By a similar analysis as in Lemma 3.3.4 of Chap. 3, we derive that

$$\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t^{\mathbf{a}}\right)^2\right] = \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \xi_i^2 (B_{t_{i+1}}^{\mathbf{a}} - B_{t_i}^{\mathbf{a}})^2\right].$$

Then from Proposition 8.1.2, we obtain that

$$\hat{\mathbb{E}}\left[\xi_j^2 (B_{t_{j+1}}^{\mathbf{a}} - B_{t_j}^{\mathbf{a}})^2 - \sigma_{\mathbf{aa}^T}^2 \xi_j^2 (t_{j+1} - t_j)\right] \le 0.$$

Thus

$$\begin{aligned}
\hat{\mathbb{E}}\left[\left(\int_0^T \eta_t dB_t^{\mathbf{a}}\right)^2\right] &= \hat{\mathbb{E}}[\sum_{i=0}^{N-1} \xi_i^2 (B_{t_N}^{\mathbf{a}} - B_{t_{N-1}}^{\mathbf{a}})^2] \\
&\le \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \xi_i^2 [(B_{t_N}^{\mathbf{a}} - B_{t_{N-1}}^{\mathbf{a}})^2 - \sigma_{\mathbf{aa}^T}^2 (t_{i+1} - t_i)]\right] + \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \sigma_{\mathbf{aa}^T}^2 \xi_i^2 (t_{i+1} - t_i)\right] \\
&\le \hat{\mathbb{E}}\left[\sum_{i=0}^{N-1} \sigma_{\mathbf{aa}^T}^2 \xi_i^2 (t_{i+1} - t_i)\right] = \sigma_{\mathbf{aa}^T}^2 \hat{\mathbb{E}}\left[\int_0^T |\eta_t|^2 dt\right],
\end{aligned}$$

which is the desired result. □

The following proposition can be verified directly by the definition of Itô's integral with respect to G-Brownian motion.

Proposition 8.1.7 *Let $\eta, \theta \in M_*^2(0,T)$. Then for any $0 \le s \le r \le t \le T$, we have:*

(i) $\int_s^t \eta_u dB_u^{\mathbf{a}} = \int_s^r \eta_u dB_u^{\mathbf{a}} + \int_r^t \eta_u dB_u^{\mathbf{a}}$;
(ii) $\int_s^t (\alpha\eta_u + \theta_u) dB_u^{\mathbf{a}} = \alpha \int_s^t \eta_u dB_u^{\mathbf{a}} + \int_s^t \theta_u dB_u^{\mathbf{a}}$, *where* $\alpha \in B_b(\Omega_s)$.

Proposition 8.1.8 *For any $\eta \in M_*^2(0,T)$, we have*

$$\hat{\mathbb{E}}\left[\sup_{0\le t\le T}\left|\int_0^t \eta_s dB_s^{\mathbf{a}}\right|^2\right] \le 4\sigma_{\mathbf{aa}^T}^2 \hat{\mathbb{E}}\left[\int_0^T \eta_s^2 ds\right]. \tag{8.1.5}$$

Proof Since for any $\alpha \in B_b(\Omega_t)$, we have

$$\hat{\mathbb{E}}\left[\alpha \int_t^T \eta_s dB_s^{\mathbf{a}}\right] = 0.$$

Then, for a fixed $P \in \mathcal{P}$, the process $\int_0^\cdot \eta_s dB_s^{\mathbf{a}}$ is a martingale on $(\Omega, \mathcal{F}_t, P)$. It follows from the classical Doob's maximal inequality (see Appendix B) that

$$E_P\left[\sup_{0\le t\le T}\left|\int_0^t \eta_s dB_s^{\mathbf{a}}\right|^2\right] \le 4E_P\left[\left|\int_0^T \eta_s dB_s^{\mathbf{a}}\right|^2\right] \le 4\hat{\mathbb{E}}\left[\left|\int_0^T \eta_s dB_s^{\mathbf{a}}\right|^2\right]$$
$$\le 4\sigma^2_{\mathbf{a}\mathbf{a}^T}\hat{\mathbb{E}}\left[\int_0^T \eta_s^2 ds\right].$$

Thus (8.1.5) holds. □

Proposition 8.1.9 *For any $\eta \in M_*^2(0,T)$ and $0 \le t \le T$, the integral $\int_0^t \eta_s dB_s^{\mathbf{a}}$ is continuous q.s., i.e., $\int_0^t \eta_s dB_s^{\mathbf{a}}$ has a modification whose paths are continuous in t.*

Proof The claim is true for $\eta \in M_{b,0}(0,T)$ since $(B_t^{\mathbf{a}})_{t\ge 0}$ is a continuous process. In the case of $\eta \in M_*^2(0,T)$, there exists $\eta^{(n)} \in M_{b,0}(0,T)$, such that $\hat{\mathbb{E}}[\int_0^T (\eta_s - \eta_s^{(n)})^2 ds] \to 0$, as $n \to \infty$. By Proposition 8.1.8, we have

$$\hat{\mathbb{E}}\left[\sup_{0\le t\le T}\left|\int_0^t (\eta_s - \eta_s^{(n)}) dB_s^{\mathbf{a}}\right|^2\right] \le 4\sigma^2_{\mathbf{a}\mathbf{a}^T}\hat{\mathbb{E}}\left[\int_0^T (\eta_s - \eta_s^{(n)})^2 ds\right] \to 0, \quad \text{as } n \to \infty.$$

Then choosing a subsequence if necessary, we can find a set $\hat{\Omega} \subset \Omega$ with $\hat{c}(\hat{\Omega}^c) = 0$ so that, for any $\omega \in \hat{\Omega}$ the sequence of processes $\int_0^\cdot \eta_s^{(n)} dB_s^{\mathbf{a}}(\omega)$ uniformly converges to $\int_0^\cdot \eta_s dB_s^{\mathbf{a}}(\omega)$ on $[0,T]$. Thus for any $\omega \in \hat{\Omega}$, we get that $\int_0^\cdot \eta_s dB_s^{\mathbf{a}}(\omega)$ is continuous in t. For any $(\omega, t) \in [0,T] \times \Omega$, we take the process

$$J_t(\omega) = \begin{cases} \int_0^t \eta_s dB_s^{\mathbf{a}}(\omega), & \omega \in \hat{\Omega}; \\ 0, & \text{otherwise}, \end{cases}$$

as the desired t-continuous modification. This completes the proof. □

We now define the integral of a process $\eta \in M_*^1(0,T)$ with respect to $\langle B^{\mathbf{a}}\rangle$. We also define a mapping:

$$Q_{0,T}(\eta) = \int_0^T \eta_t d\left\langle B^{\mathbf{a}}\right\rangle_t := \sum_{j=0}^{N-1} \xi_j(\left\langle B^{\mathbf{a}}\right\rangle_{t_{j+1}} - \left\langle B^{\mathbf{a}}\right\rangle_{t_j}) : M_b^{1,0}(0,T) \to L_*^1(\Omega_T).$$

Proposition 8.1.10 *The mapping $Q_{0,T} : M_b^{1,0}(0,T) \mapsto L_*^1(\Omega_T)$ is a continuous linear mapping and $Q_{0,T}$ can be uniquely extended to $M_*^1(0,T)$. Moreover, we have*

$$\hat{\mathbb{E}}\left[\left|\int_0^T \eta_t d\left\langle B^{\mathbf{a}}\right\rangle_t\right|\right] \le \sigma^2_{\mathbf{a}\mathbf{a}^T}\hat{\mathbb{E}}\left[\int_0^T |\eta_t| dt\right] \quad \textit{for any } \eta \in M_*^1(0,T). \tag{8.1.6}$$

Proof From the relation

$$\sigma^2_{-\mathbf{aa}^T}(t-s) \le \left\langle B^{\mathbf{a}}\right\rangle_t - \left\langle B^{\mathbf{a}}\right\rangle_s \le \sigma^2_{\mathbf{aa}^T}(t-s)$$

it follows that

$$\hat{\mathbb{E}}\left[|\xi_j|(\langle B^{\mathbf{a}}\rangle_{t_{j+1}} - \langle B^{\mathbf{a}}\rangle_{t_j}) - \sigma^2_{\mathbf{aa}^T}|\xi_j|(t_{j+1}-t_j)\right] \le 0, \quad \text{for any} \quad j = 1, \cdots, N-1.$$

Therefore, we deduce the following chain of inequalities:

$$\begin{aligned}
&\hat{\mathbb{E}}\left[\left|\sum_{j=0}^{N-1} \xi_j(\left\langle B^{\mathbf{a}}\right\rangle_{t_{j+1}} - \left\langle B^{\mathbf{a}}\right\rangle_{t_j})\right|\right] \le \hat{\mathbb{E}}\left[\sum_{j=0}^{N-1} |\xi_j|\left\langle B^{\mathbf{a}}\right\rangle_{t_{j+1}} - \left\langle B^{\mathbf{a}}\right\rangle_{t_j}\right] \\
\le &\hat{\mathbb{E}}\left[\sum_{j=0}^{N-1} |\xi_j|[(\langle B^{\mathbf{a}}\rangle_{t_{j+1}} - \langle B^{\mathbf{a}}\rangle_{t_j}) - \sigma^2_{\mathbf{aa}^T}(t_{j+1}-t_j)]\right] + \hat{\mathbb{E}}\left[\sigma^2_{\mathbf{aa}^T}\sum_{j=0}^{N-1}|\xi_j|(t_{j+1}-t_j)\right] \\
\le &\sum_{j=0}^{N-1} \hat{\mathbb{E}}[|\xi_j|[(\langle B^{\mathbf{a}}\rangle_{t_{j+1}} - \langle B^{\mathbf{a}}\rangle_{t_j}) - \sigma^2_{\mathbf{aa}^T}(t_{j+1}-t_j)]] + \hat{\mathbb{E}}\left[\sigma^2_{\mathbf{aa}^T}\sum_{j=0}^{N-1}|\xi_j|(t_{j+1}-t_j)\right] \\
\le &\hat{\mathbb{E}}\left[\sigma^2_{\mathbf{aa}^T}\sum_{j=0}^{N-1}|\xi_j|(t_{j+1}-t_j)\right] = \sigma^2_{\mathbf{aa}^T}\hat{\mathbb{E}}\left[\int_0^T |\eta_t|dt\right].
\end{aligned}$$

This completes the proof. □

From the above Proposition 8.1.9, we obtain that $\langle B^{\mathbf{a}}\rangle_t$ is continuous in t q.s.. Then for any $\eta \in M_*^1(0,T)$ and $0 \le t \le T$, the integral $\int_0^t \eta_s d\langle B^{\mathbf{a}}\rangle_s$ also has a t-continuous modification. In the sequel, we always consider the t-continuous modification of Itô's integral. Moreover, Itô's integral with respect to $\langle B^i, B^j\rangle = \langle B\rangle^{ij}$ can be similarly defined. This is left as an exercise for the readers.

Lemma 8.1.11 *Let $\eta \in M_b^2(0,T)$. Then η is Itô-integrable for every $P \in P$. Moreover,*

$$\int_0^T \eta_s dB_s^{\mathbf{a}} = \int_0^T \eta_s d_P B_s^{\mathbf{a}}, \quad P\text{-a.s.},$$

where the right hand side is the usual Itô integral.

We leave the proof of this lemma to readers as an exercise.

Lemma 8.1.12 (Generalized Burkholder-Davis-Gundy (BDG) inequality) *For any $\eta \in M_*^2(0,T)$ and $p > 0$, there exist constants c_p and C_p with $0 < c_p < C_p < \infty$, depending only on p, such that*

$$\sigma^p_{-\mathbf{aa}^T}c_p\hat{\mathbb{E}}\left[\left(\int_0^T |\eta_s|^2 ds\right)^{p/2}\right] \le \hat{\mathbb{E}}\left[\sup_{t\in[0,T]}\left|\int_0^t \eta_s dB_s^{\mathbf{a}}\right|^p\right] \le \sigma^p_{\mathbf{aa}^T}C_p\hat{\mathbb{E}}\left[\left(\int_0^T |\eta_s|^2 ds\right)^{p/2}\right].$$

Proof Observe that, under any $P \in \mathcal{P}$, $B^{\mathbf{a}}$ is a P-martingale with

$$\sigma^2_{-\mathbf{a}\mathbf{a}^T} dt \le d\langle B^{\mathbf{a}}\rangle_t \le \sigma^2_{\mathbf{a}\mathbf{a}^T} dt.$$

The proof is then a simple application of the classical BDG inequality. □

8.2 Itô's Integral for Locally Integrable Processes

So far we have considered Itô's integral $\int_0^T \eta_t dB_t^{\mathbf{a}}$ where η in $M_*^2(0,T)$. In this section we continue our study of Itô's integrals for a type of locally integrable processes.

We first give some properties of $M_*^p(0,T)$.

Lemma 8.2.1 *For any $p \ge 1$ and $X \in M_*^p(0,T)$, the following relation holds:*

$$\lim_{n\to\infty} \hat{\mathbb{E}}\left[\int_0^T |X_t|^p \mathbf{1}_{\{|X_t|>n\}} dt\right] = 0. \tag{8.2.1}$$

Proof The proof is similar to that of Proposition 6.1.22 in Chap. 6. □

Corollary 8.2.2 *For any $\eta \in M_*^2(0,T)$, let $\eta_s^{(n)} = (-n) \vee (\eta_s \wedge n)$, then, as $n \to \infty$, we have $\int_0^t \eta_s^{(n)} dB_s^{\mathbf{a}} \to \int_0^t \eta_s dB_s^{\mathbf{a}}$ in $L_*^2(0,T)$ for any $t \le T$.*

Proposition 8.2.3 *Let $X \in M_*^p(0,T)$. Then for any $\varepsilon > 0$, there exists a constant $\delta > 0$ such that for all $\eta \in M_*^p(0,T)$ satisfying $\hat{\mathbb{E}}\left[\int_0^T |\eta_t| dt\right] \le \delta$ and $|\eta_t(\omega)| \le 1$, we have $\hat{\mathbb{E}}\left[\int_0^T |X_t|^p |\eta_t| dt\right] \le \varepsilon$.*

Proof For any $\varepsilon > 0$, according to Lemma 8.2.1, there exists a number $N > 0$ such that $\hat{\mathbb{E}}\left[\int_0^T |X|^p \mathbf{1}_{\{|X|>N\}}\right] \le \varepsilon/2$. Take $\delta = \varepsilon/2N^p$. Then we derive that

$$\begin{aligned}\hat{\mathbb{E}}\left[\int_0^T |X_t|^p |\eta_t| dt\right] &\le \hat{\mathbb{E}}\left[\int_0^T |X_t|^p |\eta_t| \mathbf{1}_{\{|X_t|>N\}} dt\right] + \hat{\mathbb{E}}\left[\int_0^T |X_t|^p |\eta_t| \mathbf{1}_{\{|X_t|\le N\}} dt\right] \\ &\le \hat{\mathbb{E}}\left[\int_0^T |X_t|^p \mathbf{1}_{\{|X_t|>N\}} dt\right] + N^p \hat{\mathbb{E}}\left[\int_0^T |\eta_t| dt\right] \le \varepsilon,\end{aligned}$$

which is the desired result. □

Lemma 8.2.4 *If $p \ge 1$ and $X, \eta \in M_*^p(0,T)$ are such that η is bounded, then the product $X\eta \in M_*^p(0,T)$.*

Proof We can find $X^{(n)}, \eta^{(n)} \in M_{b,0}(0,T)$ for $n = 1, 2, \ldots$, such that $\eta^{(n)}$ is uniformly bounded and

$$\|X - X^{(n)}\|_{M^p(0,T)} \to 0, \quad \|\eta - \eta^{(n)}\|_{M^p(0,T)} \to 0, \quad \text{as} \quad n \to \infty.$$

Then we obtain that

$$\hat{\mathbb{E}}\left[\int_0^T |X_t\eta_t - X_t^{(n)}\eta_t^{(n)}|^p dt\right] \le 2^{p-1}\left(\hat{\mathbb{E}}\left[\int_0^T |X_t|^p|\eta_t - \eta_t^{(n)}|^p dt\right] + \hat{\mathbb{E}}\left[\int_0^T |X_t - X_t^{(n)}|^p|\eta_t^{(n)}|^p dt\right]\right).$$

By Proposition 8.2.3, the first term on the right-hand side tends to 0 as $n \to \infty$. Since $\eta^{(n)}$ is uniformly bounded, the second term also tends to 0. □

Now we are going to study Itô's integrals on an interval $[0, \tau]$, where τ is a stopping time relative to the G-Brownian pathes.

Definition 8.2.5 A stopping time τ relative to the filtration $(\mathcal{F}_t)$ is a map on Ω with values in $[0, T]$ such that $\{\tau \le t\} \in \mathcal{F}_t$, for every $t \in [0, T]$.

Lemma 8.2.6 *For any stopping time τ and any $X \in M_*^p(0, T)$, we have $\mathbf{1}_{[0,\tau]}(\cdot)X \in M_*^p(0, T)$.*

Proof Related to the given stopping time τ, we consider the following sequence:

$$\tau_n = \sum_{k=0}^{2^n-1} \frac{(k+1)T}{2^n}\mathbf{1}_{[\frac{kT}{2^n} \le \tau < \frac{(k+1)T}{2^n})} + T\mathbf{1}_{[\tau \ge T]}.$$

It is clear that $2^{-n} \ge \tau_n - \tau \ge 0$. It follows from Lemma 8.2.4 that any element of the sequence $\{1_{[0,\tau_n]}X\}_{n=1}^{\infty}$ is in $M_*^p(0, T)$. Note that, for $m \ge n$, we have

$$\begin{aligned}\hat{\mathbb{E}}\left[\int_0^T |\mathbf{1}_{[0,\tau_n]}(t) - \mathbf{1}_{[0,\tau_m]}(t)|dt\right] &\le \hat{\mathbb{E}}\left[\int_0^T |\mathbf{1}_{[0,\tau_n]}(t) - \mathbf{1}_{[0,\tau]}(t)|dt\right] \\ &= \hat{\mathbb{E}}[\tau_n - \tau] \le 2^{-n}T.\end{aligned}$$

Then applying Proposition 8.2.3, we derive that $\mathbf{1}_{[0,\tau]}X \in M_*^p(0, T)$ and the proof is complete. □

Lemma 8.2.7 *For any stopping time τ and any $\eta \in M_*^2(0, T)$, we have*

$$\int_0^{t\wedge\tau} \eta_s dB_s^{\mathbf{a}}(\omega) = \int_0^t \mathbf{1}_{[0,\tau]}(s)\eta_s dB_s^{\mathbf{a}}(\omega), \text{ for all } t \in [0, T] \text{ q.s.} \tag{8.2.2}$$

Proof For any $n \in \mathbb{N}$, let

$$\tau_n := \sum_{k=1}^{[t\cdot 2^n]} \frac{k}{2^n}\mathbf{1}_{[\frac{(k-1)t}{2^n} \le \tau < \frac{kt}{2^n})} + t\mathbf{1}_{[\tau \ge t]} = \sum_{k=1}^{2^n} \mathbf{1}_{A_n^k} t_n^k.$$

Here $t_n^k = k2^{-n}t$, $A_n^k = [t_n^{k-1} < t \wedge \tau \le t_n^k]$, for $k < 2^n$, and $A_n^{2^n} = [\tau \ge t]$. We see that $\{\tau_n\}_{n=1}^{\infty}$ is a decreasing sequence of stopping times which converges to $t \wedge \tau$.

We first show that

$$\int_{\tau_n}^{t} \eta_s dB_s^{\mathbf{a}} = \int_0^t \mathbf{1}_{[\tau_n,t]}(s)\eta_s dB_s^{\mathbf{a}}, \quad \text{q.s.} \tag{8.2.3}$$

By Proposition 8.1.7 we have

$$\begin{aligned}\int_{\tau_n}^{t} \eta_s dB_s^{\mathbf{a}} &= \int_{\sum_{k=1}^{2^n} \mathbf{1}_{A_n^k} t_n^k}^{t} \eta_s dB_s^{\mathbf{a}} = \sum_{k=1}^{2^n} \mathbf{1}_{A_n^k} \int_{t_n^k}^{t} \eta_s dB_s^{\mathbf{a}} \\ &= \sum_{k=1}^{2^n} \int_{t_n^k}^{t} \mathbf{1}_{A_n^k} \eta_s dB_s^{\mathbf{a}} = \int_0^t \sum_{k=1}^{2^n} \mathbf{1}_{[t_n^k,t]}(s) \mathbf{1}_{A_n^k} \eta_s dB_s^{\mathbf{a}},\end{aligned}$$

from which (8.2.3) follows. Hence we obtain that

$$\int_0^{\tau_n} \eta_s dB_s^{\mathbf{a}} = \int_0^t \mathbf{1}_{[0,\tau_n]}(s)\eta_s dB_s^{\mathbf{a}}, \text{ q.s.}$$

Observe now that $0 \le \tau_n - \tau_m \le \tau_n - t \wedge \tau \le 2^{-n}t$, for $n \le m$. Then Proposition 8.2.3 yields that $\mathbf{1}_{[0,\tau_n]}\eta$ converges in $M_*^2(0,T)$ to $\mathbf{1}_{[0,\tau\wedge t]}\eta$ as $n \to \infty$, which implies that $\mathbf{1}_{[0,\tau\wedge t]}\eta \in M_*^2(0,T)$. Consequently,

$$\lim_{n\to\infty} \int_0^{\tau_n} \eta_s dB_s^{\mathbf{a}} = \int_0^t \mathbf{1}_{[0,\tau]}(s)\eta_s dB_s^{\mathbf{a}}, \quad \text{q.s.}$$

Note that $\int_0^t \eta_s dB_s^{\mathbf{a}}$ is continuous in t, hence (8.2.2) is proved. □

The space of processes $M_*^p(0,T)$ can be further enlarged as follows.

Definition 8.2.8 For fixed $p \ge 1$, a stochastic process η is said to be in $M_w^p(0,T)$, if it is associated with a sequence of increasing stopping times $\{\sigma_m\}_{m\in\mathbb{N}}$, such that:

(i) For any $m \in \mathbb{N}$, the process $\left(\eta_t \mathbf{1}_{[0,\sigma_m]}(t)\right)_{t\in[0,T]} \in M_*^p(0,T)$;
(ii) If $\Omega^{(m)} := \{\omega \in \Omega : \sigma_m(\omega) \wedge T = T\}$ and $\hat{\Omega} := \lim_{m\to\infty} \Omega^{(m)}$, then $\hat{c}(\hat{\Omega}^c) = 0$.

Remark 8.2.9 Suppose there is another sequence of stopping times $\{\tau_m\}_{m=1}^{\infty}$ that satisfies the second condition in Definition 8.2.8. Then the sequence $\{\tau_m \wedge \sigma_m\}_{m\in\mathbb{N}}$ also satisfies this condition. Moreover, by Lemma 8.2.6, we know that for any $m \in \mathbb{N}$, $\eta \mathbf{1}_{[0,\tau_m\wedge\sigma_m]} \in M_*^p(0,T)$. This property allows to associate the same sequence of stopping times with several different processes in $M_w^p(0,T)$.

For given $\eta \in M_w^2(0,T)$ associated with $\{\sigma_m\}_{m\in\mathbb{N}}$, we consider, for any $m \in \mathbb{N}$, the t-continuous modification of the process $\left(\int_0^t \eta_s \mathbf{1}_{[0,\sigma_m]}(s) dB_s^{\mathbf{a}}\right)_{0\le t\le T}$. For any m, $n \in \mathbb{N}$ with $n > m$, by Lemma 8.2.7 we can find a polar set $\hat{A}_{m,n}$, such that for all $\omega \in (\hat{A}_{m,n})^c$, the following equalities hold:

$$\begin{aligned}\int_0^{t\wedge\sigma_m}\eta_s\mathbf{1}_{[0,\sigma_m]}(s)dB_s^{\mathbf{a}}(\omega) &= \int_0^t \eta_s\mathbf{1}_{[0,\sigma_m]}(s)dB_s^{\mathbf{a}}(\omega)\\ &= \int_0^t \eta_s\mathbf{1}_{[0,\sigma_m]}(s)\mathbf{1}_{[0,\sigma_n]}(s)dB_s^{\mathbf{a}}(\omega) \\ &= \int_0^{t\wedge\sigma_m}\eta_s\mathbf{1}_{[0,\sigma_n]}(s)dB_s^{\mathbf{a}}(\omega),\ 0\le t\le T.\end{aligned}\tag{8.2.4}$$

Define the polar set

$$\hat{A} := \bigcup_{m=1}^{\infty}\bigcup_{n=m+1}^{\infty}\hat{A}_{m,n}.$$

For any $m\in\mathbb{N}$ and any $(\omega,t)\in\Omega\times[0,T]$, we set

$$X_t^{(m)}(\omega) := \begin{cases}\int_0^t \eta_s\mathbf{1}_{[0,\sigma_m]}(s)dB_s^{\mathbf{a}}(\omega), & \omega\in\hat{A}^c\cap\hat{\Omega};\\ 0, & \text{otherwise.}\end{cases}$$

From (8.2.4), for any $m,n\in\mathbb{N}$ with $n>m$, $X^{(n)}(\omega)\equiv X^{(m)}(\omega)$ on $[0,\sigma_m(\omega)\wedge T]$ for any $\omega\in\hat{A}^c\cap\hat{\Omega}$ and $X^{(n)}(\omega)\equiv X^{(m)}(\omega)$ on $[0,T]$ for all other ω. Note that for $\omega\in\hat{A}^c\cap\hat{\Omega}$, we can find $m\in\mathbb{N}$, such that $\sigma_m(\omega)\wedge T=T$. Consequently, for any $\omega\in\Omega$, $\lim_{m\to\infty}X_t^{(m)}(\omega)$ exists for any t. From Lemma 8.2.7, it is not difficult to verify that choosing a different sequence of stopping times will only alter this limitation on the polar set. The details are left to the reader. Thus, the following definition is well posed.

Definition 8.2.10 Giving $\eta\in M_w^2([0,T])$, for any $(\omega,t)\in\Omega\times[0,T]$, we define

$$\int_0^t \eta_s dB_s^{\mathbf{a}}(\omega) := \lim_{m\to\infty}X_t^{(m)}(\omega).\tag{8.2.5}$$

For any $\omega\in\Omega$ and $t\in[0,\sigma_m]$, $\int_0^t\eta_s dB_s^{\mathbf{a}}(\omega)=X_t^{(m)}(\omega)$, $0\le t\le T$. Since each of the processes $\{X_t^{(m)}\}_{0\le t\le T}$ has t-continuous paths, we conclude that the paths of $\left(\int_0^t\eta_s dB_s^{\mathbf{a}}\right)_{0\le t\le T}$ are also t-continuous. The following theorem is an direct consequence of the above discussion.

Theorem 8.2.11 *Assume that $\eta\in M_w^2([0,T])$. Then the stochastic process $\int_0^\cdot\eta_s dB_s^{\mathbf{a}}$ is a well-defined continuous process on* $[0,T]$.

For any $\eta\in M_w^1(0,T)$, the integrals $\int_0^t\eta_s d\langle B^{\mathbf{a}}\rangle_s$ and $\int_0^t\eta_s d\langle B\rangle_s^{ij}$ are both well-defined continuous stochastic processes on $[0,T]$ by a similar analysis.

8.3 Itô's Formula for General C^2 Functions

The objective of this section is to give a very general form of Itô's formula with respect to G-Brownian motion, which is comparable with that from the classical Itô's calculus.

Consider the following G-Itô diffusion process:

$$X_t^\nu = X_0^\nu + \int_0^t \alpha_s^\nu ds + \int_0^t \eta_s^{\nu ij} d\left\langle B \right\rangle_s^{ij} + \int_0^t \beta_s^{\nu j} dB_s^j.$$

Lemma 8.3.1 *Suppose that $\Phi \in C^2(\mathbb{R}^n)$ and that all first and second order derivatives of Φ are in $C_{b,Lip}(\mathbb{R}^n)$. Let α^ν, $\beta^{\nu j}$ and $\eta^{\nu ij}$, $\nu = 1, \cdots, n$, $i, j = 1, \cdots, d$, be bounded processes in $M_*^2(0, T)$. Then for any $t \geq 0$, we have in $L_*^2(\Omega_t)$,*

$$\begin{aligned}\Phi(X_t) - \Phi(X_0) = & \int_0^t \partial_{x^\nu}\Phi(X_u)\beta_u^{\nu j} dB_u^j + \int_0^t \partial_{x^\nu}\Phi(X_u)\alpha_u^\nu du \\ & + \int_0^t [\partial_{x^\nu}\Phi(X_u)\eta_u^{\nu ij} + \frac{1}{2}\partial^2_{x^\mu x^\nu}\Phi(X_u)\beta_u^{\mu i}\beta_u^{\nu j}] d\left\langle B \right\rangle_u^{ij}.\end{aligned} \tag{8.3.1}$$

The proof is parellel to that of Proposition 6.3, in Chap. 3. The details are left as an exercise for the readers.

Lemma 8.3.2 *Suppose that $\Phi \in C^2(\mathbb{R}^n)$ and all first and second order derivatives of Φ are in $C_{b,Lip}(\mathbb{R}^n)$. Let α^ν, $\beta^{\nu j}$ be in $M_*^1(0, T)$ and $\eta^{\nu ij}$ belong to $M_*^2(0, T)$ for $\nu = 1, \cdots, n$, $i, j = 1, \cdots, d$. Then for any $t \geq 0$, relation (8.3.1) holds in $L_*^1(\Omega_t)$.*

Proof For simplicity, we only deal with the case $n = d = 1$. Let $\alpha^{(k)}$, $\beta^{(k)}$ and $\eta^{(k)}$ be bounded processes such that, as $k \to \infty$,

$$\alpha^{(k)} \to \alpha, \ \ \eta^{(k)} \to \eta \text{ in } M_*^1(0, T) \ \text{ and } \ \beta^{(k)} \to \beta \text{ in } M_*^2(0, T)$$

and let

$$X_t^{(k)} = X_0 + \int_0^t \alpha_s^{(k)} ds + \int_0^t \eta_s^{(k)} d\langle B \rangle_s + \int_0^t \beta_s^{(k)} dB_s.$$

Then applying Hölder's inequality and BDG inequality yields that

$$\lim_{k\to\infty} \hat{\mathbb{E}}[\sup_{0\leq t\leq T} |X_t^{(k)} - X_t|] = 0 \ \text{ and } \ \lim_{k\to\infty} \hat{\mathbb{E}}[\sup_{0\leq t\leq T} |\Phi(X_t^{(k)}) - \Phi(X_t)|] = 0.$$

Note that

$$\begin{aligned}
&\hat{\mathbb{E}}\left[\int_0^T |\partial_x\Phi(X_t^{(k)})\beta_t^{(k)} - \partial_x\Phi(X_t)\beta_t|^2 dt\right] \\
&\quad\le 2\hat{\mathbb{E}}\left[\int_0^T |\partial_x\Phi(X_t^{(k)})\beta_t^{(k)} - \partial_x\Phi(X_t^{(k)})\beta_t|^2 dt\right] \\
&\qquad + 2\hat{\mathbb{E}}\left[\int_0^T |\partial_x\Phi(X_t^{(k)})\beta_t - \partial_x\Phi(X_t)\beta_t|^2 dt\right] \\
&\quad\le 2C^2\hat{\mathbb{E}}\left[\int_0^T |\beta_t^{(k)} - \beta_t|^2 dt\right] + 2\hat{\mathbb{E}}\left[\int_0^T |\beta_t|^2|\partial_x\Phi(X_t^{(k)}) - \partial_x\Phi(X_t)|^2 dt\right],
\end{aligned}$$

where C is the upper bound of $\partial_x\Phi$. Since $\sup_{0\le t\le T}|\partial_x\Phi(X_t^{(k)}) - \partial_x\Phi(X_t)|^2 \le 4C^2$, we conclude that

$$\hat{\mathbb{E}}\left[\int_0^T |\partial_x\Phi(X_t^{(k)}) - \partial_x\Phi(X_t)|^2 dt\right] \to 0, \text{ as } k\to\infty.$$

Thus we can apply Proposition 8.2.3 to prove that, in $M_*^2(0,T)$, as $k\to\infty$,

$$\begin{aligned}
&\partial_x\Phi(X^{(k)})\beta^{(k)} \to \partial_x\Phi(X)\beta, \quad \partial_x\Phi(X^{(k)})\alpha^{(k)} \to \partial_x\Phi(X)\alpha, \\
&\quad \partial_x\Phi(X^{(k)})\eta^{(k)} \to \partial_x\Phi(X)\eta, \quad \partial_{xx}^2\Phi(X^{(k)})(\beta^{(k)})^2 \to \partial_{xx}^2\Phi(X)\beta^2.
\end{aligned}$$

However, from the above lemma we have

$$\begin{aligned}
\Phi(X_t^{(k)}) - \Phi(X_0^{(k)}) &= \int_0^t \partial_x\Phi(X_u^{(k)})\beta_u^{(k)} dB_u + \int_0^t \partial_x\Phi(X_u^{(k)})\alpha_u^{(k)} du \\
&\quad + \int_0^t [\partial_x\Phi(X_u^{(k)})\eta_u^{(k)} + \tfrac{1}{2}\partial_{xx}^2\Phi(X_u^{(k)})(\beta_u^{(k)})^2] d\langle B\rangle_u.
\end{aligned}$$

Therefore passing to the limit on both sides of this equality, we obtain the desired result. □

Lemma 8.3.3 *Let X be given as in Lemma 8.3.2 and let $\Phi\in C^{1,2}([0,T]\times\mathbb{R}^n)$ be such that Φ, $\partial_t\Phi$, $\partial_x\Phi$ and $\partial_{xx}^2\Phi$ are bounded and uniformly continuous on $[0,T]\times\mathbb{R}^n$. Then we have the following relation in $L_*^1(\Omega_t)$:*

$$\begin{aligned}
\Phi(t,X_t) - \Phi(0,X_0) &= \int_0^t \partial_{x^\nu}\Phi(u,X_u)\beta_u^{\nu j} dB_u^j + \int_0^t [\partial_t\Phi(u,X_u) + \partial_{x^\nu}\Phi(u,X_u)\alpha_u^\nu] du \\
&\quad + \int_0^t [\partial_{x^\nu}\Phi(u,X_u)\eta_u^{\nu ij} + \tfrac{1}{2}\partial_{x^\mu x^\nu}^2\Phi(u,X_u)\beta_u^{\mu i}\beta_u^{\nu j}] d\langle B\rangle_u^{ij}.
\end{aligned}$$

Proof Choose a sequence of functions $\{\Phi_k\}_{k=1}^\infty$ such that, Φ_k and all its first order and second order derivatives are in $C_{b,Lip}([0,T]\times\mathbb{R}^n)$. Moreover, as $n\to\infty$, Φ_n,

$\partial_t\Phi_n$, $\partial_x\Phi_n$ and $\partial^2_{xx}\Phi_n$ converge respectively to Φ, $\partial_t\Phi$, $\partial_x\Phi$ and $\partial^2_{xx}\Phi$ uniformly on $[0, T]\times\mathbb{R}$. Then we use the above Itô's formula to $\Phi_k(X^0_t, X_t)$, with $Y_t = (X^0_t, X_t)$, where $X^0_t \equiv t$:

$$\Phi_k(t, X_t) - \Phi_k(0, X_0) = \int_0^t \partial_{x^\nu}\Phi_k(u, X_u)\beta^{\nu j}_u dB^j_u + \int_0^t [\partial_t\Phi_k(u, X_u) + \partial_{x^\nu}\Phi_k(u, X_u)\alpha^\nu_u]du + \int_0^t [\partial_{x^\nu}\Phi_k(u, X_u)\eta^{\nu ij}_u + \tfrac{1}{2}\partial^2_{x^\mu x^\nu}\Phi_k(u, X_u)\beta^{\mu i}_u\beta^{\nu j}_u]d\langle B\rangle^{ij}_u .$$

It follows that, as $k\to\infty$, the following uniform convergences:

$$|\partial_{x_\nu}\Phi_k(u, X_u) - \partial_{x_\nu}\Phi(u, X_u)| \to 0,\ |\partial^2_{x_\mu x_\nu}\Phi_k(u, X_u) - \partial^2_{x_\mu x_\nu}\Phi_k(u, X_u)| \to 0,$$
$$|\partial_t\Phi_k(u, X_u) - \partial_t\Phi(u, X_u)| \to 0.$$

Sending $k\to\infty$, we arrive at the desired result. □

Theorem 8.3.4 *Suppose $\Phi\in C^{1,2}([0, T]\times\mathbb{R}^n)$. Let α^ν, $\eta^{\nu ij}$ be in $M^1_w(0, T)$ and $\beta^{\nu j}$ be in $M^2_w(0, T)$ associated with a common stopping time sequence $\{\sigma_m\}_{m=1}^\infty$. Then for any $t\geq 0$, we have q.s.*

$$\Phi(t, X_t) - \Phi(0, X_0) = \int_0^t \partial_{x^\nu}\Phi(u, X_u)\beta^{\nu j}_u dB^j_u + \int_0^t [\partial_t\Phi(u, X_u) + \partial_{x^\nu}\Phi(u, X_u)\alpha^\nu_u]du + \int_0^t [\partial_{x^\nu}\Phi(u, X_u)\eta^{\nu ij}_u + \tfrac{1}{2}\partial^2_{x^\mu x^\nu}\Phi(u, X_u)\beta^{\mu i}_u\beta^{\nu j}_u]d\langle B\rangle^{ij}_u .$$

Proof For simplicity, we only deal with the case $n = d = 1$. We set, for $k = 1, 2, \cdots$,

$$\tau_k := \inf\{t\geq 0|\ \ |X_t - X_0| > k\}\wedge\sigma_k.$$

Let Φ_k be a $C^{1,2}$-function on $[0, T]\times\mathbb{R}^n$ such that Φ_k, $\partial_t\Phi_k$, $\partial_{x_i}\Phi_k$ and $\partial^2_{x_i x_j}\Phi_k$ are uniformly bounded continuous functions satisfying $\Phi_k = \Phi$, for $|x|\leq 2k, t\in[0, T]$. It is clear that the process $\mathbf{1}_{[0,\tau_k]}\beta$ is in $M^2_w(0, T)$, while $\mathbf{1}_{[0,\tau_k]}\alpha$ and $\mathbf{1}_{[0,\tau_k]}\eta$ are in $M^1_w(0, T)$ and they are all associated to the same sequence of stopping times $\{\tau_k\}_{k=1}^\infty$. We also have

$$X_{t\wedge\tau_k} = X_0 + \int_0^t \alpha_s\mathbf{1}_{[0,\tau_k]}ds + \int_0^t \eta_s\mathbf{1}_{[0,\tau_k]}d\langle B\rangle_s + \int_0^t \beta_s\mathbf{1}_{[0,\tau_k]}dB_s$$

Then we can apply Lemma 8.3.3 to $\Phi_k(s, X_{s\wedge\tau_k})$, $s\in[0, t]$, to obtain

$$\begin{aligned}\Phi(t, X_{t\wedge\tau_k}) - \Phi(0, X_0) &= \int_0^t \partial_x\Phi(u, X_u)\beta_u\mathbf{1}_{[0,\tau_k]}dB_u + \int_0^t [\partial_t\Phi(u, X_u) + \partial_x\Phi(u, X_u)\alpha_u]\mathbf{1}_{[0,\tau_k]}du \\ &\quad + \int_0^t [\partial_x\Phi(u, X_u)\eta_u\mathbf{1}_{[0,\tau_k]} + \tfrac{1}{2}\partial^2_{xx}\Phi(u, X_u)|\beta_u|^2\mathbf{1}_{[0,\tau_k]}]d\langle B\rangle_u.\end{aligned}$$

Letting $k \to \infty$ and noticing that X_t is continuous in t, we get the desired result. □

Example 8.3.5 For given $\varphi \in C^2(\mathbb{R})$, we have

$$\Phi(B_t) - \Phi(B_0) = \int_0^t \Phi_x(B_s)dB_s + \frac{1}{2}\int_0^t \Phi_{xx}(B_s)d\langle B\rangle_s .$$

This generalizes the previous results to more general situations.

Notes and Comments

The results in this chapter were mainly obtained by Li and Peng [110, 2011]. Li and Lin [109, 2013] found a point of incompleteness and proposed to use a more essential condition (namely, Condition (ii) in Definition 8.2.8) to replace the original one which was $\int_0^T |\eta_t|^p dt < \infty$, q.s.

A difficulty hidden behind is that the G-expectation theory is mainly based on the space of random variables $X = X(\omega)$ which are quasi-continuous with respect to the G-capacity $\hat{c}$. It is not yet clear that the martingale properties still hold for random variables without quasi-continuity condition.

There are still several interesting and fundamentally important issues on G-expectation theory and its applications. It is known that stopping times play a fundamental role in classical stochastic analysis. However, it is often nontrivial to directly apply stopping time techniques in a G-expectation space. The reason is that the stopped process may not belong to the class of processes which are meaningful in the G-framework. Song [160] considered the properties of hitting times for G- martingale and, moreover the stopped processes. He proved that the stopped processes for G-martingales are still G-martingales and that the hitting times for symmetric G-martingales with strictly increasing quadratic variation processes are quasi-continuous. Hu and Peng [82] introduced a suitable definition of stopping times and obtained the optional stopping theorem.

Appendix A
Preliminaries in Functional Analysis

A.1 Completion of Normed Linear Spaces

In this section, we suppose $\mathcal{H}$ is a linear space under the norm $\|\cdot\|$.

Definition A.1.1 A sequence of elements $\{x_n\} \in \mathcal{H}$ is called a **Cauchy sequence**, if $\{x_n\}$ satisfies **Cauchy's convergence condition**:

$$\lim_{n,m\to\infty} \|x_n - x_m\| = 0.$$

Definition A.1.2 A normed linear space $\mathcal{H}$ is called a **Banach space** if it is **complete**, i.e., if every Cauchy sequence $\{x_n\}$ of $\mathcal{H}$ converges strongly to an element x_∞ of $\mathcal{H}$:

$$\lim_{n\to\infty} \|x_n - x_\infty\| = 0.$$

Such a limit point x_∞, if exists, is uniquely determined because of the triangle inequality $\|x - x'\| \leq \|x - x_n\| + \|x_n - x'\|$.

The completeness of a Banach space plays an important role in functional analysis. We introduce the following theorem of completion.

Theorem A.1.3 *Let $\mathcal{H}$ be a normed linear space which is not complete. Then $\mathcal{H}$ is isomorphic and isometric to a dense linear subspace of a Banach-space $\tilde{\mathcal{H}}$, i.e., there exists a one-to-one correspondence $x \leftrightarrow \tilde{x}$ of $\mathcal{H}$ onto a dense linear subspace of $\tilde{\mathcal{H}}$ such that*

$$\widetilde{x+y} = \tilde{x} + \tilde{y},\ \widetilde{\alpha x} = \alpha\tilde{x},\ \|\tilde{x}\| = \|x\|.$$

The space $\tilde{\mathcal{H}}$ is uniquely determined up to isometric isomorphism.

For a proof see Yosida [178] (p. 56).

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7

A.2 The Hahn-Banach Extension Theorem

Definition A.2.1 Let T_1 and T_2 be two linear operators with domains $D(T_1)$ and $D(T_2)$ both contained in a linear space $\mathcal{H}$, and the ranges $R(T_1)$ and $R(T_2)$ both contained in a linear space $\mathcal{M}$. Then $T_1 = T_2$ if and only if $D(T_1) = D(T_2)$ and $T_1 x = T_2 x$ for all $x \in D(T_1)$. If $D(T_1) \subseteq D(T_2)$ and $T_1 x = T_2 x$ for all $x \in D(T_1)$, then T_2 is called an **extension** of T_1, or T_1 is called a **restriction** of T_2.

Theorem A.2.2 (Hahn-Banach extension theorem in real linear spaces) *Let $\mathcal{H}$ be a real linear space and let $p(x)$ be a real-valued function defined on $\mathcal{H}$ and satisfying the following conditions:*

$$p(x + y) \leq p(x) + p(y) \text{ (subadditivity)};$$
$$p(\alpha x) = \alpha p(x) \text{ for } \alpha \geq 0 \text{ (positive homogeneity)}.$$

Let $\mathcal{L}$ be a real linear subspace of $\mathcal{H}$ and f_0 a real-valued linear functional defined on $\mathcal{L}$, with the property:

$$f_0(\alpha x + \beta y) = \alpha f_0(x) + \beta f_0(y) \text{ for } x, y \in \mathcal{L} \text{ and } \alpha, \beta \in \mathbb{R}.$$

Let f_0 satisfy $f_0(x) \leq p(x)$ on $\mathcal{L}$. Then there exists a real-valued linear functional F defined on $\mathcal{H}$ such that:

(i) *F is an extension of f_0, i.e., $F(x) = f_0(x)$ for all $x \in \mathcal{L}$.*

(ii) *$F(x) \leq p(x)$ for $x \in \mathcal{H}$.*

For a proof see Yosida [178] (p. 102).

Theorem A.2.3 (Hahn-Banach extension theorem in normed linear spaces) *Let $\mathcal{H}$ be a normed linear space under the norm $\|\cdot\|$, $\mathcal{L}$ a linear subspace of $\mathcal{H}$ and f_1 a continuous linear functional defined on $\mathcal{L}$. Then there exists a continuous linear functional f, defined on $\mathcal{H}$, such that:*

(i) *f is an extension of f_1.*

(ii) $\|f_1\| = \|f\|$.

For a proof see Yosida [178] (p. 106).

A.3 Dini's Theorem and Tietze's Extension Theorem

Theorem A.3.1 (Dini's theorem) *Let $\mathcal{H}$ be a compact topological space. If a monotone sequence of bounded continuous functions converges pointwise to a continuous function, then it also converges uniformly.*

Theorem A.3.2 (Tietze's extension theorem) *Let $\mathcal{L}$ be a closed subset of a normed space $\mathcal{H}$ and let $f : \mathcal{L} \mapsto \mathbb{R}$ be a continuous function. Then there exists a continuous extension of f to all $\mathcal{H}$ with values in $\mathbb{R}$.*

Appendix B
Preliminaries in Probability Theory

B.1 Kolmogorov's Extension Theorem

Let a triple $(\Omega, \mathcal{F}, P)$ be a measurable space, in which Ω is an arbitrary set, $\mathcal{F}$ is a σ-algebra of Ω and P is a measure defined on $(\Omega, \mathcal{F})$. We are mainly concerned with probability measures, namely, $P(\Omega) = 1$.

Let X be an n-dimensional random variable, i.e., $X = X(\omega)$ is an $\mathcal{F}$-measurable function with values in $\mathbb{R}^n$ defined on $(\Omega$. Denote by $\mathcal{B}$ the Borel σ-algebra on $\mathbb{R}^n$. We define X's law of distribution P_X and its expectation E_P with respect to P as follows:

$$P_X(B) := P(\{\omega : X(\omega) \in B\}), \quad \text{for } B \in \mathcal{B} \text{ and } \quad E_P[X] := \int_{-\infty}^{+\infty} x P(dx).$$

In fact, we have $P_X(B) = E_P[\mathbf{1}_B(X)]$, $B \in \mathcal{B}$, $\mathbf{1}_B(\cdot)$ is the indicator of B.

Now let $\{X_t\}_{t \in T}$ be a stochastic process with values in $\mathbb{R}^n$ defined on a probability space $(\Omega, \mathcal{F}, P)$, where the parameter set T is usually the halfline $[0, +\infty)$.

Definition B.1.1 The **finite dimensional distributions** of the process $\{X_t\}_{t \in T}$ are the measures $\mu_{t_1, \cdots, t_k}$ defined on $\mathbb{R}^{n \times k}$, $k = 1, 2, \cdots$, by

$$\mu_{t_1, \cdots, t_k}(B_1 \times \cdots \times B_k) := P[X_{t_1} \in B_1, \cdots, X_{t_k} \in B_k], \; t_i \in T, \; i = 1, 2, \cdots, k,$$

where $B_i \in \mathcal{B}$, $i = 1, 2, \cdots, k$.

Many (but not all) important properties of the process $\{X_t\}_{t \in T}$ can be derived in terms of the family of all finite-dimensional distributions.

Conversely, given a family $\{\nu_{t_1, \cdots, t_k} : t_i \in T, i = 1, 2, \cdots, k, k \in \mathbb{N}\}$ of probability measures on $\mathbb{R}^{nk}$, it is important to be able to construct a stochastic process $(Y_t)_{t \in T}$ with $\nu_{t_1, \cdots, t_k}$ being its finite-dimensional distributions. The following famous theorem states that this can be done provided that $\{\nu_{t_1, \cdots, t_k}\}$ satisfy two natural consistency conditions.

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7

Theorem B.1.2 (Kolmogorov's extension theorem) *For all $t_1, t_2, \cdots, t_k$, $k \in \mathbb{N}$, let $\nu_{t_1,\cdots,t_k}$ be probability measures on $\mathbb{R}^{n\times k}$ such that*

$$\nu_{t_{\pi(1)},\cdots,t_{\pi(k)}}(B_1 \times \cdots \times B_k) = \nu_{t_1,\cdots,t_k}(B_{\pi^{-1}(1)} \times \cdots \times B_{\pi^{-1}(k)})$$

for all permutations π of the indices $\{1, 2, \cdots, k\}$ and

$$\nu_{t_1,\cdots,t_k}(B_1 \times \cdots \times B_k) = \nu_{t_1,\cdots,t_k,t_{k+1},\cdots,t_{k+m}}(B_1 \times \cdots \times B_k \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n)$$

for all $m \in \mathbb{N}$, where the set argument on the right hand side has totally $k+m$ factors.

Then there exists a probability space $(\Omega, \mathcal{F}, P)$ and a stochastic process $(X_t)_{t\in T}$ defined on Ω, $X_t : \Omega \mapsto \mathbb{R}^n$, such that

$$\nu_{t_1,\cdots,t_k}(B_1 \times \cdots \times B_k) = P(X_{t_1} \in B_1, \cdots, X_{t_k} \in B_k)$$

for all $t_i \in T$ and all Borel sets B_i, $i = 1, 2, \cdots, k$, $k \in \mathbb{N}$.

For a proof see Kolmogorov [103] (p. 29).

B.2 Kolmogorov's Criterion

Definition B.2.1 Suppose that $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$ are two stochastic processes defined on $(\Omega, \mathcal{F}, P)$. Then we say that X is a **version** (or modification) of Y, if

$$P(\{\omega : X_t(\omega) = Y_t(\omega)\}) = 1, \quad \text{for all } t \in T.$$

Theorem B.2.2 (Kolmogorov's continuity criterion) *Suppose that the process $X = \{X_t\}_{t\geq 0}$ satisfies the following condition: for all $T > 0$ there exist positive constants α, β, D such that*

$$E[|X_t - X_s|^\alpha] \leq D|t-s|^{1+\beta}, \quad 0 \leq s, t \leq T.$$

Then there exists a continuous version of X.

For a proof see Stroock and Varadhan [168] (p. 51).

Let E be a metric space and $\mathcal{B}$ be the Borel σ-algebra on E. We recall a few facts about the weak convergence of probability measures on $(E, \mathcal{B})$. If P is such a measure, we say that a subset A of E is a P-continuity set if $P(\partial A) = 0$, where ∂A is the boundary of A.

Proposition B.2.3 *For probability measures P_n, $n \in \mathbb{N}$, and P, the following conditions are equivalent:*

(i) *For every bounded continuous function f on E,*

$$\lim_{n\to\infty}\int f\,dP_n = \int f\,dP.$$

(ii) *For any bounded uniformly continuous function f on E,*

$$\lim_{n\to\infty}\int f\,dP_n = \int f\,dP.$$

(iii) *For any closed subset F of E,* $\limsup_{n\to\infty} P_n(F) \le P(F)$.
(iv) *For any open subset G of E,* $\liminf_{n\to\infty} P_n(G) \ge P(G)$.
(v) *For any P-continuity set A,* $\lim_{n\to\infty} P_n(A) = P(A)$.

Definition B.2.4 If P_n and P satisfy one of the equivalent conditions in Proposition B.2.3, we say that (P_n) **converges weakly** to P as $n \to \infty$.

Now let $\mathcal{P}$ be a family of probability measures on $(E, \mathcal{B})$.

Definition B.2.5 A family $\mathcal{P}$ is **weakly relatively compact** if every sequence of $\mathcal{P}$ contains a weakly convergent subsequence.

Definition B.2.6 A family $\mathcal{P}$ is **tight** if for any $\varepsilon \in (0, 1)$, there exists a compact set K_ε such that

$$P(K_\varepsilon) \ge 1 - \varepsilon \quad \text{for every} \quad P \in \mathcal{P}.$$

With this definition, we have the following theorem.

Theorem B.2.7 (Prokhorov's criterion) *If a family $\mathcal{P}$ is tight, then it is weakly relatively compact. If E is a Polish space (i.e., a separable completely metrizable topological space), then a weakly relatively compact family is tight.*

Definition B.2.8 If $(X_n)_{n\in\mathbb{N}}$ and X are random variables taking their values in a metric space E, we say that (X_n) **converges in distribution** or **converges in law** to X if their laws P_{X_n} converge weakly to the law P_X of X.

We stress on the fact that the random variables $(X_n)_{n\in\mathbb{N}}$ and X need not be defined on the same probability space.

Theorem B.2.9 (Kolmogorov's criterion for weak compactness) *Let $\{X^{(n)}\}$ be a sequence of $\mathbb{R}^d$-valued continuous processes defined on probability spaces $(\Omega^{(n)}, \mathcal{F}^{(n)}, P^{(n)})$ such that:*

(i) *The family $\{P^{(n)}_{X_0^{(n)}}\}$ of initial laws is tight in $\mathbb{R}^d$.*
(ii) *There exist three strictly positive constants α, β, γ such that for any $s, t \in \mathbb{R}_+$ and any n,*

$$E_{P^{(n)}}[|X_s^{(n)} - X_t^{(n)}|^\alpha] \le \beta |s - t|^{\gamma+1}.$$

Then the set $\{P^{(n)}_{X^{(n)}}\}$ of the laws of (X_n) is weakly relatively compact.

For the proof see Revuz and Yor [151] (p. 517)

B.3 Daniell-Stone Theorem

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, on which one can define integration. One essential property of the integration is its linearity, thus it can be seen as a linear functional on $L^1(\Omega, \mathcal{F}, \mu)$. This idea leads to another approach to define integral, this is the so-called Daniell's integral.

Definition B.3.1 Let Ω be an abstract set and $\mathcal{H}$ a linear space formed by a family of real valued functions. The space $\mathcal{H}$ is called a **vector lattice** if

$$f \in \mathcal{H} \Rightarrow |f| \in \mathcal{H},\ f \wedge 1 \in \mathcal{H}.$$

Definition B.3.2 Suppose that $\mathcal{H}$ is a vector lattice on Ω and I is a positive linear functional on $\mathcal{H}$, i.e.,

$$\begin{aligned} &f, g \in \mathcal{H}, \alpha, \beta \in \mathbb{R} \Rightarrow I(\alpha f + \beta g) = \alpha I(f) + \beta I(g); \\ &f \in \mathcal{H},\ f \geq 0 \Rightarrow I(f) \geq 0. \end{aligned}$$

If I satisfies the following condition:

$$f_n \in \mathcal{H},\ f_n \downarrow 0, \quad \text{as} \ \ n \to \infty \Rightarrow I(f_n) \to 0,$$

or equivalently,

$$f_n \in \mathcal{H},\ f_n \uparrow f \in \mathcal{H} \Rightarrow I(f) = \lim_{n\to\infty} I(f_n),$$

then I is called a **Daniell's integral** on $\mathcal{H}$.

Theorem B.3.3 (Daniell-Stone theorem) *Suppose that $\mathcal{H}$ is a vector lattice on Ω and I a Daniell's integral on $\mathcal{H}$. Then there exists a measure μ on $\mathcal{F}$, where $\mathcal{F} := \sigma(f : f \in \mathcal{H})$, such that $\mathcal{H} \subset L^1(\Omega, \mathcal{F}, \mu)$ and $I(f) = \mu(f)$, $\forall f \in \mathcal{H}$. Furthermore, if $1 \in \mathcal{H}_+^*$, where $\mathcal{H}_+^* := \{f : \exists f_n \geq 0,\ f_n \in \mathcal{H}$ such that $f_n \uparrow f\}$, then this measure μ is unique and is σ-finite.*

For the proof see Ash [4], Dellacherie and Meyer [43] (p. 59), Dudley [55] (p. 142), or Yan [176] (p. 74).

B.4 Some Important Inequalities

Let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, P)$ be a probability space. A process $\{M_t\}_{t\geq 0}$ is called a martingale if $M_t \in L^1(\mathcal{F}_t)$ for any t and $E_P[M_t|\mathcal{F}_s] = M_s$ for all $s \leq t$.

Theorem B.4.1 (Doob's maximal inequality) *Let $\{M_t\}_{t\geq 0}$ be a right-continuous martingale. Then for any $p > 1$,*

$$\left(E_P[\sup_t |M_t|^p]\right)^{1/p} \le \frac{p}{p-1} \sup_t (E_P[|M_t|^p])^{1/p}.$$

For the proof see e.g., Revuz and Yor [151] (p. 54).

Theorem B.4.2 (Burkholder-Davis-Gundy inequality) *For any* $p > 0$*, there exists a constant* $c_p > 0$ *such that, for all continuous martingale* $\{M_t\}_{t\ge 0}$ *vanishing at zero,*

$$c_p E_P[\langle M\rangle_\infty^{p/2}] \le E_P\left[\sup_t |M_t|^p\right] \le \frac{1}{c_p} E_P[\langle M\rangle_\infty^{p/2}],$$

where $\langle M\rangle$ *is the quadratic variation process of* M.

For the proof see Revuz and Yor [151] (p. 160).

Appendix C
Solutions of Parabolic Partial Differential Equation

C.1 The Definition of Viscosity Solutions

The notion of viscosity solutions was firstly introduced by Crandall and Lions [37, 38] (see also Evans's contribution [61, 62]) for the first-order Hamilton-Jacobi equation, with uniqueness proof given in [38]. The proof for second-order Hamilton-Jacobi-Bellman equations was firstly developed by Lions [112, 113] using stochastic control verification arguments. A breakthrough was achieved in the second-order PDE theory by Jensen [94]. For all other important contributions in the developments of this theory we refer to the well-known user's guide by Crandall et al. [39]. For reader's convenience, we systematically interpret some parts of [39] required in this book into it's parabolic version. However, up to my knowledge, the presentation and the related proof for the domination theorems seems to be a new generalization of the maximum principle presented in [39]. Books on this theory are, among others, Barles [9], Fleming, and Soner [66], Yong and Zhou [177].

Let $T > 0$ be fixed and let $\mathcal{O} \subset [0, T] \times \mathbb{R}^d$. We set

$$USC(\mathcal{O}) = \{\text{upper semicontinuous functions } u : \mathcal{O} \mapsto \mathbb{R}\},$$

$$LSC(\mathcal{O}) = \{\text{lower semicontinuous functions } u : \mathcal{O} \mapsto \mathbb{R}\}.$$

Consider the following parabolic PDE:

$$\begin{cases} \text{(E) } \partial_t u - G(t, x, u, Du, D^2 u) = 0 \text{ on } (0, T) \times \mathbb{R}^d, \\ \text{(IC) } u(0, x) = \varphi(x) \text{ for } x \in \mathbb{R}^d, \end{cases} \tag{C.1.1}$$

where $G : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$, $\varphi \in C(\mathbb{R}^d)$. We always suppose that the function G is continuous and satisfies the following degenerate elliptic condition:

$$G(t, x, r, p, X) \geq G(t, x, r, p, Y) \text{ whenever } X \geq Y. \tag{C.1.2}$$

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7

Next we recall the definition of viscosity solutions from Crandall et al. [39]. Let $u : (0, T) \times \mathbb{R}^d \to \mathbb{R}$ and $(t, x) \in (0, T) \times \mathbb{R}^d$. We denote by $\mathcal{P}^{2,+}u(t, x)$ (the "**parabolic superjet**" of u at (t, x)) the set of triplets $(a, p, X) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d)$ such that

$$u(s, y) \leq u(t, x) + a(s - t) + \langle p, y - x\rangle \\ + \frac{1}{2}\langle X(y - x), y - x\rangle + o(|s - t| + |y - x|^2).$$

We define

$$\begin{aligned}\bar{\mathcal{P}}^{2,+}u(t, x) :=\{&(a, p, X) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d) : \exists(t_n, x_n, a_n, p_n, X_n) \\ &\text{such that } (a_n, p_n, X_n) \in \mathcal{P}^{2,+}u(t_n, x_n) \text{ and} \\ &(t_n, x_n, u(t_n, x_n), a_n, p_n, X_n) \to (t, x, u(t, x), a, p, X)\}.\end{aligned}$$

Similarly, we define $\mathcal{P}^{2,-}u(t, x)$ (the "**parabolic subjet**" of u at (t, x)) by $\mathcal{P}^{2,-}u(t, x) := -\mathcal{P}^{2,+}(-u)(t, x)$ and $\bar{\mathcal{P}}^{2,-}u(t, x)$ by $\bar{\mathcal{P}}^{2,-}u(t, x) := -\bar{\mathcal{P}}^{2,+}(-u)(t, x)$.

Definition C.1.1 (i) A **viscosity subsolution** of (E) on $(0, T) \times \mathbb{R}^d$ is a function $u \in USC((0, T) \times \mathbb{R}^d)$ such that for any $(t, x) \in (0, T) \times \mathbb{R}^d$,

$$a - G(t, x, u(t, x), p, X) \leq 0, \quad \text{for} \quad (a, p, X) \in \mathcal{P}^{2,+}u(t, x).$$

Likewise, a **viscosity supersolution** of (E) on $(0, T) \times \mathbb{R}^d$ is a function $v \in LSC((0, T) \times \mathbb{R}^d)$ such that for any $(t, x) \in (0, T) \times \mathbb{R}^d$,

$$a - G(t, x, v(t, x), p, X) \geq 0 \ \text{ for } (a, p, X) \in \mathcal{P}^{2,-}v(t, x).$$

A **viscosity solution** of (E) on $(0, T) \times \mathbb{R}^d$ is a function that is simultaneously a viscosity subsolution and a viscosity supersolution of (E) on $(0, T) \times \mathbb{R}^d$.

(ii) A function $u \in USC([0, T) \times \mathbb{R}^d)$ is called a **viscosity subsolution** of (C.1.1) on $[0, T) \times \mathbb{R}^d$ if u is a viscosity subsolution of (E) on $(0, T) \times \mathbb{R}^d$ and $u(0, x) \leq \varphi(x)$ for $x \in \mathbb{R}^d$; the corresponding notions of viscosity supersolution and viscosity solution of (C.1.1) on $[0, T) \times \mathbb{R}^d$ are then obvious.

We now give the following equivalent definition (see Crandall et al. [39]).

Definition C.1.2 A viscosity subsolution of (E), or G-subsolution, on $(0, T) \times \mathbb{R}^d$ is a function $u \in USC((0, T) \times \mathbb{R}^d)$ such that for all $(t, x) \in (0, T) \times \mathbb{R}^d$, $\phi \in C^2((0, T) \times \mathbb{R}^d)$ such that $u(t, x) = \phi(t, x)$ and $u < \phi$ on $(0, T) \times \mathbb{R}^d \backslash (t, x)$, we have

$$\partial_t \phi(t, x) - G(t, x, \phi(t, x), D\phi(t, x), D^2\phi(t, x)) \leq 0.$$

Likewise, a viscosity supersolution of (E), or G-supersolution, on $(0, T) \times \mathbb{R}^d$ is a function $v \in LSC((0, T) \times \mathbb{R}^d)$ such that for all $(t, x) \in (0, T) \times \mathbb{R}^d$, $\phi \in$

$C^2((0,T)\times\mathbb{R}^d)$ such that $u(t,x)=\phi(t,x)$ and $u>\phi$ on $(0,T)\times\mathbb{R}^d\backslash(t,x)$, we have

$$\partial_t\phi(t,x)-G(t,x,\phi(t,x),D\phi(t,x),D^2\phi(t,x))\geq 0.$$

Finally, a viscosity solution of (E) on $(0,T)\times\mathbb{R}^d$ is a function that is simultaneously a viscosity subsolution and a viscosity supersolution of (E) on $(0,T)\times\mathbb{R}^d$. The definition of a viscosity solution of (C.1.1) on $[0,T)\times\mathbb{R}^d$ is the same as in the above definition.

C.2 Comparison Theorem

We will use the following well-known result in viscosity solution theory (see Theorem 8.3 of Crandall et al. [39]).

Theorem C.2.1 *Let $u_i\in USC((0,T)\times\mathbb{R}^{d_i})$ for $i=1,\cdots,k$. Let φ be a function defined on $(0,T)\times\mathbb{R}^{d_1+\cdots+d_k}$ such that $(t,x_1,\ldots,x_k)\mapsto\varphi(t,x_1,\ldots,x_k)$ is once continuously differentiable in t and twice continuously differentiable in $(x_1,\cdots,x_k)\in\mathbb{R}^{d_1+\cdots+d_k}$. Suppose that $\hat{t}\in(0,T)$, $\hat{x}_i\in\mathbb{R}^{d_i}$ for $i=1,\cdots,k$ and*

$$\begin{aligned}w(t,x_1,\cdots,x_k)&:=u_1(t,x_1)+\cdots+u_k(t,x_k)-\varphi(t,x_1,\cdots,x_k)\\&\leq w(\hat{t},\hat{x}_1,\cdots,\hat{x}_k)\end{aligned}$$

for $t\in(0,T)$ and $x_i\in\mathbb{R}^{d_i}$. Assume, moreover, that there exists a constant $r>0$ such that for every $M>0$ there exists a constant C such that for $i=1,\cdots,k$,

$$\begin{aligned}&b_i\leq C,\quad \text{whenever}\quad (b_i,q_i,X_i)\in\mathcal{P}^{2,+}u_i(t,x_i),\\&|x_i-\hat{x}_i|+|t-\hat{t}|\leq r,\quad \text{and}\quad |u_i(t,x_i)|+|q_i|+\|X_i\|\leq M.\end{aligned}\tag{C.2.1}$$

Then for any $\varepsilon>0$, there exist $X_i\in\mathbb{S}(d_i)$ such that:
(i) $(b_i,D_{x_i}\varphi(\hat{t},\hat{x}_1,\cdots,\hat{x}_k),X_i)\in\overline{\mathcal{P}}^{2,+}u_i(\hat{t},\hat{x}_i),\quad i=1,\cdots,k;$
(ii)

$$-(\frac{1}{\varepsilon}+\|A\|)I\leq\begin{bmatrix}X_1&\cdots&0\\\vdots&\ddots&\vdots\\0&\cdots&X_k\end{bmatrix}\leq A+\varepsilon A^2;$$

(iii) $b_1+\cdots+b_k=\partial_t\varphi(\hat{t},\hat{x}_1,\cdots,\hat{x}_k)$, *where* $A=D_x^2\varphi(\hat{t},\hat{x})\in\mathbb{S}(d_1+\cdots+d_k)$.

Observe that the above condition (C.2.1) will be guaranteed if having each u_i being a subsolution of a parabolic equation as given in two theorems, see below.

In this section we will give comparison theorem for G-solutions with different functions G_i.

(G). We assume that

$$G_i : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}, \quad i = 1, \cdots, k,$$

is continuous in the following sense: for any $t \in [0, T)$, $v \in \mathbb{R}$, x, y, $p \in \mathbb{R}^d$ and $X \in \mathbb{S}(d)$,

$$\begin{aligned} &|G_i(t, x, v, p, X) - G_i(t, y, v, p, X)| \\ &\leq \bar{\omega}\left(1 + (T-t)^{-1} + |x| + |y| + |v|\right) \cdot \omega\left(|x-y| + |p| \cdot |x-y|\right), \end{aligned}$$

where $\omega, \bar{\omega} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ are given continuous functions with $\omega(0) = 0$.

Theorem C.2.2 (Domination Theorem) *We are given constants* $\beta_i > 0$, $i=1, \cdots, k$. *Let* $u_i \in USC([0, T] \times \mathbb{R}^d)$ *be subsolutions of*

$$\partial_t u - G_i(t, x, u, Du, D^2u) = 0, \quad i = 1, \cdots, k, \tag{C.2.2}$$

on $(0, T) \times \mathbb{R}^d$ *such that* $\left(\sum_{i=1}^k \beta_i u_i(t, x)\right)^+ \to 0$, *uniformly as* $|x| \to \infty$. *We assume that the functions* $\{G_i\}_{i=1}^k$ *satisfy assumption (G) and that the following domination condition hold:*

$$\sum_{i=1}^k \beta_i G_i(t, x, v_i, p_i, X_i) \leq 0, \tag{C.2.3}$$

for any $(t, x) \in (0, T) \times \mathbb{R}^d$ *and* $(v_i, p_i, X_i) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d)$ *such that* $\sum_{i=1}^k \beta_i v_i \geq 0$, $\sum_{i=1}^k \beta_i p_i = 0$, $\sum_{i=1}^k \beta_i X_i \leq 0$.

Then a domination also holds for the solutions: if the sum of initial values $\sum_{i=1}^k \beta_i u_i(0, \cdot)$ *is a non-positive function on* $\mathbb{R}^d$, *then* $\sum_{i=1}^k \beta_i u_i(t, \cdot) \leq 0$, *for all* $t > 0$.

Proof We first observe that for $\bar{\delta} > 0$ and for any $1 \leq i \leq k$, the function defined by $\tilde{u}_i := u_i - \bar{\delta}/(T-t)$ is a subsolution of

$$\partial_t \tilde{u}_i - \tilde{G}_i(t, x, \tilde{u}_i, D\tilde{u}_i, D^2\tilde{u}_i) \leq -\frac{\bar{\delta}}{(T-t)^2}.$$

Here $\tilde{G}_i(t, x, v, p, X) := G_i(t, x, v + \bar{\delta}/(T-t), p, X)$. It is easy to check that the functions $\tilde{G}_i$ satisfy the same conditions as G_i. Since $\sum_{i=1}^k \beta_i u_i \leq 0$ follows from $\sum_{i=1}^k \beta_i \tilde{u}_i \leq 0$ in the limit $\bar{\delta} \downarrow 0$, it suffices to prove the theorem under the additional assumptions:

$$\begin{aligned} \partial_t u_i - G_i(t, x, u_i, Du_i, D^2u_i) \leq -c, \quad \text{where } c = \bar{\delta}/T^2, \\ \text{and } \lim_{t \to T} u_i(t, x) = -\infty \text{ uniformly on } \mathbb{R}^d. \end{aligned} \tag{C.2.4}$$

To prove the theorem, we assume to the contrary that

$$\sup_{(t,x)\in[0,T)\times\mathbb{R}^d} \sum_{i=1}^{k} \beta_i u_i(t,x) = m_0 > 0.$$

We will apply Theorem C.2.1 for $x = (x_1, \cdots, x_k) \in \mathbb{R}^{k\times d}$ and

$$w(t,x) := \sum_{i=1}^{k} \beta_i u_i(t,x_i), \quad \varphi(x) = \varphi_\alpha(x) := \frac{\alpha}{2}\sum_{i=1}^{k-1} |x_{i+1} - x_i|^2.$$

For any large $\alpha > 0$, the maximum of $w - \varphi_\alpha$ is achieved at some point (t^α, x^α) inside a compact subset of $[0, T) \times \mathbb{R}^{k\times d}$. Indeed, since

$$M_\alpha = \sum_{i=1}^{k} \beta_i u_i(t^\alpha, x_i^\alpha) - \varphi_\alpha(x^\alpha) \geq m_0,$$

we conclude that t^α must be inside an interval $[0, T_0]$, $T_0 < T$ and x^α must be inside a compact set $\{x \in \mathbb{R}^{k\times d} : \sup_{t\in[0,T_0]} w(t,x) \geq \frac{m_0}{2}\}$. We can check that (see [39] Lemma 3.1)

$$\begin{cases} \text{(i)} \ \lim_{\alpha\to\infty} \varphi_\alpha(x^\alpha) = 0, \\ \text{(ii)} \ \lim_{\alpha\to\infty} M_\alpha = \lim_{\alpha\to\infty} \beta_1 u_1(t^\alpha, x_1^\alpha) + \cdots + \beta_k u_k(t^\alpha, x_k^\alpha)) \\ \qquad\qquad = \sup_{(t,x)\in[0,T)\times\mathbb{R}^d} [\beta_1 u_1(t,x) + \cdots + \beta_k u_k(t,x)] \\ \qquad\qquad = [\beta_1 u_1(\hat{t}, \hat{x}) + \cdots + \beta_k u_k(\hat{t}, \hat{x})] = m_0, \end{cases} \tag{C.2.5}$$

where $(\hat{t}, \hat{x})$ is a limit point of (t^α, x^α). Since $u_i \in \text{USC}$, for sufficiently large α, we have

$$\beta_1 u_1(t^\alpha, x_1^\alpha) + \cdots + \beta_k u_k(t^\alpha, x_k^\alpha) \geq \frac{m_0}{2}.$$

If $\hat{t} = 0$, we have $\limsup_{\alpha\to\infty} \sum_{i=1}^{k} \beta_i u_i(t^\alpha, x_i^\alpha) = \sum_{i=1}^{k} \beta_i u_i(0, \hat{x}) \leq 0$. We know that $\hat{t} > 0$ and thus t^α must be strictly positive for large α. It follows from Theorem C.2.1 that, for any $\varepsilon > 0$ there exist $b_i^\alpha \in \mathbb{R}$, $X_i \in \mathbb{S}(d)$ such that

$$(b_i^\alpha, \beta_i^{-1} D_{x_i}\varphi(x^\alpha), X_i) \in \bar{\mathcal{P}}^{2,+} u_i(t^\alpha, x_i^\alpha), \ \sum_{i=1}^{k} \beta_i b_i^\alpha = 0 \text{ for } i = 1, \cdots, k, \tag{C.2.6}$$

and

$$-(\frac{1}{\varepsilon} + \|A\|)I \leq \begin{pmatrix} \beta_1 X_1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \beta_{k-1}X_{k-1} & 0 \\ 0 & \ldots & 0 & \beta_k X_k \end{pmatrix} \leq A + \varepsilon A^2, \tag{C.2.7}$$

where $A = D^2\varphi_\alpha(x^\alpha) \in \mathbb{S}(k \times d)$ is explicitly given by

$$A = \alpha J_{kd}, \text{ where } J_{kd} = \begin{pmatrix} I_d & -I_d & \cdots & \cdots & 0 \\ -I_d & 2I_d & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 2I_d & -I_d \\ 0 & \cdots & \cdots & -I_d & I_d \end{pmatrix}.$$

The second inequality in (C.2.7) implies that $\sum_{i=1}^k \beta_i X_i \leq 0$. Set

$$\begin{aligned} p_1^\alpha &= \beta_1^{-1} D_{x_1}\varphi_\alpha(x^\alpha) = \beta_1^{-1}\alpha(x_1^\alpha - x_2^\alpha), \\ p_2^\alpha &= \beta_2^{-1} D_{x_2}\varphi_\alpha(x^\alpha) = \beta_2^{-1}\alpha(2x_2^\alpha - x_1^\alpha - x_3^\alpha), \\ &\vdots \\ p_{k-1}^\alpha &= \beta_{k-1}^{-1} D_{x_{k-1}}\varphi_\alpha(x^\alpha) = \beta_{k-1}^{-1}\alpha(2x_{k-1}^\alpha - x_{k-2}^\alpha - x_k^\alpha), \\ p_k^\alpha &= \beta_k^{-1} D_{x_k}\varphi_\alpha(x^\alpha) = \beta_k^{-1}\alpha(x_k^\alpha - x_{k-1}^\alpha). \end{aligned}$$

Thus $\sum_{i=1}^k \beta_i p_i^\alpha = 0$. From this together with (C.2.6) and (C.2.4), it follows that

$$b_i^\alpha - G_i(t^\alpha, x_i^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) \leq -c, \quad i = 1, \cdots, k.$$

By (C.2.5) (i), we also have $\lim_{\alpha\to\infty} |p_i^\alpha| \cdot |x_i^\alpha - x_1^\alpha| \to 0$. This, together with the domination condition (C.2.3) for G_i, imply that

$$\begin{aligned} -c\sum_{i=1}^k \beta_i &= -\sum_{i=1}^k \beta_i b_i^\alpha - c\sum_{i=1}^k \beta_i \geq -\sum_{i=1}^k \beta_i G_i(t^\alpha, x_i^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) \\ &\geq -\sum_{i=1}^k \beta_i G_i(t^\alpha, x_1^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) \\ &\quad - \sum_{i=1}^k \beta_i |G_i(t^\alpha, x_i^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i) - G_i(t^\alpha, x_1^\alpha, u_i(t^\alpha, x_i^\alpha), p_i^\alpha, X_i)| \\ &\geq -\sum_{i=1}^k \beta_i \bar{\omega}\left(1 + (T - T_0)^{-1} + |x_1^\alpha| + |x_i^\alpha| + |u_i(t^\alpha, x_i^\alpha)|\right) \\ &\quad \cdot \omega\left(|x_i^\alpha - x_1^\alpha| + |p_i^\alpha| \cdot |x_i^\alpha - x_1^\alpha|\right). \end{aligned}$$

Notice now that the right side tends to zero as $\alpha \to \infty$, which leads to a contradiction. The proof is complete. □

Theorem C.2.3 *Given a family of functions* $G_i = G_i(t, x, v, p, X)$, $i = 1, \cdots, k+1$ *for which assumption* **(G)** *holds. Assume further that* G_i, $i = 1, \cdots, k+1$, *for any* $(t, x) \in [0, \infty) \times \mathbb{R}^d$ *and* $(v_i, p_i, X_i) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d)$, *satisfies the following conditions:*

(i) positive homogeneity:

$$G_i(t, x, \lambda v, \lambda p, \lambda X) = \lambda G_i(t, x, v, p, X), \quad \textit{for all}\, \lambda \geq 0;$$

(ii) there exists a constant $\bar{C}$ *such that*

$$|G_i(t, x, v_1, p_1, X_1) - G_i(t, x, v_2, p_2, X_2)| \leq \bar{C}\,(|v_1 - v_2| + |p_1 - p_2| + |X_1 - X_2|)\,;$$

(iii) G_{k+1} *dominates* $\{G_i\}_{i=1}^k$ *in the following sense:*

$$\sum_{i=1}^k G_i(t, x, v_i, p_i, X_i) \leq G_{k+1}\left(t, x, \sum_{i=1}^k v_i, \sum_{i=1}^k p_i, \sum_{i=1}^k X_i\right). \tag{C.2.8}$$

Let $u_i \in USC\,([0, T] \times \mathbb{R}^d)$ *be a* G_i*-subsolution and* $u \in LSC\,([0, T] \times \mathbb{R}^d)$ *a* G*-supersolution such that* u_i *and* u *satisfy polynomial growth condition. Then* $\sum_{i=1}^k u_i(t, x) \leq u(t, x)$ *on* $[0, T) \times \mathbb{R}^d$ *provided that* $\sum_{i=1}^k u_i|_{t=0} \leq u|_{t=0}$.

Proof For a fixed and large constant $\lambda > C_1$, where C_1 is a positive constant determined later, we set $\xi(x) := (1 + |x|^2)^{l/2}$ and

$$\tilde{u}_i(t, x) := u_i(t, x)\xi^{-1}(x)e^{-\lambda t},\ i = 1, \cdots, k,\ \ \tilde{u}_{k+1}(t, x) := -u(t, x)\xi^{-1}(x)e^{-\lambda t}.$$

The constant λ is chosen large enough so that $\sum |\tilde{u}_i(t, x)| \to 0$ uniformly as $|x| \to \infty$. It is easy to check that, for any $i = 1, \cdots, k+1$, $\tilde{u}_i$ is a subsolution of the equation

$$\partial_t \tilde{u}_i + \lambda \tilde{u}_i - \tilde{G}_i(t, x, \tilde{u}_i, D\tilde{u}_i, D^2\tilde{u}_i) = 0,$$

where, for any $i = 1, \cdots, k$, the function $\tilde{G}_i(t, x, v, p, X)$ is given by

$$e^{-\lambda t}\xi^{-1} G_i(t, x, e^{\lambda t} v\xi, e^{\lambda t}(p\xi(x) + v\eta(x)), e^{\lambda t}(X\xi + p \otimes \eta(x) + \eta(x) \otimes p + v\kappa(x))),$$

while $\tilde{G}_{k+1}(t, x, v, p, X)$ is given by

$$-e^{-\lambda t}\xi^{-1} G_{k+1}(t, x, -e^{\lambda t} v\xi, -e^{\lambda t}(p\xi(x) + v\eta(x)), -e^{\lambda t}(X\xi + p \otimes \eta(x) + \eta(x) \otimes p + v\kappa(x))).$$

Here

$$\eta(x) := \xi^{-1}(x)D\xi(x) = l(1 + |x|^2)^{-1}x,$$
$$\kappa(x) := \xi^{-1}(x)D^2\xi(x) = l(1 + |x|^2)^{-1}I + l(l-2)(1 + |x|^2)^{-2}x \otimes x.$$

Note that both $\eta(x)$ and $\kappa(x)$ are uniformly bounded Lipschitz functions. Then one can easily check that $\tilde{G}_i$ satisfies assumption **(G)** by conditions (i) and (ii).

From the domination condition (C.2.8), for any $(v_i, p_i, X_i) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d)$, $i = 1, \cdots, k+1$, such that $\sum_{i=1}^{k+1} v_i = 0$, $\sum_{i=1}^{k+1} p_i = 0$, and $\sum_{i=1}^{k+1} X_i = 0$, we have

$$\sum_{i=1}^{k+1} \tilde{G}_i(t, x, v_i, p_i, X_i) \leq 0.$$

For $v, r \in \mathbb{R}$, $p \in \mathbb{R}^d$ and $X, Y \in \mathbb{S}(d)$ such that $Y \geq 0$ and $r > 0$, since $\tilde{G}$ is still monotone in X, by condition (ii), we get

$$\begin{aligned}
&\tilde{G}_{k+1}(t, x, v, p, X) - \tilde{G}_{k+1}(t, x, v - r, p, X + Y) \\
&\leq \tilde{G}_{k+1}(t, x, v, p, X) - \tilde{G}_{k+1}(t, x, v - r, p, X) \\
&\leq C_1 r,
\end{aligned}$$

where the constant C_1 depends only on $\bar{C}$. We then apply the above theorem by choosing $\beta_i = 1, i = 1, \cdots, k+1$. Thus $\sum_{i=1}^{k+1} \tilde{u}_i|_{t=0} \leq 0$. Moreover for any $v_i \in \mathbb{R}$, $p_i \in \mathbb{R}^d$ and $X_i \in \mathbb{S}(d)$ such that $\hat{v} = \sum_{i=1}^{k+1} v_i \geq 0$, $\sum_{i=1}^{k+1} p_i = 0$ and $\hat{X} = \sum_{i=1}^{k+1} X_i \leq 0$, we have

$$\begin{aligned}
&-\lambda \sum_{i=1}^{k+1} v_i + \sum_{i=1}^{k+1} \tilde{G}_i(t, x, v_i, p_i, X_i) \\
&= -\lambda \hat{v} + \sum_{i=1}^{k} \tilde{G}_i(t, x, v_i, p_i, X_i) + \tilde{G}_{k+1}(t, x, v_{k+1} - \hat{v}, p_{k+1}, X_{k+1} - \hat{X}) \\
&\quad + \tilde{G}_{k+1}(t, x, v_{k+1}, p_{k+1}, X_{k+1}) - \tilde{G}_{k+1}(t, x, v_{k+1} - \hat{v}, p_{k+1}, X_{k+1} - \hat{X}) \\
&\leq -\lambda \hat{v} + \tilde{G}_{k+1}(t, x, v_{k+1}, p_{k+1}, X_{k+1}) - \tilde{G}_{k+1}(t, x, v_{k+1} - \hat{v}, p_{k+1}, X_{k+1}) \\
&\leq -\lambda \hat{v} + C_1 \hat{v} \leq 0.
\end{aligned}$$

It follows that all conditions in Theorem C.2.2 are satisfied. Thus we have $\sum_{i=1}^{k+1} \tilde{u}_i \leq 0$, or equivalently, $\sum_{i=1}^{k+1} u_i(t, x) \leq u(t, x)$ for $(t, x) \in [0, T) \times \mathbb{R}^d$. □

Remark C.2.4 We can replace the above Condition (i) by the following Condition **(U)**: $u_i, i = 1, \ldots, k$ and u satisfy:

$$|v(t, x)| \to 0, \text{ uniformly (in } t) \text{ as } x \to \infty.$$

The above theorem still holds true. Its proof is quite similar.

The following comparison theorem is a direct consequence of the above Domination Theorem C.2.3.

Theorem C.2.5 (Comparison Theorem) *Given two functions $G = G(t, x, v, p, X)$ and $G_1 = G_1(t, x, v, p, X)$ satisfying condition (**G**) and conditions (i), (ii) in Theorem C.2.3. We also assume that, for any $(t, x, v, p, X) \in [0, \infty) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$,*

$$G(t, x, v, p, X) \geq G_1(t, x, v, p, X). \tag{C.2.9}$$

Let $u_1 \in USC([0, T] \times \mathbb{R}^d)$ be a G_1-subsolution and $u \in LSC([0, T] \times \mathbb{R}^d)$ a G-supersolution on $(0, T) \times \mathbb{R}^d$ satisfying the polynomial growth condition. Then $u \geq u_1$ on $[0, T) \times \mathbb{R}^d$ provided that $u|_{t=0} \geq u_1|_{t=0}$. In particular this comparison holds in the case when $G \equiv G_1$.

The following special case of Theorem C.2.3 is also very useful.

Theorem C.2.6 (Domination Theorem) *We assume that G_1 and G satisfy the same conditions given in Theorem C.2.5 except that the condition (C.2.9) is replaced by the following one: for any $(t, x) \in [0, \infty) \times \mathbb{R}^d$ and (v, p, X), $(v', p', X') \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d)$,*

$$G_1(t, x, v, p, X) - G_1(t, x, v', p', X') \leq G(t, x, v - v', p - p', X - X')$$

Let $u \in USC([0, T] \times \mathbb{R}^d)$ be a G_1-subsolution and $v \in LSC([0, T] \times \mathbb{R}^d)$ a G_1-supersolution on $(0, T) \times \mathbb{R}^d$ and w is a G-supersolution. They all satisfy the polynomial growth condition. If $(u - v)|_{t=0} = w|_{t=0}$, then $u - v \leq w$ on $[0, T) \times \mathbb{R}^d$.

Remark C.2.7 According to Remark C.2.4, Theorems C.2.5 and C.2.6 hold still true if we replace Condition (i) in Theorem C.2.3 by Condition (**U**).

The following theorem is frequently used in this book. Let $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a given continuous sublinear function, monotone in $A \in \mathbb{S}(d)$. Obviously, G satisfies the conditions in Theorem C.2.3. We consider the following G-equation:

$$\partial_t u - G(Du, D^2 u) = 0, \quad u(0, x) = \varphi(x). \tag{C.2.10}$$

Theorem C.2.8 *Let $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a given continuous sublinear function, monotone in $A \in \mathbb{S}(d)$. Then we have two statements:*

(i) *If $u \in \mathrm{USC}([0, T] \times \mathbb{R}^d)$ is a viscosity subsolution of* (C.2.10) *and $v \in \mathrm{LSC}([0, T] \times \mathbb{R}^d)$ is a viscosity supersolution of* (C.2.10)*, and both u and v have polynomial growth, then $u \leq v$;*
(ii) *If $u^\varphi \in C([0, T] \times \mathbb{R}^d)$ denotes the solution of* (C.2.10) *with initial condition φ and u^φ has polynomial growth, then $u^{\varphi+\psi} \leq u^\varphi + u^\psi$, and, for any $\lambda \geq 0$, $u^{\lambda\varphi} = \lambda u^\varphi$.*

Proof By the above theorems, it is easy to obtain the results. □

In fact, the assumption (**G**) is not necessary for the validity of the Comparison Theorem. To see this we consider the following condition instead of (**G**):

(G') Assume

$$G : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R},$$

is continuous in the following sense: for any $t \in [0, T)$, $v \in \mathbb{R}$, $x, y, p \in \mathbb{R}^d$, and $X, Y \in \mathbb{S}(d)$ with

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq 3\alpha \begin{pmatrix} I_d & -I_d \\ -I_d & I_d \end{pmatrix}, \tag{C.2.11}$$

we have

$$\begin{aligned} & G(t, x, v, \alpha(x - y), X) - G(t, y, v, \alpha(x - y), -Y) \\ & \leq \bar{\omega}(1 + (T - t)^{-1} + |x| + |y| + |v|)\omega(|x - y| + \alpha|x - y|^2), \end{aligned}$$

where $\omega, \bar{\omega} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ are given continuous functions with $\omega(0) = 0$. Moreover, $G(t, x, v, p, A)$ is non-increasing in v.

Taking $k = 2$, $\beta_1 = \beta_2 = 1$ and $G_1 = G$, $G_2(t, x, v, p, A) = -G(t, x, -v, -p, -A)$, we have that Theorem C.2.2 holds. Indeed, it follows from the proof of Theorem C.2.2, line by line, that there exist $b_i^\alpha \in \mathbb{R}$, $X_i \in \mathbb{S}(d)$ such that ($\varepsilon = 1/\alpha$)

$$(b_i^\alpha, D_{x_i}\varphi(x^\alpha), X_i) \in \bar{\mathcal{P}}^{2,+}u_i(t^\alpha, x_i^\alpha), \text{ for } i = 1, 2, \tag{C.2.12}$$

and

$$\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \leq 3\alpha \begin{pmatrix} I_d & -I_d \\ -I_d & I_d \end{pmatrix}.$$

Then we obtain the following chain of relations:

$$\begin{aligned} & -2c \\ & \geq G(t^\alpha, x_2^\alpha, -u_2(t^\alpha, x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2) - G(t^\alpha, x_1^\alpha, u_1(t^\alpha, x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1) \\ & \geq G(t^\alpha, x_2^\alpha, u_1(t^\alpha, x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2) - G(t^\alpha, x_1^\alpha, u_1(t^\alpha, x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1) \\ & \geq -\bar{\omega}(1 + (T - T_0)^{-1} + |x_1^\alpha| + |x_2^\alpha| + |u_1(t^\alpha, x_1^\alpha)|) \cdot \omega(|x_2^\alpha - x_1^\alpha| + \alpha|x_2^\alpha - x_1^\alpha|^2), \end{aligned}$$

which implies the desired result. Thus we can also establish the Comparison Theorem C.2.5 under assumption (G').

As an important case, we consider:

$$G^*(t, x, v, p, A) := \sup_{\gamma \in \Gamma} \left\{ \tfrac{1}{2}\mathrm{tr}[\sigma(t, x, \gamma)\sigma^T(t, x, \gamma)A] + \langle b(t, x, \gamma), p\rangle + f(t, x, v, p, \gamma) \right\}.$$

Here Γ is a compact metric space, $b(t, x, \gamma) : [0, T] \times \mathbb{R}^d \times \Gamma \mapsto \mathbb{R}^d$, $\sigma(t, x, \gamma) : [0, T] \times \mathbb{R}^d \times \Gamma \mapsto \mathbb{R}^{d\times d}$ and $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \Gamma \mapsto \mathbb{R}^n$ are given continuous function. Moreover, for all $\gamma \in \Gamma$,

$$|b(t, x, \gamma) - b(t, y, \gamma)| + |\sigma(t, x, \gamma) - \sigma(t, y, \gamma)| \leq \bar{L}|x - y|,$$

$$|f(t,x,v_1,p,\gamma)-f(t,x,v_2,p,\gamma)|\le \bar{L}|v_1-v_2|$$

with a constant $\bar{L}$ and

$$|f(t,x,v,p,\gamma)-f(t,y,v,p,\gamma)|\le \bar{\omega}(1+(T-t)^{-1}+|x|+|y|+|v|)\omega(|x-y|+|p|\cdot|x-y|),$$

where $\omega,\bar{\omega}:\mathbb{R}^+\mapsto\mathbb{R}^+$ are given continuous functions with $\omega(0)=0$.

From (C.2.11), for any $q,q'\in\mathbb{R}^d$,

$$\langle Xq,q\rangle+\langle Yq',q'\rangle\le 3\alpha|q-q'|^2$$

and

$$\begin{aligned}
&\mathrm{tr}[\sigma(t,x,\gamma)\sigma^T(t,x,\gamma)X+\sigma(t,y,\gamma)\sigma^T(t,y,\gamma)Y]\\
&=\mathrm{tr}[\sigma^T(t,x,\gamma)X\sigma(t,x,\gamma)+\sigma^T(t,y,\gamma)Y\sigma(t,y,\gamma)]\\
&=\sum_{i=1}^d[\langle X\sigma(t,x,\gamma)e_i,\sigma(t,x,\gamma)e_i\rangle+\langle Y\sigma(t,y,\gamma)e_i,\sigma(t,y,\gamma)e_i\rangle]\\
&\le 3\alpha d\bar{L}^2|x-y|^2.
\end{aligned}$$

Then we can derive that the function G^* satisfies assumption (**G**$'$). Moreover, we have the following result.

Theorem C.2.9 *Let $u\in USC([0,T]\times\mathbb{R}^d)$ be a G^*-subsolution and $v\in LSC([0,T]\times\mathbb{R}^d)$ a G^*-supersolution on $(0,T)\times\mathbb{R}^d$ both satisfying the polynomial growth condition. Then $u\le v$ on $[0,T)\times\mathbb{R}^d$ provided that $u|_{t=0}\le v|_{t=0}$.*

Proof We set $\xi(x):=(1+|x|^2)^{l/2}$ and, for a fixed and large enough constant λ, let

$$\tilde{u}_1(t,x):=u(t,x)\xi^{-1}(x)e^{-\lambda t},\quad \tilde{u}_2(t,x):=-v(t,x)\xi^{-1}(x)e^{-\lambda t},$$

where l is chosen to be large enough such that $\sum|\tilde{u}_i(t,x)|\to 0$ uniformly as $|x|\to\infty$. It is easy to check that, for any $i=1,2$, the function $\tilde{u}_i$ is a subsolution of

$$\partial_t\tilde{u}_i+\lambda\tilde{u}_i-\tilde{G}_i^*(t,x,\tilde{u}_i,D\tilde{u}_i,D^2\tilde{u}_i)=0.$$

Here the function $\tilde{G}_1^*(t,x,v,p,X)$ is given by

$$e^{-\lambda t}\xi^{-1}G^*(t,x,e^{\lambda t}v\xi,e^{\lambda t}(p\xi(x)+v\eta(x)),e^{\lambda t}(X\xi+p\otimes\eta(x)+\eta(x)\otimes p+v\kappa(x))),$$

while $\tilde{G}_2(t,x,v,p,X)$ is given by

$$-e^{-\lambda t}\xi^{-1}G^*(t,x,-e^{\lambda t}v\xi,-e^{\lambda t}(p\xi(x)+v\eta(x)),-e^{\lambda t}(X\xi+p\otimes\eta(x)+\eta(x)\otimes p+v\kappa(x))).$$

In the expressions for $\tilde{G}_1^*$ and $\tilde{G}^2$, we have used that

$$\eta(x) := \xi^{-1}(x)D\xi(x) = l(1+|x|^2)^{-1}x,$$
$$\kappa(x) := \xi^{-1}(x)D^2\xi(x) = l(1+|x|^2)^{-1}I + l(l-2)(1+|x|^2)^{-2}x\otimes x.$$

Then one can easily check that $\tilde{G}_i$ satisfies assumption (**G**$'$) by the conditions of G^*. Thus, from the proof of Theorem C.2.3 and the above discussion, we arrive at the desired result. □

C.3 Perron's Method and Existence

The combination of Perron's method and viscosity solutions was introduced by Ishii [91]. For the convenience of the readers, we interpret the proof provided in Crandall et al. [39] in its parabolic situation.

We consider the following parabolic PDE:

$$\begin{cases} \partial_t u - G(t,x,u,Du,D^2u) = 0 \text{ on } (0,\infty)\times\mathbb{R}^d, \\ u(0,x) = \varphi(x) \text{ for } x\in\mathbb{R}^d, \end{cases} \tag{C.3.1}$$

where $G : [0,\infty)\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\times\mathbb{S}(d)\mapsto\mathbb{R}$, $\varphi\in C(\mathbb{R}^d)$.

For our discussion on Perron's method, we need the following notations: if $u : \mathcal{O}\mapsto[-\infty,\infty]$ where $\mathcal{O}\subset[0,\infty)\times\mathbb{R}^d$, then

$$\begin{cases} u^*(t,x) = \lim_{r\downarrow 0}\sup\{u(s,y) : (s,y)\in\mathcal{O} \text{ and } \sqrt{|s-t|+|y-x|^2}\le r\}, \\ u_*(t,x) = \lim_{r\downarrow 0}\inf\{u(s,y) : (s,y)\in\mathcal{O} \text{ and } \sqrt{|s-t|+|y-x|^2}\le r\}. \end{cases} \tag{C.3.2}$$

The functionn u^* is called an **upper semicontinuous envelope** of u; it is the smallest upper semicontinuous function (with values in $[-\infty,\infty]$) satisfying $u\le u^*$. Similarly, u_* is the **lower semicontinuous envelope** of u.

Theorem C.3.1 (Perron's Method) *Let the comparison property holds for* (C.3.1)*, i.e., if w is a viscosity subsolution of* (C.3.1) *and v is a viscosity supersolution of* (C.3.1)*, then $w\le v$. Suppose also that there is a viscosity subsolution $\underline{u}$ and a viscosity supersolution $\bar{u}$ of* (C.3.1) *that satisfy the condition $\underline{u}_*(0,x) = \bar{u}^*(0,x) = \varphi(x)$ for $x\in\mathbb{R}^d$. Then the function*

$$W(t,x) = \sup\{w(t,x) : \underline{u}\le w\le\bar{u} \ \textit{ with } \ w \textit{ being a viscosity subsolution of (3.1)}\}$$

is a viscosity solution of (C.3.1).

The proof of Theorem C.3.1 is based on two lemmas. Their proofs can be found in [1]. We start with the first lemma:

Lemma C.3.2 *Let $\mathcal{F}$ be a family of viscosity subsolutions of* (C.3.1) *on* $(0,\infty)\times\mathbb{R}^d$. *Let $w(t,x) = \sup\{u(t,x) : u\in\mathcal{F}\}$ and assume that $w^*(t,x) < \infty$ for $(t,x)\in(0,\infty)\times\mathbb{R}^d$. Then w^* is a viscosity subsolution of* (C.3.1) *on* $(0,\infty)\times\mathbb{R}^d$.

Proof Let $(t, x) \in (0, \infty) \times \mathbb{R}^d$ and consider a sequence s_n, y_n, $u_n \in \mathcal{F}$ such that $\lim_{n\to\infty}(s_n, y_n, u_n(s_n, y_n)) = (t, x, w^*(t, x))$. There exists $r > 0$ such that the set $N_r = \{(s, y) \in (0, \infty) \times \mathbb{R}^d : \sqrt{|s-t| + |y-x|^2} \le r\}$ is compact. For $\phi \in C^2$ such that $\phi(t, x) = w^*(t, x)$ and $w^* < \phi$ on $(0, \infty) \times \mathbb{R}^d \backslash (t, x)$, let (t_n, x_n) be a maximum point of $u_n - \phi$ over N_r, hence $u_n(s, y) \le u_n(t_n, x_n) + \phi(s, y) - \phi(t_n, x_n)$ for $(s, y) \in N_r$. Suppose that (passing to a subsequence if necessary) $(t_n, x_n) \to (\bar{t}, \bar{x})$ as $n \to \infty$. Putting $(s, y) = (s_n, y_n)$ in the above inequality and taking the limit inferior as $n \to \infty$, we obtain

$$\begin{aligned} w^*(t, x) &\le \liminf_{n\to\infty} u_n(t_n, x_n) + \phi(t, x) - \phi(\bar{t}, \bar{x}) \\ &\le w^*(\bar{t}, \bar{x}) + \phi(t, x) - \phi(\bar{t}, \bar{x}). \end{aligned}$$

From the above inequalities and the assumption on ϕ, we get $\lim_{n\to\infty}(t_n, x_n, u_n(t_n, x_n)) = (t, x, w^*(t, x))$ (without passing to a subsequence). Since u_n is a viscosity subsolution of (C.3.1), by definition we have

$$\partial_t \phi(t_n, x_n) - G(t_n, x_n, u_n(t_n, x_n), D\phi(t_n, x_n), D^2\phi(t_n, x_n)) \le 0.$$

Letting $n \to \infty$, we conclude that

$$\partial_t \phi(t, x) - G(t, x, w^*(t, x), D\phi(t, x), D^2\phi(t, x)) \le 0.$$

Therefore, w^* is a viscosity subsolution of (C.3.1) by definition. □

The second step in the proof of Theorem C.3.1 is a simple " bump" construction that we now describe. Suppose that u is a viscosity subsolution of (C.3.1) on $(0, \infty) \times \mathbb{R}^d$ and that u_* is not a viscosity supersolution of (C.3.1), so that there is $(t, x) \in (0, \infty) \times \mathbb{R}^d$ and $\phi \in C^2$ with $u_*(t, x) = \phi(t, x)$, $u_* > \phi$ on $(0, \infty) \times \mathbb{R}^d \backslash (t, x)$ and

$$\partial_t \phi(t, x) - G(t, x, \phi(t, x), D\phi(t, x), D^2\phi(t, x)) < 0.$$

The continuity of G provides $r, \delta_1 > 0$ such that $N_r = \{(s, y) : \sqrt{|s-t| + |y-x|^2} \le r\}$ is compact and

$$\partial_t \phi - G(s, y, \phi + \delta, D\phi, D^2\phi) \le 0$$

for all $s, y, \delta \in N_r \times [0, \delta_1]$. Lastly, we obtain $\delta_2 > 0$ for which $u_* > \phi + \delta_2$ on the boundary ∂N_r. Setting $\delta_0 = \min(\delta_1, \delta_2) > 0$, we define

$$U = \begin{cases} \max(u, \phi + \delta_0), & \text{on } N_r, \\ u, & \text{elsewhere.} \end{cases}$$

By the above inequalities and Lemma C.3.2, it is easy to check that U is a viscosity subsolution of (C.3.1) on $(0, \infty) \times \mathbb{R}^d$. Obviously, $U \ge u$. Finally, observe that

$U_*(t,x) \geq \max(u_*(t,x), \phi(t,x)+\delta_0) > u_*(t,x)$; hence there exists (s,y) such that $U(s,y) > u(s,y)$. We summarize the above discussion as the following lemma.

Lemma C.3.3 *Let u be a viscosity subsolution of (C.3.1) on $(0,\infty)\times\mathbb{R}^d$. If u_* fails to be a viscosity supersolution at some point (s,z), then for any small $\kappa > 0$ there is a viscosity subsolution U_κ of (C.3.1) on $(0,\infty)\times\mathbb{R}^d$ satisfying*

$$\begin{cases} U_\kappa(t,x) \geq u(t,x) \ \ \textit{and} \ \ \sup(U_\kappa - u) > 0, \\ U_\kappa(t,x) = u(t,x) \quad \textit{for} \sqrt{|t-s|+|x-z|^2} \geq \kappa. \end{cases}$$

Proof of Theorem C.3.1 With the notation in the theorem, we observe that $\underline{u}_* \leq W_* \leq W \leq W^* \leq \bar{u}^*$ and, in particular, $W_*(0,x) = W(0,x) = W^*(0,x) = \varphi(x)$ for $x \in \mathbb{R}^d$. By Lemma C.3.2, W^* is a viscosity subsolution of (C.3.1) and hence, by comparison, $W^* \leq \bar{u}$. It then follows from the definition of W that $W = W^*$ (so W is a viscosity subsolution). If W_* fails to be a viscosity supersolution at some point $(s,z) \in (0,\infty)\times\mathbb{R}^d$, let W_κ be provided by Lemma C.3.3. Clearly $\underline{u} \leq W_\kappa$ and $W_\kappa(0,x) = \varphi(x)$ for sufficiently small κ. By comparison, $W_\kappa \leq \bar{u}$ and since W is the maximal viscosity subsolution between $\underline{u}$ and $\bar{u}$, we arrive at the contradiction $W_\kappa \leq W$. Hence W_* is a viscosity supersolution of (C.3.1) and then, by comparison for (C.3.1), $W^* = W \leq W_*$, showing that W is continuous and is a viscosity solution of (C.3.1). The proof is complete. □

Let $G : \mathbb{R}^d \times \mathbb{S}(d) \to \mathbb{R}$ be a given continuous sublinear function monotone in $A \in \mathbb{S}(d)$. We now consider the existence of viscosity solution of the following G-equation:

$$\partial_t u - G(Du, D^2u) = 0, \ \ u(0,x) = \varphi(x). \tag{C.3.3}$$

Case 1: If $\varphi \in C_b^2(\mathbb{R}^d)$, then $\underline{u}(t,x) = \underline{M}t + \varphi(x)$ and $\bar{u}(t,x) = \overline{M}t + \varphi(x)$ are respectively the classical subsolution and supersolution of (C.3.3), where

$$\underline{M} = \inf_{x\in\mathbb{R}^d} G(D\varphi(x), D^2\varphi(x)),$$
$$\overline{M} = \sup_{x\in\mathbb{R}^d} G(D\varphi(x), D^2\varphi(x)).$$

Obviously, $\underline{u}$ and $\bar{u}$ satisfy all the conditions in Theorem C.3.1. By Theorem C.2.8, we know that the comparison holds for (C.3.3). Thus by Theorem C.3.1, we obtain that the G-equation (C.3.3) has a viscosity solution.

Case 2: If $\varphi \in C_b(\mathbb{R}^d)$ with $\lim_{|x|\to\infty}\varphi(x) = 0$, then we can choose a sequence $\varphi_n \in C_b^2(\mathbb{R}^d)$ which uniformly converge to φ as $n \to \infty$. For φ_n, by Case 1, there exists a viscosity solution u^{φ_n}. By the comparison theorem, it is easy to show that u^{φ_n} is uniformly convergent, the limit denoted by u. Similarly to the proof of Lemma C.3.2, it is easy to show that u is a viscosity solution of the G-equation (C.3.3) with initial condition φ.

Case 3: If $\varphi \in C(\mathbb{R}^d)$ with polynomial growth, then we can choose a large $l > 0$ such that $\tilde{\varphi}(x) = \varphi(x)\xi^{-1}(x)$ satisfies the condition in Case 2, where $\xi(x) = (1 +$

$|x|^2)^{l/2}$. It is easy to check that u is a viscosity solution of the G-equation (C.3.3) if and only if $\tilde{u}(t,x) = u(t,x)\xi^{-1}(x)$ is a viscosity solution of the following PDE:

$$\partial_t \tilde{u} - \tilde{G}(x, \tilde{u}, D\tilde{u}, D^2\tilde{u}) = 0, \ \ \tilde{u}(0,x) = \tilde{\varphi}, \tag{C.3.4}$$

where $\tilde{G}(x, v, p, X) = G(p + v\eta(x), X + p \otimes \eta(x) + \eta(x) \otimes p + v\kappa(x))$. Here

$$\eta(x) := \xi^{-1}(x) D\xi(x) = l(1+|x|^2)^{-1}x,$$
$$\kappa(x) := \xi^{-1}(x) D^2\xi(x) = l(1+|x|^2)^{-1}I + l(l-2)(1+|x|^2)^{-2}x \otimes x.$$

Similarly to the above discussion, we obtain that there exists a viscosity solution of (C.3.4) with initial condition $\widetilde{\varphi}$. Thus there exists a viscosity solution of the G-equation (C.3.3).

We summarize the above discussions as a theorem.

Theorem C.3.4 *Let $\varphi \in C(\mathbb{R}^d)$ with polynomial growth. Then there exists a unique viscosity solution of the G-equation (C.3.3) with initial condition $u(t,\cdot)|_{t=0} = \varphi(\cdot)$.*

Moreover, let (X, η) be a $2d$-dimensional G-distributed random variable on a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ with

$$G(p, A) = \hat{\mathbb{E}}\left[\frac{1}{2}\langle AX, X\rangle + \langle p, \eta\rangle\right], \ \forall (p, A) \in \mathbb{R}^d \times \mathbb{S}(d).$$

From Sect. 2.1 in Chap. 2 and the above theorem, it is easy to check that the G-solution with $u|_{t=0} = \varphi$ is given by $u(t,x) = \hat{\mathbb{E}}[\varphi(x + \sqrt{t}X + t\eta)]$.

The next is to discuss the case of a nonlinear function $\widetilde{G}$ without the sublinear assumption. The results can be applied to obtain a wide class of $\widetilde{G}$-nonlinear expectations.

Theorem C.3.5 *Let the function G be given as in Theorem C.3.4, and let $\widetilde{G}(p, X)$: $\mathbb{R}^N \times \mathbb{S}(N) \mapsto \mathbb{R}$ be a given function satisfying condition* **(G)** *(see Sect. C.2) and $\widetilde{G}(0,0) = 0$. We assume that this function $\widetilde{G}$ is dominated by G in the sense that*

$$\widetilde{G}(p, X) - \widetilde{G}(p', X') \le G(p - p', X - X'), \quad \text{for } p, p' \in \mathbb{R}^d \text{ and } \ X, X' \in \mathbb{S}(d).$$

Then there exists a unique family of functions

$$\{\widetilde{u}^{\varphi} | \ \ \varphi \in C(\mathbb{R}^N) \text{ with polynomial growth } \}$$

such that $\widetilde{u}^{\varphi}$ is a viscosity solution of

$$\partial_t \widetilde{u}^{\varphi} - \widetilde{G}(D\widetilde{u}^{\varphi}, D^2\widetilde{u}^{\varphi}) = 0, \quad \text{with } \ \widetilde{u}^{\varphi}|_{t=0} = \varphi, \tag{C.3.5}$$

satisfying

$$\widetilde{u}^0 \equiv 0, \quad \widetilde{u}^{\varphi} - \widetilde{u}^{\psi} \leq u^{\varphi-\psi}, \tag{C.3.6}$$

where u^{φ} is the viscosity solution of the G-equation (C.3.3) *with the initial condition $u^{\varphi}|_{t=0} = \varphi$. Moreover, it holds that*

$$\widetilde{u}^{\varphi} \leq \widetilde{u}^{\psi}, \quad \text{if } \varphi \leq \psi. \tag{C.3.7}$$

Proof We denote $G_*(p, X) := -G(-p, -X)$, it is clear that

$$G_*(p - p', A - A') \leq \widetilde{G}(p, A) - \widetilde{G}(p', A') \leq G(p - p', A - A'), \quad \forall p, p' \in \mathbb{R}^d, \quad A, A' \in \mathbb{S}(d).$$

Thus $\widetilde{G}$ satisfies Lipschitz condition in (p, A) and $\widetilde{G}(0, 0) = 0$. The uniqueness is obvious since that $u^0 \equiv 0$. Moreover, for $\varphi \leq \psi$, we have $u^{\varphi-\psi} \leq 0$ which implies that $\widetilde{u}^{\varphi} \leq \widetilde{u}^{\psi}$. Next we will consider the existence.

Note that the G-solution with $u|_{t=0} = \varphi$ is given by $\bar{u}(t, x) = \hat{\mathbb{E}}[\varphi(x + \sqrt{t}X + t\eta)]$ and it is also a $\widetilde{G}$-supersolution. Similarly denoting $G_*(p, X) := -G(-p, -X)$, the G_*-solution with $u|_{t=0} = \varphi$ is given by $\underline{u}(t, x) = -\hat{\mathbb{E}}[-\varphi(x + \sqrt{t}X + t\eta)]$, which is a $\widetilde{G}$-subsolution.

Case 1: If $\varphi \in C(\mathbb{R}^N)$ vanishing at infinity, then we can find some constant C so that for each $m > 0$

$$|\varphi(x + z)| \leq \sup_{y \in [-m,m]} |\varphi(x + y)| + \frac{C}{m}|z|,$$

which implies that

$$\sup_{t \in [0,T]} \hat{\mathbb{E}}[|\varphi(x + \sqrt{t}X + t\eta)] \leq \sup_{y \in [-m,m]} |\varphi(x + y)| + \frac{C}{m} \sup_{t \in [0,T]} \hat{\mathbb{E}}[\sqrt{t}|X| + t|\eta|].$$

Letting $x \to \infty$ and then $m \to \infty$, we could get that $\bar{u}$ and $\underline{u}$ both satisfy condition **(U)**. In sprit of Remark C.2.7, the comparison holds true for $\widetilde{G}$-solution under condition (**G**). We then can apply Theorem C.3.1 to prove that $\widetilde{G}$-solution u with $u|_{t=0} = \varphi$ exists. Moreover, it follows from Remark C.2.7 that equation (C.3.6) holds true.

Case 2: If $\varphi \in C(\mathbb{R}^N)$ satisfying $|\varphi(x)| \leq C(1 + |x|^p)$ for some constants C and p, then we can find a sequence $\varphi_n \in C(\mathbb{R}^N)$ vanishing at infinity which converges to φ on each compact set and satisfies $|\varphi_n| \leq C(1 + |x|^p)$. By the result of Case 1, the viscosity solution u^{φ_n} with initial condition φ_n exists and

$$\widetilde{u}^{\varphi_n} - \widetilde{u}^{\varphi_m} \leq u^{\varphi_n - \varphi_m}.$$

Note that $u^{\varphi_n - \varphi_m} = \hat{\mathbb{E}}[\varphi_n(x + \sqrt{t}X + t\eta) - \varphi_m(x + \sqrt{t}X + t\eta)]$. By a similar analysis as in Step 1, we can obtain that $(\widetilde{u}^{\varphi_n})_{n=1}^{\infty}$ is uniformly convergent on each compact set of $[0, T] \times \mathbb{R}^N$. We then denote

$$\widetilde{u}^{\varphi} := \lim_{n \to \infty} \widetilde{u}^{\varphi_n}.$$

Using the approach in the proof of Lemma C.3.2, it is easy to check that $\widetilde{u}^{\varphi}$ is a viscosity solution with initial condition φ (stability properties, see also [39]). Moreover, relation (C.3.6) remains true. □

C.4 Krylov's Regularity Estimate for Parabolic PDEs

The proof of our new central limit theorem is based on powerful $C^{1+\alpha/2,2+\alpha}$-regularity estimates for fully nonlinear parabolic PDE obtained by Krylov [105]. A more recent result of Wang [171] (the version for elliptic PDE was initially introduced in Cabre and Caffarelli [24]), using viscosity solution arguments, can also be applied.

For simplicity, we only consider the following type of PDE:

$$\partial_t u + G(D^2 u, Du, u) = 0, \quad u(T, x) = \varphi(x), \tag{C.4.1}$$

where $G : \mathbb{S}(d) \times \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ is a given function and $\varphi \in C_b(\mathbb{R}^d)$.

Following Krylov [105], we fix constants $K \geq \varepsilon > 0$, $T > 0$ and set $Q = (0, T) \times \mathbb{R}^d$. Now we give the definition of two sets of functions $\mathcal{G}(\varepsilon, K, Q)$ and $\bar{\mathcal{G}}(\varepsilon, K, Q)$.

The next definition is according to Definition 5.5.1 in Krylov [105].

Definition C.4.1 Let $G : \mathbb{S}(d) \times \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ be given; we write G as follows: $G(u_{ij}, u_i, u), i, j = 1, \ldots, d$. We say that $G \in \mathcal{G}(\varepsilon, K, Q)$ if G is twice continuously differentiable with respect to all arguments (u_{ij}, u_i, u) and, for any real $u_{ij} = u_{ji}$, $\tilde{u}_{ij} = \tilde{u}_{ji}$, u_i, $\tilde{u}_i$, u, $\tilde{u}$ and λ_i, the following inequalities hold:

$$\varepsilon|\lambda|^2 \leq \sum_{i,j} \lambda_i \lambda_j \partial_{u_{ij}} G \leq K|\lambda|^2,$$

$$|G - \sum_{i,j} u_{ij} \partial_{u_{ij}} G| \leq M_1^G(u)(1 + \sum_i |u_i|^2),$$

$$|\partial_u G| + (1 + \sum_i |u_i|) \sum_i |\partial_{u_i} G| \leq M_1^G(u)(1 + \sum_i |u_i|^2 + \sum_{i,j} |u_{ij}|),$$

$$\begin{aligned}[M_2^G(u, u_k)]^{-1} G_{(\eta)(\eta)} \leq & \sum_{i,j} |\tilde{u}_{ij}| \Big[\sum_i |\tilde{u}_i| + (1 + \sum_{i,j} |u_{ij}|)|\tilde{u}| \Big] \\ & + \sum_i |\tilde{u}_i|^2 (1 + \sum_{i,j} |u_{ij}|) + (1 + \sum_{i,j} |u_{ij}|^3)|\tilde{u}|^2.\end{aligned}$$

Here the arguments (u_{ij}, u_i, u) of G and its derivatives are omitted, $\eta = (\tilde{u}_{ij}, \tilde{u}_i, \tilde{u})$, and

$$G_{(\eta)(\eta)} := \sum_{i,j,r,s} \tilde{u}_{ij}\tilde{u}_{rs}\partial^2_{u_{ij}u_{rs}} G + 2\sum_{i,j,r} \tilde{u}_{ij}\tilde{u}_r\partial^2_{u_{ij}u_r} G + 2\sum_{i,j} \tilde{u}_{ij}\tilde{u}\partial^2_{u_{ij}u} G$$
$$+ \sum_{i,j} \tilde{u}_i\tilde{u}_j\partial^2_{u_iu_j} G + 2\sum_i \tilde{u}_i\tilde{u}\partial^2_{u_iu} G + |\tilde{u}|^2\partial^2_{uu} G,$$

$M_1^G(u)$ and $M_2^G(u, u_k)$ are some continuous functions which grow with $|u|$ and $u_k u_k$ and $M_2^G \geq 1$.

Remark C.4.2 Let $\varepsilon I \leq A = (a_{ij}) \leq KI$. It is easy to check that

$$G(u_{ij}, u_i, u) = \sum_{i,j} a_{ij}u_{ij} + \sum_i b_i u_i + cu$$

belongs to the set $\mathcal{G}(\varepsilon, K, Q)$.

The next definition is Definition 6.1.1 in Krylov [105].

Definition C.4.3 Let a function $G = G(u_{ij}, u_i, u) : \mathbb{S}(d) \times \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ be given. We write $G \in \bar{\mathcal{G}}(\varepsilon, K, Q)$ if there exists a sequence $G_n \in \mathcal{G}(\varepsilon, K, Q)$ converging to G as $n \to \infty$ at every point $(u_{ij}, u_i, u) \in \mathbb{S}(d) \times \mathbb{R}^d \times \mathbb{R}$ and such that:

(i) $M_i^{G_1} = M_i^{G_2} = \cdots =: M_i^G$, $i = 1, 2$;
(ii) for any $n = 1, 2, \ldots$, the function G_n is infinitely differentiable with respect to (u_{ij}, u_i, u);
(iii) there exist constants $\delta_0 =: \delta_0^G > 0$ and $M_0 =: M_0^G > 0$ such that

$$G_n(u_{ij}, 0, -M_0) \geq \delta_0, \;\; G_n(-u_{ij}, 0, M_0) \leq -\delta_0$$

for any $n \geq 1$ and symmetric nonnegative matrices (u_{ij}).

The following theorem is Theorem 6.4.3 in Krylov [105], and it plays an important role in the proof of our central limit theorem.

Theorem C.4.4 *Suppose that* $G \in \bar{\mathcal{G}}(\varepsilon, K, Q)$ *and* $\varphi \in C_b(\mathbb{R}^d)$ *with* $\sup_{x\in\mathbb{R}^d} |\varphi(x)| \leq M_0^G$. *Then the PDE* (C.4.1) *has a solution* u *possessing the following properties:*

(i) $u \in C([0, T] \times \mathbb{R}^d)$, $|u| \leq M_0^G$ *on* Q;
(ii) *there exists a constant* $\alpha \in (0, 1)$ *depending only on* d, K, ε *such that for any* $\kappa > 0$,

$$||u||_{C^{1+\alpha/2,2+\alpha}([0,T-\kappa]\times\mathbb{R}^d)} < \infty. \quad \text{(C.4.2)}$$

Now we consider the G-equation. Let $G : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ be a given continuous, sublinear function monotone in $A \in \mathbb{S}(d)$. Then there exists a bounded, convex and closed subset $\Sigma \subset \mathbb{R}^d \times \mathbb{S}_+(d)$ such that

$$G(p, A) = \sup_{(q,B)\in\Sigma} \left[\frac{1}{2}\mathrm{tr}[AB] + \langle p, q\rangle\right] \quad \text{for } (p, A) \in \mathbb{R}^d \times \mathbb{S}(d). \tag{C.4.3}$$

We are interested in the following G-equation:

$$\partial_t u + G(Du, D^2u) = 0, \quad u(T, x) = \varphi(x). \tag{C.4.4}$$

The idea is to set

$$\tilde{u}(t, x) = e^{t-T} u(t, x) \tag{C.4.5}$$

and easily check that $\tilde{u}$ satisfies the following PDE:

$$\partial_t \tilde{u} + G(D\tilde{u}, D^2\tilde{u}) - \tilde{u} = 0, \quad \tilde{u}(T, x) = \varphi(x). \tag{C.4.6}$$

Suppose that there exists a constant $\varepsilon > 0$ such that for any $A, \bar{A} \in \mathbb{S}(d)$ with $A \geq \bar{A}$, we have

$$G(0, A) - G(0, \bar{A}) \geq \varepsilon\mathrm{tr}[A - \bar{A}]. \tag{C.4.7}$$

Since G is continuous, it is easy to prove that there exists a constant $K > 0$ such that for any $A, \bar{A} \in \mathbb{S}(d)$ with $A \geq \bar{A}$, we have

$$G(0, A) - G(0, \bar{A}) \leq K\mathrm{tr}[A - \bar{A}].$$

Thus for any $(q, B) \in \Sigma$, the following relations hold:

$$2\varepsilon I \leq B \leq 2KI.$$

By Remark C.4.2, it is easy to check that $\tilde{G}(u_{ij}, u_i, u) := G(u_i, u_{ij}) - u \in \bar{\mathcal{G}}(\varepsilon, K, Q)$ and $\delta_0^G = M_0^G$ can be any positive constant. By Theorem C.4.4 and relation (C.4.5), we obtain the following regularity estimate for the G-equation (C.4.4).

Theorem C.4.5 *Let G satisfy* (C.4.3) *and* (C.4.7)*, $\varphi \in C_b(\mathbb{R}^d)$ and let u be a solution of the G-equation* (C.4.4)*. Then there exists a constant $\alpha \in (0, 1)$ depending only on d, G, ε such that for any $\kappa > 0$,*

$$||u||_{C^{1+\alpha/2,2+\alpha}([0,T-\kappa]\times\mathbb{R}^d)} < \infty.$$

References

1. Alvarez, O., Tourin, A.: Viscosity solutions of nonlinear integro-differential equations. Ann. Inst. H. Poincaré Anal. Non Linéaire **13**(3), 293–317 (1996)
2. Artzner, Ph, Delbaen, F., Eber, J.M., Heath, D.: Thinking coherently. RISK **10**(11), 68–71 (1997)
3. Artzner, Ph, Delbaen, F., Eber, J.M., Heath, D.: Coherent measures of risk. Math. Financ. **9**(3), 203–228 (1999)
4. Ash, R.: Real Analysis and Probability. Academy Press, New York, London (1972)
5. Atlan, M.: Localizing Volatilities (2006). arXiv:0604316v1
6. Avellaneda, M., Lévy, A., Paras, A.: Pricing and hedging derivative securities in markets with uncertain volatilities. Appl. Math. Financ. **2**(2), 73–88 (1995)
7. Bachelier, L.: Théorie de la spéculaton. Ann. Scientifiques de l'École Normale Supérieure **17**, 21–86 (1900)
8. Barenblatt, G.I.: Similarity. Consultants Bureau, Self-Similarity and Intermediate Asymptotics (1979)
9. Barles, G.: Solutions de viscosité des équations de Hamilton-Jacobi. Collection "Math ématiques et Applications" de la SMAI, vol. 17. Springer (1994)
10. Barrieu, P., El Karoui, N.: Pricing, hedging and optimally designing derivatives via minimization of risk measures, Preprint. In: Carmona, R. et al. (eds.) Volume on Indifference Pricing. Princeton University Press (in press)
11. Bayraktar, E., Munk, A.: α-Stable limit theorem under sublinear expectation. Bernoulli **22**(4), 2548–2578 (2016)
12. Bensoussan, A.: Lectures on Stochastic Control, LNM 972. Springer (1981)
13. Bensoussan, A.: Stochastic Control by Functional Analysis Methods, North-Holland (1982)
14. Beissner, P., Riedel, F.: Non-implementability of Arrow-Debreu equilibria by continuous trading under volatility uncertainty. Financ. Stoch. **22**, 603–620 (2018)
15. Beissner, P., Riedel, F.: Equilibria under Knightian Price Uncertainty. Econometrica **87**(1), 37–64 (2019)
16. Beissner, P., Denis, L.: Duality, the theory of value and Asset Pricing under Knightian Uncertainty. SIAM J. Financ. Math. **9**(1), 381–400 (2018)
17. Bismut, J.M.: Conjugate convex functions in optimal stochastic control. J. Math. Anal. Appl. **44**(2), 384–404 (1973)
18. Biagini, F., Mancin, J., Meyer Brandis, T.: Robust mean variance hedging via G-expectation. Stoch. Process. Appl. **129**(4), 1287–1325 (2019)
19. Bismut, J.M.: Contrôle des systemes linéaires quadratiques: applications de l'integrale stochastique, Sémin. Proba. XII., LNM, vol. 649, pp. 180–264. Springer (1978)

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7

20. Briand, Ph, Coquet, F., Hu, Y., Mémin, J., Peng, S.: A converse comparison theorem for BSDEs and related properties of g-expectations. Electron. Commun. Probab. **5**(13), 101–117 (2000)
21. Bogachev, V.I.: Gaussian measures. Amer. Math. Soc. Math. Surv. Monogr. 62 (1998)
22. Bion-Nadal, J., Kervarec, M.: Risk measuring under model uncertainty. Ann. Appl. Probab. **22**(1), 213–238 (2012)
23. Bouleau, N., Hirsch, F.: Dirichlet Forms and Analysis on Wiener Space. De Gruyter (1991)
24. Cabré, X., Caffarelli, L.A.: (1997) Fully Nonlinear elliptic partial differential equations. Am. Math. Soc
25. Cao, D., Tang, S.: Reflected Quadratic BSDEs driven by G-Brownian Motions. Chinese Annals of Mathematics, Series B (2019). in press
26. Chen, Z.: A property of backward stochastic differential equations. C.R. Acad. Sci. Paris, Sér.I **326**(4), 483–488 (1998)
27. Chen, Z.: Strong laws of large numbers for sub-linear expectations. Sci. China Math. **59**, 945–954 (2016)
28. Chen, Z., Epstein, L.: Ambiguity, risk and asset returns in continuous time. Econometrica **70**(4), 1403–1443 (2002)
29. Chen, Z., Kulperger, R., Jiang, L.: Jensen's inequality for g-expectation: part 1. C. R. Acad. Sci. Paris **337**(11), 725–730 (2003)
30. Chen, Z., Peng, S.: A nonlinear Doob-Meyer type decomposition and its application. SUT J. Math. (Japan) **34**(2), 197–208 (1998)
31. Chen, Z., Peng, S.: A general downcrossing inequality for g-martingales. Statist. Probab. Lett. **46**(2), 169–175 (2000)
32. Cheridito, P., Soner, H.M., Touzi, N., Victoir, N.: Second order backward stochastic differential equations and fully non-linear parabolic PDEs. Comm. Pure Appl. Math. **60**(7), 1081–1110 (2007)
33. Choquet, G. (1953) Theory of capacities. Ann. Inst. Fourier (Grennoble) **5**, 131–295
34. Chung, K.L., Williams, R.: Introduction to Stochastic Integration, 2nd edn. Birkhäuser (1990)
35. Coquet, F., Hu, Y., Mémin, J., Peng, S.: A general converse comparison theorem for Backward stochastic differential equations. C. R. Acad. Sci. Paris **333**(6), Serie I, 577–581 (2001)
36. Coquet, F., Hu, Y., Mémin, J., Peng, S.: Filtration-consistent nonlinear expectations and related g-expectations. Probab. Theory Relat. Fields **123**(1), 1–27 (2002)
37. Crandall, M.G., Lions, P.L.: Condition d'unicité pour les solutions generalisées des équations de Hamilton-Jacobi du premier ordre. C. R. Acad. Sci. Paris Sér. I Math. **292**, 183–186 (1981)
38. Crandall, M.G., Lions, P.L.: Viscosity solutions of Hamilton-Jacobi equations. Trans. Am. Math. Soc. **277**(1), 1–42 (1983)
39. Crandall, M.G., Ishii, H., Lions, P.L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. **27**(1), 1–67 (1992)
40. Crepey, S.: Financial modeling. A Backward Stochastic Differential Equations Prospective. Springer Finance Textbook (2013)
41. Daniell, P.J.: A general form of integral. Ann. Math. **19**(4), 279–294 (1918)
42. Dellacherie, C.: Capacités et Processus Stochastiques. Springer (1972)
43. Dellacherie, C., Meyer, P.A.: Probabilities and Potential A and B, North–Holland (1978 and 1982)
44. Delbaen, F.: Representing martingale measures when asset prices are continuous and bounded. Math. Financ. **2**(2), 107–130 (1992)
45. Delbaen, F.: Coherent Risk Measures (Lectures given at the Cattedra Galileiana at the Scuola Normale di Pisa, March 2000). Published by the Scuola Normale di Pisa (2002)
46. Delbaen, F., Peng, S., Rosazza Gianin, E.: Representation of the penalty term of dynamic concave utilities. Financ. Stoch. **14**(3), 449–472 (2010)
47. Denis, L., Hu, M., Peng, S.: Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. Potential Anal. **34**(2), 139–161 (2011)
48. Denis, L., Martini, C.: A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. Ann. Appl. Probab. **16**(2), 827–852 (2006)

49. Denneberg, D.: Non-Additive Measure and Integral. Kluwer (1994)
50. Denk, R., Kupper, M., Nendel, M.: A semigroup approach to nonlinear Lévy processes. Stoch. Process. Appl. (2019). in press
51. Doob, J.L.: Classical Potential Theory and its Probabilistic Counterpart. Springer (1984)
52. Dolinsky, Y., Nutz, M., Soner, H.M.: Weak approximation of G-expectations. Stoch. Process. Appl. **122**(2), 664–675 (2012)
53. Dolinsky, Y.: Numerical schemes for G-expectations. Electron. J. Probab. **17**(98), 1–15 (2012)
54. Dolinsky, Y.: Hedging of game options under model uncertainty in discrete time. Electron. Commun. Probab. **19**(19), 1–11 (2014)
55. Dudley, R.M.: Real Analysis and Probability. Wadsworth, Brooks& Cole (1989)
56. Einstein, A. (1905) On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. English translation in Investigations on the Theory of the Brownian Movement. Repr. Dover (1956)
57. El Karoui, N., Quenez, M.C.: Dynamic programming and pricing of contingent claims in incomplete market. SIAM J. Control Optim. **33**(1), 29–66 (1995)
58. El Karoui, N., Peng, S., Quenez, M.C.: Backward stochastic differential equation in finance. Math. Financ. **7**(1), 1–71 (1997)
59. Epstein, L., Ji, S.: Ambiguous volatility and asset pricing in continuous time. Rev. Financ. Stud. **26**(7), 1740–1786 (2013)
60. Epstein, L., Ji, S.: Ambiguous volatility, possibility and utility in continuous time. J. Math. Econom. **50**, 269–282 (2014)
61. Evans, L.C.: A convergence theorem for solutions of nonlinear second order elliptic equations. Indiana Univ. Math. J. **27**(5), 875–887 (1978)
62. Evans, L.C.: On solving certain nonlinear differential equations by accretive operator methods. Israel J. Math. **36**(3–4), 225–247 (1980)
63. Fang, X., Peng, S.G., Shao, Q.M., Song, Y.S.: Limit Theorems with Rate of Convergence Under Sublinear Expectations. arXiv:1711.10649
64. Feller, W.: An Introduction to Probability Theory and its Applications, 1 (3rd edn.); 2(2nd edn.). Wiley (1968,1971)
65. Feyel, D., de La Pradelle, A.: Espaces de Sobolev gaussiens. Ann. Inst. Fourier (Grenoble) **39**, 875–908 (1989)
66. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions. Springer (1992)
67. Fouque, J.-P., Pun, C.S., Wong, H.Y.: Portfolio optimization withambiguous correlation and stochastic volatilities. SIAM. J. Control. Optim. **54**(5), 2309–2338 (2016)
68. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Financ. Stoch. **6**(4), 429–447 (2002)
69. Föllmer, H., Schied, A.: Statistic Finance, An Introduction in Discrete Time, 2nd Edn. de Gruyter (2004)
70. Frittelli, M., Rossaza Gianin, E.: Putting order in risk measures. J. Bank. Financ. **26**(7), 1473–1486 (2002)
71. Frittelli, M., Rossaza Gianin, E.: Dynamic convex risk measures. Risk Measures for the 21st Century, pp. 227–247. Wiley (2004)
72. Gao, F.: Pathwise properties and homeomorphic flows for stochastic differential equations driven by G-Brownian motion. Stoch. Process. Appl. **119**(10), 3356–3382 (2009)
73. Gao, F., Jiang, H.: Large deviations for stochastic differential equations driven by G-Brownian motion. Stoch. Process. Appl. **120**(11), 2212–2240 (2009)
74. He, S.W., Wang, J.G., Yan, J.A.: Semimartingale Theory and Stochastic Calculus. CRC Press (1992)
75. Hu, M., Li, X.: Independence under the G-expectation framework. J. Theor. Probab. **27**(3), 1011–1020 (2014)
76. Hu, M., Ji, S.: Dynamic programming principle for stochastic recursive optimal control problem driven by a G-Brownian motion. Stoch. Process. Appl. **127**(1), 107–134 (2017)

77. Hu, M., Li, H., Wang, F., Zheng, G.: Invariant and ergodic nonlinear expectations for G-diffusion processes. Electron. Commun. Probab. **20**(30), 1–15 (2015)
78. Hu, M., Ji, S., Peng, S., Song, Y.: Backward stochastic differential equations driven by G-Brownian motion. Stoch. Process. Appl. **124**(1), 759–784 (2014)
79. Hu, M., Ji, S., Peng, S., Song, Y.: Comparison theorem, Feynman-Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion. Stoch. Process. Appl. **124**(2), 1170–1195 (2014)
80. Hu, M., Peng, S.: On representation theorem of G-Expectations and paths of G-Brownian motion. Acta Math. Appl. Sinica, Engl. Ser. **25**(3), 539–546 (2009)
81. Hu, M., Peng, S.: G-Lévy Processes Under Sublinear Expectations (2009). arXiv:0911.3533
82. Hu, M., Peng, S.: Extended Conditional G-expectations and Related Stopping Times (2013). arXiv:1309.3829
83. Hu, M.S., Peng, S.G., Song, Y.S.: Stein Type Characterization for G-normal Distributions (2016). arXiv:1603.04611v1
84. Hu, M., Wang, F.: Ergodic BSDEs driven by G-Brownian motion and applications. Stoch. Dyn. **18**(6), 1–35 (2018)
85. Hu, M., Wang, F., Zheng, G.: Quasi-continuous random variables and processes under the G-expectation framework. Stoch. Process. Appl. **126**(8), 2367–2387 (2016)
86. Hu, Y.: On Jensen's inequality for g-expectation and for nonlinear expectation. Arch. der Math. **85**(6), 572–580 (2005)
87. Hu, Y., Lin, Y., Soumana Hima, A.: Quadratic backward stochastic differential equations driven by G-Brownian motion: discrete solutions and approximation. Stoch. Process. Appl. **128**(11), 3724–3750 (2018)
88. Huber, P.J.: Robust Statistics. Wiley (1981)
89. Huber, P., Strassen, V.: Minimax tests and the Neyman-Pearson lemma for capacity. Ann. Statist. **1**(2), 252–263 (1973)
90. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes, North–Holland (1981)
91. Ishii, H.: Perron's method for Hamilton-Jacobi equations. Duke Math. J. **55**(2), 369–384 (1987)
92. Itô, K.: Differential equations determining a Markoff process. J. Pan–Jpn. Math. Coll. No. 1077 (1942). In Kiyosi Itô: Selected Papers, Springer (1987)
93. Itô, K., McKean, M.: Diffusion Processes and their Sample Paths, Springer (1965)
94. Jensen, R.: The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. Arch. Rat. Mech. Anal. **101**(1), 1–27 (1988)
95. Jia, G., Peng, S.: Jensen's inequality for g-convex function under g-expectation. Probab. Theory Relat. Fields **147**(1–2), 217–239 (2010)
96. Jiang, L.: Some results on the uniqueness of generators of backward stochastic differential equations. C. R. Acad. Sci. Paris **338**(7), 575–580 (2004)
97. Jiang, L., Chen, Z.: A result on the probability measures dominated by g-expectation, Acta Math. Appl. Sinica Engl. Ser. **20**(3), 507–512 (2004)
98. Jiang, L., Chen, Z.: On Jensen's inequality for g-expectation. Chin. Ann. Math. **25B**(3), 401–412 (2004)
99. Jin, H., Peng, S.: Optimal Unbiased Estimation for Maximal Distribution (2016). arXiv:1611.07994v1
100. Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer (2002)
101. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer (1988)
102. Knight, F.H.: Risk, Uncertainty, and Profit. Houghton Mifflin, Boston (1921)
103. Kolmogorov, A.N.: Foundations of the Theory of Probability. Chelsea (1956); 2nd edn. Osnovnye poniatiya teorii veroyatnostei, "Nauka", Moscow, 1974 (1933)
104. Krylov, N.V.: Controlled Diffusion Processes. Springer (1980)
105. Krylov, N.V.: (1987) Nonlinear Parabolic and Elliptic Equations of the Second Order, Reidel. Original Russian version by Nauka, Moscow (1985)
106. Krylov, N.V.: On Shige Peng's Central Limit Theorem (2018). arXiv:1806.11238v1

107. Lévy, P.: Calcul des Probabilités. Gautier-Villars (1925)
108. Lévy, P.: Processus Stochastiques et Mouvement Brownian, 2ème éd., Gautier-Villars (1965)
109. Li, X., Lin, Y.: Localization method for stochastic differential equations driven by G-Brownian motion. Preprint (2013)
110. Li, X., Peng, S.: Stopping times and related Itô calculus with G-Brownian motion. Stoch. Process. Appl. **121**, 1492–1508 (2011)
111. Li, H., Peng, S., Soumana Hima, A.: Reflected solutions of backward stochastic differential equations driven by G-Brownian motion. Sci. China Math. **61**(1), 1–26 (2018)
112. Lions, P.L.: Some recent results in the optimal control of diffusion processes, In: Stochastic analysis, Proceedings of the Taniguchi International Symposium on Stochastic Analysis (Katata and Kyoto 1982), Kinokuniya, Tokyo (1982)
113. Lions, P.L.: Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 1: the dynamic programming principle and applications and Part 2: viscosity solutions and uniqueness. Commun. Partial Differ. Equ. **8**, 1101–1174 and 1229–1276 (1983)
114. Lyons, T.: Uncertain volatility and the risk free synthesis of derivatives. Appl. Math. Financ. **2**(2), 7–133 (1995)
115. Marinacci, M.: Limit laws for non-additive probabilities and their frequentist interpretation. J. Econ. Theory **84**(2), 145–195 (1999)
116. Maccheroni, F., Marinacci, M.: A strong law of large numbers for capacities. Ann. Probab. **33**(3), 1171–1178 (2005)
117. Neufeld, A., Nutz, M.: Nonlinear Lévy processes and their characteristics. Trans. Am. Math. Soc. **369**(1), 69–95 (2017)
118. Nisio, M.: On a nonlinear semigroup attached to optimal stochastic control. Publ. RIMS Kyoto Univ. **12**(2), 513–537 (1976)
119. Nisio, M. (1976) On stochastic optimal controls and envelope of Markovian semi–groups. In: Proceedings of the International Japan-Soviet Symposium Kyoto, pp. 297–325. Springer
120. Nutz, M.: Random G-expectations. Ann. Appl. Probab. **23**(5), 1755–1777 (2013)
121. Nutz, M., Van Handel, R.: Constructing sublinear expectations on path space. Stoch. Process. Appl. **123**(8), 3100–3121 (2013)
122. Øksendal, B.: Stochastic Differential Equations, 5th edn. Springer (1998)
123. Osuka, E.: Girsanov's formula for G-Brownian motion. Stoch. Process. Appl. **123**(4), 1301–1318 (2013)
124. Pardoux, E., Peng, S.: Adapted solutions of a backward stochastic differential equation. Syst. Control Lett. **14**(1), 55–61 (1990)
125. Pardoux, E., Rascanu, A.: Stochastic Differential Equations, Backward SDEs, Partial Differential Equations, in Stochastic Modelling and Applied Probability, vol. 69. Springer (2014)
126. Possamaï, D., Tan, X., Zhou, C.: Stochastic control for a class of nonlinear kernels and applications. Ann. Probab. **46**(1), 551–603 (2018)
127. Peng, S.: Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. Stochastics **37**(1–2), 61–74 (1991)
128. Peng, S.: A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation. Stoch. Stoch. Rep. **38**(2), 119–134 (1992)
129. Peng, S.: A nonlinear Feynman-KAC formula and applications. In: Proceedings of Symposium of System Sciences Control Theory, Chen & Yong (eds.), pp. 173–184. World Scientific (1992)
130. Peng, S.: Backward SDE and related g–expectations, In: El Karoui, M. (ed.) Backward Stochastic Differential Equations, Pitman Research Notes in Mathematics Series, No. 364, pp. 141–159 (1997)
131. Peng, S. (1997) BSDE and Stochastic Optimizations, In: Yan, J., Peng, S., Fang, S., Wu, L.M. (eds.) Topics in Stochastic Analysis, Ch.2, (Chinese vers.). Science Press, Beijing
132. Peng, S.: Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer's type. Prob. Theory Rel. Fields **113**(4), 473–499 (1999)
133. Peng, S.: Nonlinear expectation, nonlinear evaluations and risk measurs. In: Back, K., Bielecki, T.R., Hipp, C., Peng, S., Schachermayer, W. (eds.) Stochastic Methods in Finance, pp. 143–217, LNM 1856. Springer (2004)

134. Peng, S.: Filtration consistent nonlinear expectations and evaluations of contingent claims. Acta Math. Appl. Sinica, Engl. Ser. **20**(2), 1–24 (2004)
135. Peng, S.: Dynamical evaluations. C. R. Acad. Sci. Paris, Ser. I, **339**(8), 585–589 (2004)
136. Peng, S.: Nonlinear expectations and nonlinear Markov chains. Chin. Ann. Math. **26B**(2), 159–184 (2005)
137. Peng, S.: Dynamically Consistent Nonlinear Evaluations and Expectations (2005). arXiv:math/0501415
138. Peng, S.: G–Expectation, G–Brownian Motion and Related Stochastic Calculus of Itô's type. In: Benth, et al. (eds.) The Abel Symposium 2005, Abel Symposia, pp. 541–567. Springer (2007)
139. Peng, S.: Law of Large Numbers and Central Limit Theorem Under Nonlinear Expectations (2007). arXiv:0702358
140. Peng, S.: Lecture Notes: G-Brownian Motion and Dynamic Risk Measure Under Volatility Uncertainty (2007). arXiv:0711.2834
141. Peng, S.: Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. Stoch. Process. Appl. **118**(12), 2223–2253 (2008)
142. Peng, S.: A New Central Limit Theorem Under Sublinear Expectations (2008). arXiv:0803.2656
143. Peng, S.: Survey on normal distributions, central limit theorem, Brownian motion and the related stochastic calculus under sublinear expectations. Sci. China Ser. A: Math. **52**(7), 1391–1411 (2009)
144. Peng, S.: Nonlinear Expectations and Stochastic Calculus Under Uncertainty, preprint (2010). arXiv:1002.4546
145. Peng, S.: Tightness, Weak Compactness of Nonlinear Expectations and Application to CLT (2010). arxiv:1006.2541
146. Peng, S.: G-Gaussian Processes Under Sublinear Expectations and q-Brownian Motion in Quantum Mechanics (2011). arxiv:1105.1055
147. Peng, S., Song, Y.: G-expectation weighted Sobolev spaces, backward SDE and path dependent PDE. J. Math. Soc. Jpn. **67**(4), 1725–1757 (2015)
148. Peng, S., Xu, M.: The Numerical Algorithms and Simulations for BSDEs (2006). arXiv:0611864
149. Peng, S., Yang, S., Yao: Improving Value-at-Risk Prediction Under Model Uncertainty (2018). arXiv:1805.03890v3
150. Protter, P.H.: Stochastic Integration and Differential Equations. Springer (1990)
151. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer (1999)
152. Rosazza, G.E.: Some examples of risk measures via g-expectations. Insur. Math. Econ. **39**(1), 19–34 (2006)
153. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales, 1 (2nd edn.) 2. Cambridge University Press (2000)
154. Shiryaev, A.N.: Probability, 2nd edn. Springer (1995)
155. Soner, M., Touzi, N., Zhang, J.: Martingale representation theorem under G-expectation. Stoch. Process. Appl. **121**(2), 265–287 (2011)
156. Soner, H.: Mete; Touzi, Nizar; Zhang, Jianfeng Quasi-sure stochastic analysis through aggregation. Electron. J. Probab. **16**(67), 1844–1879 (2011)
157. Soner, H.M., Touzi, N., Zhang, J.: Wellposedness of second order backward SDEs. Probab. Theory Relat. Fields **153**(1–2), 149–190 (2012)
158. Soner, H.M., Touzi, N., Zhang, J.: Dual formulation of second order target problems. Ann. Appl. Probab. **23**(1), 308–347 (2013)
159. Song, Y.: Some properties on G-evaluation and its applications to G-martingale decomposition. Sci. China Math. **54**(2), 287–300 (2011)
160. Song, Y.: Properties of hitting times for G-martingales and their applications. Stoch. Process. Appl. **121**(8), 1770–1784 (2011)
161. Song, Y.: Uniqueness of the representation for G-martingales with finite variation. Electron. J. Probab. **17**(24), 1–15 (2012)

162. Song, Y.: Characterizations of processes with stationary and independent increments under G-expectation. Ann. l'Instit. H. Poincaré, Probabilités et Statistiques **49**(1), 252–269 (2013)
163. Song, Y.: Gradient Estimates for Nonlinear Diffusion Semigroups by Coupling Methods (2014). arxiv:1407.5426
164. Song, Y.: Normal Approximation by Stein's Method Under Sublinear Expectations (2017). arxiv:1711.05384
165. Song, Y.: Properties of G-martingales with finite variation and the application to G-Sobolev spaces. Stoch. Process. Appl. **129**(6), 2066–2085 (2019)
166. Song, Y.: Normal Approximation by Stein's Method Under Sublinear Expectations (2019). arXiv:1711.05384
167. Song, Y.S.: Stein's Method for Law of Large Numbers Under Sublinear Expectations (2019). arXiv:1904.04674
168. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Springer (1979)
169. Szegö, G. et al. (Eds.): Risk Measures for the 21 Centrury. Wiley (2004)
170. Vorbrink, J.: Financial markets with volatility uncertainty. J. Math. Econ. **53**, 64–78 (2014)
171. Wang, L.: On the regularity of fully nonlinear parabolic equations: II. Commun. Pure Appl. Math. **45**, 141–178 (1992)
172. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall (1991)
173. Wiener, N.: Differential space. J. Math. Phys. **2**, 131–174 (1923)
174. Xu, J., Zhang, B.: Martingale characterization of G-Brownian motion. Stoch. Process. Appl. **119**(1), 232–248 (2009)
175. Xu, J., Zhang, B.: Martingale property and capacity under G-framework. Electron. J. Probab. **15**(67), 2041–2068 (2010)
176. Yan, J.-A.: Lecture Note on Measure Theory (Chinese version). Science Press (1998)
177. Yong, J., Zhou, X.: Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer (1999)
178. Yosida, K.: Functional Analysis, 6th edn. Springer (1980)
179. Zhang, J.F.: Backward Stochastic Differential Equations, from Linear to Fully Nonlinear Theory, in Springer "Probability Theory and Stochastic Modelling 86" (2017)
180. Zhang, L.X.: Rosenthal's inequalities for independent and negatively dependent random variables under sublinear expectations with applications. Sci. China Math. **59**(4), 751–768 (2016)

Index of Symbols

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7

Author Index

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7

Subject Index

S. Peng, *Nonlinear Expectations and Stochastic Calculus under Uncertainty*,
Probability Theory and Stochastic Modelling 95,
https://doi.org/10.1007/978-3-662-59903-7