DYNAMIC CONIC FINANCE VIA BACKWARD STOCHASTIC DIFFERENCE EQUATIONS AND RECURSIVE CONSTRUCTION OF CONFIDENCE REGIONS

BY

TAO CHEN

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Applied Mathematics
in the Graduate College of the
Illinois Institute of Technology

Approved ______________________
Advisor

Approved ______________________
Co-Advisor

Chicago, Illinois
July 2016

ACKNOWLEDGMENT

I owe this thesis enormously to my advisors, Professor Tomasz R. Bielecki and Professor Igor Cialenco. Their excellent guidance and rigorous approach allow me to develop a solid foundation of intellectual and technical skills needed for my research. I am deeply grateful for their support, help, and care through my hard time. There are not enough words to express my gratitude. It is an honor to be their student.

I want to thank Professor Fred Hickernell, Professor Gregory Fasshauer, Professor Ruoting Gong, and Professor Kevin Cassel for serving on my comprehensive exam committee and thesis committee. I am grateful for their time and effort spent towards this commitment.

I appreciate the financial support from the Applied Mathematics Department, my Ph.D. advisors, and my family.

I am also thankful to all the teachers and professors I met through my studies at Illinois Institute of Technology, for the mathematics and finance knowledge they taught and suggestions they provided.

During my time at Illinois Institute of Technology, I made many good friends who helped me in various ways. Their friendship is invaluable to me.

I owe my parents for their love, unconditional support, and continuous encouragement. Without them, none of these would have been possible.

I dedicate this thesis to my parents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF SYMBOLS FOR CHAPTER 2

Symbol	Description
$\mathcal{T}$	Set of time indices
g	Driver for backward stochastic difference equations
$L^p(\mathcal{F}_t)$	L^p integrable $\mathcal{F}_t$-measurable random variables
W	Square integrable martingale process
M	Martingale process strongly orthogonal to W
(Y, Z, M)	Solution to backward stochastic difference equations
$\mathcal{E}_g$	g-expectation
$\mathcal{D}$	Set of all real valued square integrable adapted processes
ρ	Dynamic convex risk measure
α	Dynamic acceptability index
D^0	Dividend process associated with banking account
D^{ask}	Dividend process associated with holding a long position of security
D^{bid}	Dividend process associated with holding a short position of security
P^{ask}	Ex-dividend price of purchasing cash flows
P^{ask}	Ex-dividend price of selling cash flows
ϕ, φ, ξ	Trading strategies
$\widetilde{V}$	Set-up cost process
V	Liquidation value process
$\mathcal{S}$	Trading strategies initiated with zero cost
$(\mathcal{M}, P^{\text{ask}}, P^{\text{bid}})$	Market model

$\mathcal{H}^0$	Set of cash flows generated by strategies in $\mathcal{S}$
$\mathcal{H}$	Set of cash flows that can be super-hedged by strategies in $\mathcal{S}$ at zero cost
a^g	Acceptability ask price
b^g	Acceptability bid price
$\widehat{a}^g$	Acceptability ask price in extended market model
$\widehat{b}^g$	Acceptability bid price in extended market model

LIST OF SYMBOLS FOR CHAPTER 3

Symbol	Description
$\boldsymbol{\Theta}$	Space of parameters
θ	Parameters
θ^*	True parameter
$\mathbb{P}_\theta$	Probability measure associated with θ
$\mathbb{E}_\theta$	Expectation associated with θ
Z	Underlying Markov process
p_θ	Transition density function of Z
∇	Gradient vector with respect to θ
H	Hessian matrix with respect to θ
Id	Identity matrix
$I(\theta)$	Fisher information
$\tilde{\theta}$	Base estimator
$\hat{\theta}$	Quasi-asymptotically linear estimator
χ^2_d	Random variable that is chi-squared distributed with d degrees of freedom
$\mathcal{N}$	Gaussian random variable
$\mathcal{T}$	Confidence region
α	Confidence level
ρ	Correlation coefficient of bivariate Gaussian random variable

ABSTRACT

This thesis consists of two major parts, and it contributes to the fields of mathematical finance and statistics.

The contribution to mathematical finance is made via developing new theoretical results in the area of conic finance. Specifically, we have advanced dynamic aspects of conic finance by developing an arbitrage free theoretical framework for modeling bid and ask prices of dividend paying securities using the theory of dynamic acceptability indices. This has been done within the framework of general probability spaces and discrete time. In the process, we have advanced the theory of dynamic subscale invariant performance measures. In particular, we proved a representation theorem of such measures in terms of a family of dynamic convex risk measures, and provided a representation of dynamic risk measures in terms of BSΔEs.

The contribution to statistics is of fundamental importance as it initiates the theory underlying recursive computation of confidence regions for finite dimensional parameters in the context of stochastic dynamical systems. In the field of engineering, particularly in the field of control engineering, the area of recursive point estimation came to great prominence in the last forty years. However, there has been no work done with regard to recursive computation of confidence regions. To partially fill this gap, the second part of the thesis is devoted to recursive construction of confidence regions for parameters characterizing the one-step transition kernel of a time-homogeneous Markov chain.

CHAPTER 1
INTRODUCTION

Risk management and no-arbitrage pricing are among the core research and applications areas in mathematical finance. Recursive estimation of unknown parameters is one of the core research and applications areas in the statistics of stochastic processes. This thesis contributes to these areas, and it consists of two main parts. In the first part – Chapter 2 – we develop a unified pricing theory for modeling bid and ask prices of dividend paying securities in a discrete time market model with frictions. In particular, we contribute here to the so called dynamic conic finance theory. In the second part – Chapter 3 – motivated by the applications to the problem of adaptive robust hedging, we develop a methodology for recursive construction of confidence regions.

Dynamic conic finance theory was originated in Bielecki et al. [BCIR13]. As in [BCIR13], we extend here the conic finance methodology initiated by Cherny and Madan [CM10] in the static case. The idea behind conic finance is to use coherent acceptability indices to define bid/ask prices in the spirit of the no-good deal method proposed by Cochrane and Saa-Requejo [CSR00]. The coherent acceptability index is a measure of performance of financial portfolios, and it is essentially a generalization of the well known measures of performance such as the Sharpe Ratio or the Gain-to-Loss Ratio. Also in [CM10], it was shown that the conic finance pricing framework can be used as a tool to shrink the arbitrage–free price interval.

The extension of conic finance to multiperiod markets is quite delicate, especially if the underlying securities pay dividends and bear transaction costs themselves. The main challenges are due to the fact that the wealth process associated with a self-financing trading strategy is not a linear functional of trading strategies. Based on the dynamic version of coherent acceptability indices introduced in Bielecki, Cialenco, and

Zhang [BCZ14] (see also Biagini and Bion-Nadal [BBN14]), Bielecki et al. [BCIR13] studied dynamic conic finance theory. This was done for the case of discrete time and a finite probability space. There, the authors also investigated the connection between dynamic conic finance framework and classical arbitrage theory, based on the arbitrage theory for the corresponding markets developed in Bielecki, Cialenco, and Rodriguez [BCR15].

Although dynamic conic finance theory is a flexible nonlinear pricing framework, it does not fully capture the liquidity risk. More precisely, due to the scale invariance of the dynamic coherent acceptability indices, the bid/ask prices are homogeneous in the number of shares traded. However, the typical market phenomenon is that the more shares one buys the higher price per share one pays; similarly, more shares one sells, lower price per share is received. It turns out, as observed in Rosazza Gianin and Sgarra [RGS13], and Bion-Nadal [BN09], that replacing the scale invariance postulate by sub-scale invariance yields a pricing framework that captures the liquidity charge describe above. Accordingly, a 'dynamic conic finance' framework generated by sub-scale invariant acceptability indices was develped in [RGS13]. They consider a continuous time set-up for pricing terminal payoffs defined on a general probability space. Similarly to the original conic finance case of [CM10], the authors derive a representation theorem for bid/ask prices in terms of convex risk measures, and consequently in terms of solutions of some Backward Stochastic Differential Equations (BSDEs) and g-expectations.

Our work in Chapter 2 builds upon the ideas described above. By applying a time consistent, quasi-concave acceptability based approach, we develop a nonlinear pricing framework on a general probability space, in the discrete time set-up. The pricing framework leads to arbitrage-free prices, and it takes into account the market impact effect. The advantage of establishing the link between Backward Stochastic

Difference Equations (BSΔEs) and dynamic acceptability indices comes up in two ways. On one hand, we derive robust representations of bid and ask prices via g-expectations (which are solutions of BSΔEs). On the other hand, BSΔEs allow for efficient numerical computations. The results presented in this chapter are the basis for Bielecki, Cialenco, and Chen [BCC15].

Chapter 3 develops a recursive construction of confidence regions for a finite dimensional parameter in a discrete time stochastic dynamical systems. Motivated by discrete time adaptive robust stochastic control problems subject to model uncertainty (cf. Bielecki, Cialenco, Chen, Cousin, and Jeanblanc [BCC$^+$16b]), we consider in this chapter discrete time, time-homogeneous Markov chain models. The set of possible one-step transition kernels of the Markov chain models is parameterized in terms of a finite dimensional parameter θ taking values in the known parameter space. We postulate that all these models are possible descriptions of some reality, and that only one of the models, say the one corresponding to θ^*, is the adequate, or true, description of this reality. The true parameter θ^* is unknown. In Chapter 3 we derive a recursive (in time) construction of confidence regions for θ^* that satisfy some desired properties, such as desired asymptotic properties, when the time series of observations increases. The results presented in this chapter underlie Bielecki, Cialenco, and Chen [BCC16a].

There is a vast literature devoted to recursive computation, also known as on-line computation, of point estimators. It is fair to say though, that we are the first to study a recursive construction of confidence regions. The geometric idea that underlies such recursive construction is motivated by recursive representation of confidence intervals for the mean of one dimensional Gaussian distribution with known variance, and by recursive representation of confidence ellipsoids for the mean and variance of a one dimensional Gaussian distribution, where in both cases observations are generated by i.i.d. random variables. The recursive representation is straightforward in the former case, but it is not so any more in the latter one.

The key step to the recursive construction of confidence regions for θ^* is to establish an appropriate recursive point estimator and prove its asymptotic normality. For this purpose we developed an appropriate version of the so called stochastic approximation algorithm. Initiated by Robbins and Monro [RM51], the stochastic approximation methodology is one of the most widely used recursive point estimation procedures. It essentially amounts to an iterative approximation of the root of an unknown function, using an iterative sequence of approximations of that function that can be computed from the observed data.

There is a vast literature that studies consistency and asymptotic normality of point estimators obtained via stochastic approximation (cf. Khas'minskii and Nevelson [KN76]; Fabian [Fab78]; Ljung and Söderström [LS87]; Englund, Holst, and Ruppert [EHR89]; Sharia [Sha98]). These approaches for proving asymptotic normality only apply to the case when the error of approximation of the unknown function is a martingale difference process. We cannot use these approaches in our set-up as such requirement is not satisfied.

We thus proceed differently, building upon the concept of the (local) asymptotic linearity of the point estimator. This property is frequently used in the proof of asymptotic normality of estimators. Detailed discussion of the literature on this subject can be found in Barndorff-Nielsen and Sørensen [BNS94], Heyde [Hey97], and Prakasa Rao [PR99]. Unfortunately, in general, asymptotic linearity can not be reconciled with the full recursiveness of a point estimator, the property, which is the key property involved in recursive construction of confidence regions. Therefore, one of the major contributions made in Chapter 3 is that we propose the concept of quasi-asymptotic linearity, which not only applies to the fully recursive (modified) point estimator introduced in Section 3.3, but, importantly, allows us to prove the asymptotic normality of this estimator.

The recursive construction of confidence regions is needed not only for the purpose of speeding up the computation of the successive confidence regions, but, primarily, for the ability to apply the dynamic programming principle in the context of robust stochastic control methodology introduced in [BCC$^+$16b]. Other potential applications of the results of Chapter 3 are far reaching.

The thesis is organized as follows. Chapter 2 studies dynamic conic finance via BSΔEs. In Section 2.1 we establish existence and uniqueness of the solutions for a (large) class of BSΔEs, and define the g-expectation, in terms of the solution of a BSΔE. In Section 2.2 we introduce and study the notion of Dynamic Acceptability Index (DAI) and show that a DAI can be generated by a family of dynamic convex risk measures, and consequently by solutions of BSΔE with convex drivers. Section 2.3 is the main section of this chapter, and it is devoted to dynamic conic finance. We start with Section 2.3.1 by defining a market model consisting of a banking account and K securities. The prices of the securities are given by bid and ask pricing operators, and the banking account is the only asset that trades with no transaction costs. Next, in Section 2.3.2, we introduce the relevant financial definitions in this market model, such as value process, self-financing trading strategy, and arbitrage. Then, in Section 2.3.3, we define the main objects of this chapter – *the acceptability bid and ask prices* – by using the DAIs. We provide a representation of acceptability bid and ask prices in terms of g-expectations associated with the corresponding family of convex drivers. Subsequently, in Section 2.3.4 we build an arbitrage-free market model by using the acceptability bid and ask prices, and prove a series of fundamental properties of these prices. We conclude Chapter 2 with Section 2.4, where we introduce the notions of *no good deal* and arbitrage in an extended market model. We prove that there are no good deals if and only if the acceptability prices in the extended market are arbitrage free prices.

We consider the problem of recursive construction of confidence regions in Chapter 3. Section 3.1 introduces the Markov chain framework relevant for the present study, and it provides an important technical result, Proposition 3.1.1, that is crucial for recursive identification of the true Markov chain model. Section 3.2 is devoted to the recursive construction of the *base (recursive) point estimator* of the true parameter θ^*. Here we prove the strong consistency of the base point estimator. The key step to the desired recursive construction of confidence regions for θ^* is to establish the asymptotic normality of the underlying recursive point estimator. Therefore, in Section 3.3, we appropriately modify our base point estimator, so to construct a quasi-asymptotically linear (recursive) point estimator, for which we prove weak consistency and asymptotic normality. The main section of this chapter is Section 3.4, which is devoted to recursive construction of confidence regions for θ^*, and to studying their asymptotic properties. We show that confidence regions derived from quasi-asymptotically linear point estimators preserve a desired geometric structure. Such structure guarantees that we can represent the confidence regions in a recursive way in the sense that the region produced at step n is fully determined by the region produced at step $n-1$ and by the the newly arriving observation of the underlying reality. We finish Chapter 3 with Section 3.5, where illustrating examples are provided.

1.1 Contributions of the Thesis

The major contributions of Chapter 2 are:

- We develop the theory of dynamic sub-scale invariant performance measures on a general probability spaces, and in the discrete time set-up. We prove a representation theorem for such measures in terms of a family of dynamic convex risk measures. Moreover, we provide a representation of dynamic risk measures in terms of g-expectations, given as part of solutions of BSΔEs with convex drivers. We demonstrate existence and uniqueness of the solutions of the relevant BSΔEs, and we provide a comparison theorem for corresponding BSΔEs.

- We construct a market model for dividend paying securities. For this model we introduce the pricing operators that are defined in terms of dynamic acceptability indices, and we prove various properties of these operators. Using these pricing operators, we first define the bid and ask prices for the underlying securities and then we define the bid and ask prices for the derivative securities in this market market. We prove a series of important and desired properties of these prices. In particular, we show that the obtained market model is arbitrage free.

The major contributions of Chapter 3 are:

- We introduce the concept of quasi-asymptotic linearity of a point estimator of the true parameter. This concept is related to the classic definition of asymptotic linearity of a point estimator, but it requires less stringent properties, which are satisfied by the recursive point estimation scheme that we develop in Section 3.3.
- Starting from what we call the base recursive point estimation scheme, we design a quasi-asymptotically linear recursive point estimation scheme, and we prove the weak consistency and asymptotic normality of the point estimator generated by this scheme.
- We provide the relevant recursive construction of confidence regions for the true parameter. We prove that these confidence regions are weakly consistent, that

is, they converge in probability (in Hausdorff metric) to the true parameter.

CHAPTER 2

DYNAMIC CONIC FINANCE VIA BACKWARD STOCHASTIC DIFFERENCE EQUATIONS

2.1 Backward Stochastic Difference Equations

Let T be a fixed and finite time horizon, and let $\mathcal{T} := \{0, 1, \ldots, T\}$. We consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \mathbb{P})$, with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F} = \mathcal{F}_T$. Throughout, we will use the notations $L^p(\mathcal{F}_t) := L^p(\Omega, \mathcal{F}_t, \mathbb{P})$, $p \geq 1$, $t \in \mathcal{T}$. Also, we will denote by $\mathcal{X}$ the set of all adapted and square integrable stochastic processes on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \mathbb{P})$. We reserve the notation Δ for the backward difference operator $\Delta X_t := X_t - X_{t-1}$, $t \in \mathcal{T}$, where X is a stochastic process, and we also take the convention $\Delta X_0 := X_0$. In what follows, all equalities and inequalities will be understood in $\mathbb{P}$-almost surely sense. We recall that the predictable quadratic variation $\langle X \rangle_t$ of a stochastic process X is defined as a predictable process, starting at zero, and such that $X_t^2 - \langle X \rangle_t$ is a martingale with respect to filtration $\{\mathcal{F}_t\}$. It can be shown that $\Delta \langle X \rangle_t = \mathbb{E}[(\Delta X_t)^2 | \mathcal{F}_{t-1}]$.

In the sequel, the function $g : \mathcal{T} \times \Omega \times \mathbb{R} \to \mathbb{R}$ will play the role of a driver for considered Backward Stochastic Difference Equations (BSΔEs), and we will assume that it satisfies

Assumption A:

A1. the mapping $(t, \omega) \mapsto g(t, \omega, z)$ is predictable for any $z \in \mathbb{R}$;

A2. the function $z \mapsto g(t, \omega, z)$ is uniformly Lipschitz continuous, i.e. there exists a finite constant $K > 0$ such that

$$|g(t, \omega, z_1) - g(t, \omega, z_2)| \leq K|z_1 - z_2|;$$

A3. $g(t, \omega, 0) = 0$ for any $t \in \mathcal{T}$.

Throughout the chapter we will denote by $c_t(\omega)$ the Lipschitz coefficient of mapping $z \mapsto g(t,\omega,z)$, that is

$$c_t(\omega) = \operatorname{ess\,inf}\Big\{ l_t(\omega) \in L^\infty(\mathcal{F}_{t-1}) : \\ |g(t,\omega,z_1) - g(t,\omega,z_2)| \le l_t(\omega)|z_1 - z_2|,\ \omega \in \Omega,\ z_1,\ z_2 \in \mathbb{R}\Big\},$$

for any $t \in \mathcal{T}$. Note that in view of condition A2. $c_t(\omega)$ is well defined (in particular, $c_t(\omega) \le K$).

Also, we will suppress the explicit dependence on ω, if no confusion arises; for example, we may write $g(t,z)$ instead of $g(t,\omega,z)$.

We consider the following Backward Stochastic Difference Equation (BSΔE),

$$Y_t = Y_T + \sum_{t<s\le T} g(s, Z_s)\Delta\langle W\rangle_s - \sum_{t<s\le T} Z_s \Delta W_s + M_T - M_t, \quad t \in \mathcal{T}, \tag{2.1}$$

with terminal condition $Y_T \in L^2(\mathcal{F}_T)$, and where W_t is a fixed square integrable martingale process with increment ΔW_t independent of $\mathcal{F}_{t-1}$, and such that $\Delta\langle W\rangle_t \neq 0$ for any $t \in \mathcal{T}$. As already mentioned, the function g is usually referred to as the driver of the BSΔE (2.1).

As one may expect, due to its 'backward' nature, and similar to continuous time BSDEs, a solution is a triple of processes, rather than just an adapted process. Next, we give the precise definition of a solution of BSΔE (2.1).

Definition 2.1.1. *A solution to BSΔE* (2.1) *is a triple of processes* (Y, Z, M) *such that:* $(Y_t, Z_t, M_t) \in L^2(\mathcal{F}_t) \times L^2(\mathcal{F}_{t-1}) \times L^2(\mathcal{F}_t)$*, it satisfies equality* (2.1) *for all* $t \in \mathcal{T}$*, and* M *is a martingale process strongly orthogonal*[1] *to* W*.*

In general, for fixed Z and M, Y that satisfies (2.1) is not necessarily an adapted process. For example, by taking the terminal condition $X \in \mathcal{F}_T$, and putting

[1]We say that the process M is strongly orthogonal to W if the process $W_t M_t$ is a martingale.

$Z_t = M_t = 0$ for all $t \in \mathcal{T}$, then $Y_t = X$ for any $t \in \mathcal{T}$. In this case, Y is not an adapted process unless $\mathcal{F}_t = \mathcal{F}_T$, $t \in \mathcal{T}$. However, due to Galtchouk-Kunita-Watanabe decomposition (cf. [FS04, Theorem 10.18]), there exists Z and M along with Y such that they (Y, Z, M) is a solution of (2.1). We recall that W is said to have the predictable representation property, if $\mathcal{F} = \mathcal{F}^W$ implies that for any square integrable martingale process X, there exists a predictable process Z, such that $\Delta X_t = Z_t \Delta W_t$. It can be shown that if $\mathcal{F} = \mathcal{F}^W$ and W has predictable representation property, then $M_t = 0$ for all $t \in \mathcal{T}$. If W does not satisfy the predictable representation property, then a martingale process M which is orthogonal to W is indeed needed to ensure that the solution of (2.1) is well defined.

With a slight abuse of notation, we will sometimes refer to process Y as solution of BSΔE (2.1), rather than saying process Y from the solution (Y, Z, M).

It is fair to say, we believe, that the theory of backward stochastic difference equations is in many respects analogous to the theory of the backward stochastic differential equations (BSDEs). However, BSΔEs that we use in this work can not be just considered as time discretized BSDEs. So, even though existence, uniqueness, as well as other properties of the solution of BSDEs are well studied and understood (cf. [PP90, Pen97, EKQ97, BCH$^+$00]), corresponding issues for BSΔEs need to be studied in their own right. Accordingly, one goal of this chapter is to present some relevant properties of solutions of these type of equations, that tailored to our needs. Similar work on BSΔEs has already been done in [CE11], [CE10], [CS13], and [Sta09], where quite general BSΔEs were studied, in particular with drivers depending on Y_t. The driver g considered in our work does not depend on Y_t. The main reason to consider this type of drivers comes from the fact that, as proved in the remark on page 114 in [BCH$^+$00], any driver $\overline{g}(t, \omega, y, z)$ such that the mapping $(y, z) \mapsto \overline{g}(t, \omega, y, z)$ is Lipschitz continuous, $\overline{g}(t, \omega, y, 0) = 0$, and $\overline{g}(t, \omega, y, z)$ is convex with respect to y

and z, does not depend on y. Since we will apply the theory of BSΔE to dynamic risk measures which require convexity on the drivers, then we focus only on drivers of the form $g(t,z)$. It needs to be stressed though that our set-up is not just a special case of the set-up considered in the references mentioned above, primarily because we work under a set of assumptions which is not nested in the assumptions used in these papers.

2.1.1 Existence, Uniqueness and Comparison Results. In this section we will prove some general results regarding existence, uniqueness and comparison of the solutions of BSΔE (2.1).

Theorem 2.1.1. *Assume that the driver g satisfies Assumption A, and that the terminal condition $Y_T \in L^2(\mathcal{F}_T)$. Then, there exists a unique solution of equation (2.1).*

Proof. First, we will prove the existence using backward induction argument. Given $Y_T \in L^2(\mathcal{F}_T)$, we consider the equation

$$Y_{T-1} = Y_T + g(T, Z_T)\Delta\langle W\rangle_T - Z_T\Delta W_T + \Delta M_T, \tag{2.2}$$

where $Z_T, \Delta M_T$ and Y_{T-1} are the unknowns. Since $Y_T \in L^2(\mathcal{F}_T)$, then

$$\mathbb{E}[\mathbb{E}[Y_T|\mathcal{F}_{T-1}]^2] \leq \mathbb{E}[\mathbb{E}[(Y_T)^2|\mathcal{F}_{T-1}]] = \mathbb{E}[(Y_T)^2] < \infty.$$

Hence, $Y_T - \mathbb{E}[Y_T|\mathcal{F}_{T-1}]$ is a square integrable martingale difference, so it admits the Galtchouk-Kunita-Watanabe decomposition, which implies that there exist $Z_T \in \mathcal{F}_{T-1}$, $Z_T\Delta W_T \in L^2(\mathcal{F}_T)$, $\Delta M_T \in L^2(\mathcal{F}_T)$ such that $\mathbb{E}[\Delta M_T|\mathcal{F}_{T-1}] = 0$, $\mathbb{E}[\Delta M_T\Delta W_T|\mathcal{F}_{T-1}] = 0$ and

$$Y_T - \mathbb{E}[Y_T|\mathcal{F}_{T-1}] = Z_T\Delta W_T - \Delta M_T. \tag{2.3}$$

We multiply both sides of last identity by ΔW_T, and then apply $\mathbb{E}[\,\cdot\,|\mathcal{F}_{T-1}]$ to both sides. This implies that $\mathbb{E}[Y_T \Delta W_T | \mathcal{F}_{T-1}] = Z_T \Delta \langle W \rangle_T$. Therefore,

$$Z_T = \frac{\mathbb{E}[Y_T \Delta W_T | \mathcal{F}_{T-1}]}{\Delta \langle W \rangle_T}. \tag{2.4}$$

Since W has independent increments, then $\Delta \langle W \rangle_T = \mathbb{E}[\Delta W_T^2 | \mathcal{F}_{T-1}] = \mathbb{E}[\Delta W_T^2] =: C$, and by our initial assumption $C \neq 0$. Hence, we deduce

$$\mathbb{E}[Z_T^2] = \frac{\mathbb{E}[\mathbb{E}[Y_T \Delta W_T | \mathcal{F}_{T-1}]^2]}{C^2} \leq \frac{\mathbb{E}[C \mathbb{E}[Y_T^2 | \mathcal{F}_{T-1}]]}{C^2} = \frac{\mathbb{E}[Y_T^2]}{C} < \infty.$$

With Z_T and ΔM_T known, taking into account (2.3) and (2.2), we conclude that Y_{T-1} must be given by

$$Y_{T-1} = \mathbb{E}[Y_T | \mathcal{F}_{T-1}] + g(T, Z_T) \Delta \langle W \rangle_T. \tag{2.5}$$

From here, due to A1, we get that $Y_{T-1} \in \mathcal{F}_{T-1}$. Also, using A2, A3 and the fact that $Z_T \in L^2(\mathcal{F}_{T-1})$, we have

$$\begin{aligned}
\mathbb{E}[(g(T, Z_T) \Delta \langle W \rangle_T)^2] &= \mathbb{E}[g(T, Z_T)^2 \mathbb{E}[\Delta \langle W \rangle_T^2 | \mathcal{F}_{T-1}]] \\
&\leq \mathbb{E}[c_T^2 Z_T^2 \mathbb{E}[\Delta \langle W \rangle_T^2 | \mathcal{F}_{T-1}]] \\
&\leq C^2 \|c_T\|_\infty^2 \mathbb{E}[Z_T^2] < \infty
\end{aligned}$$

and thus $Y_{T-1} \in L^2(\mathcal{F}_{T-1})$. Therefore, we determined Y_{T-1}, Z_T and ΔM_T.

We continue this backward procedure for any finite number of steps smaller than T: having $Y_{t+1} \in L^2(\mathcal{F}_{t+1})$ for some fixed $t \in \{0, 1, \ldots, T-1\}$, by similar arguments as above, we find $(Y_t, Z_{t+1}, \Delta M_{t+1}) \in L^2(\mathcal{F}_t) \times L^2(\mathcal{F}_t) \times L^2(\mathcal{F}_{t+1})$, such that

$$Y_t = Y_{t+1} + g(t+1, Z_{t+1}) \Delta \langle W \rangle_{t+1} - Z_{t+1} \Delta W_{t+1} + \Delta M_{t+1}, \quad t = 0, \ldots, T-1, \tag{2.6}$$

with

$$\begin{aligned}
Y_{t+1} - \mathbb{E}[Y_{t+1} | \mathcal{F}_t] &= Z_{t+1} \Delta W_{t+1} - \Delta M_{t+1}, \\
Z_{t+1} &= \frac{\mathbb{E}[Y_{t+1} \Delta W_{t+1} | \mathcal{F}_t]}{\Delta \langle W \rangle_{t+1}}, \\
Y_t &= \mathbb{E}[Y_{t+1} | \mathcal{F}_t] + g(t+1, Z_{t+1}) \Delta \langle W \rangle_{t+1}.
\end{aligned} \tag{2.7}$$

By taking the convention $Z_0 = 0$, $M_0 = 0$, and letting $M_t := M_0 + \sum_{s=1}^{t} \Delta M_s$, we have that (2.1) holds true for all $t \in \mathcal{T}$. Moreover, M is a square integrable martingale process. Finally, since

$$\begin{aligned}\mathbb{E}[M_t W_t | \mathcal{F}_{t-1}] &= \mathbb{E}[(\sum_{s=1}^{t} \Delta M_s) W_t | \mathcal{F}_{t-1}] \\ &= \sum_{s=1}^{t-1} \Delta M_s \mathbb{E}[W_t | \mathcal{F}_{t-1}] + \mathbb{E}[\Delta M_t (W_{t-1} + \Delta W_t) | \mathcal{F}_{t-1}] \\ &= M_{t-1} W_{t-1}, \quad t = 1, \dots, T,\end{aligned}$$

we conclude that M is strongly orthogonal to W, which concludes the proof of existence of the solution.

Next, we prove the uniqueness. Assume there are two solutions (Y_t^1, Z_t^1, M_t^1) and (Y_t^2, Z_t^2, M_t^2), $t \in \mathcal{T}$, of BSΔE (2.1) with terminal condition Y_T. Then

$$Y_{T-1}^1 - Y_{T-1}^2 = (g(T, Z_T^1) - g(T, Z_T^2)) \Delta \langle W \rangle_T - (Z_T^1 - Z_T^2) + \Delta M_T^1 - \Delta M_T^2. \quad (2.8)$$

From (2.9), we have that $Z_T^1 = \frac{\mathbb{E}[Y_T \Delta W_T | \mathcal{F}_{T-1}]}{\Delta \langle W \rangle_T} = Z_T^2$. Hence, by (2.7), we get

$$\Delta M_T^1 = -Y_T + \mathbb{E}[Y_T | \mathcal{F}_{T-1}] + Z_T^1 \Delta W_T = -Y_T + \mathbb{E}[Y_T | \mathcal{F}_{T-1}] + Z_T^2 \Delta W_T = \Delta M_T^2.$$

From here, in view of (2.8), we immediately conclude that $Y_{T-1}^1 = Y_{T-1}^2$. Inductively, and by using the convention that $Z_0^1 = Z_0^2 = 0$, $M_0^1 = M_0^2 = 0$, we get that $(Y_t^1, Z_t^1, M_t^1) = (Y_t^2, Z_t^2, M_t^2)$, for any $t \in \mathcal{T}$. Therefore, the solution is unique, and this concludes the proof. $\square$

Remark 2.1.1. *Here, we make a note of one step in the proof that amounts to derivation of the following backward recurrence relations (starting from Y_T),*

$$Z_{t+1} = \frac{\mathbb{E}[Y_{t+1} \Delta W_{t+1} | \mathcal{F}_t]}{\Delta \langle W \rangle_{t+1}}, \quad (2.9)$$

$$Y_t = \mathbb{E}[Y_{t+1} | \mathcal{F}_t] + g(t+1, Z_{t+1}) \Delta \langle W \rangle_{t+1}, \quad (2.10)$$

for $t = 0, \dots, T-1$. We will make use of these formulae later on.

Remark 2.1.2. *Note that, if $\mathcal{F} = \mathcal{F}^W$ and if W has the predictable representation property, then we have that $\Delta M_t = 0$, and since $M_0 = 0$, we conclude that $M_t = 0$, $t \in \mathcal{T}$, in* (2.1). *Saying differently, if $\mathcal{F} = \mathcal{F}^W$ and W has the predictable representation property then there exists a pair of processes (Y_t, Z_t), $t \in \mathcal{T}$, that is the unique solution of equation*

$$Y_t = Y_T + \sum_{t<s\leq T} g(s, Z_s)\Delta\langle W\rangle_s - \sum_{t<s\leq T} Z_s \Delta W_s, \qquad t \in \mathcal{T}.$$

One important example of martingale W that has the predictable representation property is the symmetric random walk.

We now proceed with presenting comparison results between the solutions of BSΔEs. These results, besides being of fundamental importance for the theory of BSΔEs itself, will also serve as key ingredients for describing the risk measures developed later on in this chapter. More precisely, the version of the comparison theorem provided here is tailored to our needs and it will be used to prove the monotonicity property of proposed risk measures.

We start with an auxiliary result.

Lemma 2.1.1. *Consider BSΔE* (2.1)*, and assume that the driver g satisfies Assumption A, and that the terminal condition $Y_T \geq 0$. Also, suppose that for a fixed $t \in \mathcal{T}$, $g(s,z) = x_s z$, $s \in \{t, \dots, T\}$, where x is such that $1 + x_s \Delta W_s > 0$, for any $s \in \{t, \dots, T\}$. Then, $Y_s \geq 0$ for all $s \in \{t, \dots, T\}$. Moreover, if $Y_t = 0$ on $A \in \mathcal{F}_t$, then $Y_s = 0$ on A, for all $s \in \{t, \dots, T\}$.*

Proof. First note that Assumption A and Theorem 2.1.1 guarantee that the solution (Y, Z, M) of (2.1) exists. Fix $t \in \mathcal{T}$, assume that $g(s, z) = x_s z, s \in \{t, \dots, T\}$ and

$1 + x_s \Delta W_s > 0, s \in \{t, \dots, T\}$. Then, by (2.10) and (2.9),

$$\begin{aligned} Y_{s-1} &= \mathbb{E}[Y_s|\mathcal{F}_{s-1}] + x_s Z_s \Delta\langle W\rangle_s \\ &= \mathbb{E}[Y_s|\mathcal{F}_{s-1}] + x_s \frac{\mathbb{E}[Y_s \Delta W_s|\mathcal{F}_{s-1}]}{\Delta\langle W\rangle_s} \Delta\langle W\rangle_s \\ &= \mathbb{E}[Y_s|\mathcal{F}_{s-1}] + x_s \mathbb{E}[Y_s \Delta W_s|\mathcal{F}_{s-1}]. \end{aligned}$$

Recall that g is predictable, and since $g(s, z) = x_s z$, we have that x_s is $\mathcal{F}_{s-1}$-measurable. Thus,

$$Y_{s-1} = \mathbb{E}[Y_s(1 + x_s \Delta W_s)|\mathcal{F}_{s-1}]. \tag{2.11}$$

From here, if $Y_s \geq 0$ for some $s \in \{t+1, \dots, T\}$, using the assumption that $1 + x_s \Delta W_s > 0$, we get that $Y_{s-1} \geq 0$. Hence, since $Y_T \geq 0$, we conclude that $Y_s \geq 0$ for all $s \in \{t, \dots, T\}$.

If $\mathbb{1}_A Y_t = 0$ for some $A \in \mathcal{F}_t$, then, by the above, $\mathbb{1}_A Y_s \geq 0$ for all $s \in \{t, \dots, T\}$. Moreover, by (2.11) we also have $\mathbb{1}_A Y_t = \mathbb{E}[\mathbb{1}_A Y_{t+1}(1 + x_{t+1}\Delta W_{t+1}|\mathcal{F}_t] = 0$, and since $1 + x_{t+1}\Delta W_{t+1} 0$, we get $\mathbb{1}_A Y_{t+1}(\omega) = 0$. Similarly, we deduce that $\mathbb{1}_A Y_s = 0, s \in \{t+2, \dots, T\}$.

This concludes the proof. □

Lemma 2.1.1 depicts the comparison result for drivers g of the form $g(t, z) = x_t z$. Using this result, we will prove next the comparison theorem for a general BSΔE.

Theorem 2.1.2. *Assume that g^1, g^2 satisfy Assumption A, and Y_T^1, $Y_T^2 \in L^2(\mathcal{F}_T)$, and suppose that for every $t \in \mathcal{T}$, the following conditions hold true:*

1) $Y_T^1 \geq Y_T^2$;

2) $g^1(s, z) \geq g^2(s, z)$, $s \in \{t, \dots, T\}$, $z \in \mathbb{R}$;

3) $|c_s^1 \Delta W_s| < 1$, $s \in \{t, \dots, T\}$, *where c^1 is the Lipschitz coefficient of g^1 as defined in A2.*

Denote by $Y^i, i = 1,2$, *the solution of* (2.1), *that corresponds to driver* g^i, *and terminal condition* Y^i_T, *for* $i = 1,2$. *Then,* $Y^1_s \geq Y^2_s$, *for all* $s \in \{t, \dots, T\}$. *Moreover, the comparison is strict, in the sense that if* $Y^1_t = Y^2_t$ *on* $A \in \mathcal{F}_t$, *then* $\mathbb{1}_A Y^1_s = \mathbb{1}_A Y^2_s$ *and* $\mathbb{1}_A g^1(s, Z^2_s) = \mathbb{1}_A g^2(s, Z^2_s)$, *for all* $s \in \{t, \dots, T\}$.

Proof. We prove by backward induction that $Y^1_t \geq Y^2_t$, for every $t \in \mathcal{T}$. By condition 1), the statement holds true for $t = T$. Assume that $Y^1_t \geq Y^2_t$ for some fixed $t \in \mathcal{T}$. Then, by (2.6)

$$\begin{aligned}
Y^1_{s-1} - Y^2_{s-1} =& Y^1_s - Y^2_s + (g^1(s, Z^1_s) - g^2(s, Z^2_s))\Delta\langle W\rangle_s \\
& - (Z^1_s - Z^2_s)\Delta W_s + \Delta(M^1_s - M^2_s) \\
=& \Big[Y^1_s - Y^2_s + (g^1(s, Z^1_s) - g^1(s, Z^2_s))\Delta\langle W\rangle_s \\
& - (Z^1_s - Z^2_s)\Delta W_s + \Delta(M^1_s - M^2_s)\Big] \\
& + \Big[(g^1(s, Z^2_s) - g^2(s, Z^2_s))\Delta\langle W\rangle_s\Big] \\
=:& \widetilde{Y}_{s-1} + \overline{Y}_{s-1}.
\end{aligned}$$

Clearly, condition 2) implies that $\overline{Y}_{s-1} \geq 0$. As for $\widetilde{Y}_{s-1}$, we write it as

$$\begin{aligned}
\widetilde{Y}_{s-1} = Y^1_s - Y^2_s + & \frac{g^1(s, Z^1_s) - g^1(s, Z^2_s)}{Z^1_s - Z^2_s}(Z^1_s - Z^2_s)\Delta\langle W\rangle_s \\
& - (Z^1_s - Z^2_s)\Delta W_s + \Delta(M^1_s - M^2_s),
\end{aligned} \tag{2.12}$$

where, as usually, $0/0 = 0$. Let us now define $\widetilde{g}(t, z) = \frac{g^1(t, Z^1_t) - g^1(t, Z^2_t)}{Z^1_t - Z^2_t} z$, for $t \in \mathcal{T}$ and $z \in \mathbb{R}$. Since $g^1(t, z)$ satisfies Assumption A, then $\widetilde{g}(t, z)$ is predictable, and by Assumption A2

$$|\widetilde{g}(t, z_1) - \widetilde{g}(t, z_2)| = |\frac{g^1(t, Z^1_t) - g^1(t, Z^2_t)}{Z^1_t - Z^2_t}||z_1 - z_2| \leq c^1_t |z_1 - z_2|.$$

Moreover, $\widetilde{g}(t, 0) = 0$, and hence $\widetilde{g}(t, z)$ satisfies Assumption A, and $\widetilde{Y}_{s-1}$ is the solution to BSΔE (2.12) with driver $\widetilde{g}(t, z)$ and terminal condition $Y^1_s - Y^2_s$. Since $|\frac{g^1(t, Z^1_t) - g^1(t, Z^2_t)}{Z^1_t - Z^2_t}| \leq c^1_t$, in view of Assumption 3), $|\frac{g^1(t, Z^1_t) - g^1(t, Z^2_t)}{Z^1_t - Z^2_t}\Delta W_t| < 1$, and thus $1 + \frac{g^1(t, Z^1_t) - g^1(t, Z^2_t)}{Z^1_t - Z^2_t}\Delta W_t > 0$. From here, using Lemma 2.1.1, we get that $\widetilde{Y}_{s-1} \geq 0$.

From the above arguments, we have that $Y^1_{s-1} - Y^2_{s-1} = \widetilde{Y}_{s-1} + \overline{Y}_{s-1} \geq 0$, and consequently, by induction argument $Y^1_s \geq Y^2_s$, $s \in \{t, \dots, T\}$.

Finally, if $\mathbb{1}_A Y^1_t = \mathbb{1}_A Y^2_t$, for some $A \in \mathcal{F}_t$, then $\mathbb{1}_A \widetilde{Y}_t = \mathbb{1}_A \overline{Y}_t = 0$. By Lemma 2.1.1, $\mathbb{1}_A Y^1_{t+1} = \mathbb{1}_A Y^2_{t+1}$. Since $\mathbb{1}_A \overline{Y}_t = 0$ and $\mathbb{1}_A g^1(t+1, Z^2_{t+1}) \geq \mathbb{1}_A g^2(t+1, Z^2_{t+1})$, then $\mathbb{1}_A g^1(t+1, Z^2_{t+1}) = \mathbb{1}_A g^2(t+1, Z^2_{t+1})$. Similarly, for $t+1 < s \leq T$, one gets that $\mathbb{1}_A Y^1_s = \mathbb{1}_A Y^2_s$ and $\mathbb{1}_A g^1(s, Z^2_s) = \mathbb{1}_A g^2(s, Z^2_s)$.

The proof is complete. □

Corollary 2.1.1. *Let g be a driver that satisfies Assumption A, and let $Y^1_T, Y^2_T \in L^2(\mathcal{F}_T)$ be two terminal conditions such that $Y^1_T \geq Y^2_T$ Assume that $|c_t \Delta W_t| < 1$, $t \in \mathcal{T}$, where c_t is the Lipschitz coefficient of g. Then, $Y^1_t \geq Y^2_t$, $t \in \mathcal{T}$. Moreover, the comparison is strict in the sense that if $\mathbb{1}_A Y^1_t = \mathbb{1}_A Y^2_t$, for some $t \in \mathcal{T}$, $A \in \mathcal{F}_t$, then $\mathbb{1}_A Y^1_s = \mathbb{1}_A Y^2_s$, $s = t, \dots, T$.*

Remark 2.1.3. *Using the same ideas, one can show that Theorem 2.1.2 holds true if condition 3) is replaced by the following assumption:*

$$g^1\left(s, \frac{\mathbb{E}[Y^1 \Delta W_s | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s}\right) - g^1\left(s, \frac{\mathbb{E}[Y^2 \Delta W_s | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s}\right) \geq \frac{\mathbb{E}[Y^2 - Y^1 | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s}, \quad (2.13)$$

for $Y^1, Y^2 \in L^2(\mathcal{F}_s), Y^1 \geq Y^2$, $t < s \leq T$, and the equality reached if and only if $Y^1 = Y^2$.

Assumption (2.13) *is weaker than condition 3) in Theorem 2.1.2. Indeed, assuming that $|c^1_s \Delta W_s| < 1$, and $Y^1 \geq Y^2$, we then have that*

$$\begin{aligned}
\frac{\mathbb{E}[Y^1 - Y^2 | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s} &\geq \frac{\mathbb{E}[|c^1_s \Delta W_s| (Y^1 - Y^2) | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s} \\
&\geq c^1_s \Big| \frac{\mathbb{E}[\Delta W_s (Y^1 - Y^2) | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s} \Big| \\
&= c^1_s \Big| \frac{\mathbb{E}[Y^1 \Delta W_s | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s} - \frac{\mathbb{E}[Y^2 \Delta W_s | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s} \Big| \\
&\geq \Big| g^1\Big(s, \frac{\mathbb{E}[Y^1 \Delta W_s | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s}\Big) - g^1\Big(s, \frac{\mathbb{E}[Y^2 \Delta W_s | \mathcal{F}_{s-1}]}{\Delta \langle W \rangle_s}\Big) \Big|.
\end{aligned}$$

Hence, inequality (2.13) *holds true.*

We feel that assumption (2.13) *is less intuitive, and cumbersome, and for sake of ease of exposition, in what follows we will use assumption 3) from Theorem 2.1.2.*

Remark 2.1.4. *In [Sta09], the author proves a comparison result for BSΔEs in the limit sense, as time step goes to zero. It was assumed that the driver g and the martingale process W are such that*

$$|g(t,z_1) - g(t,z_2)| \leq K(1 + (|z_1|_\infty \vee |z_2|_\infty)^{\alpha/2})|z_1 - z_2|_\infty;$$

$$\lim_{\Delta t \to 0} \|\Delta W_t\|_\infty \Delta\langle W\rangle_t^{-(\alpha/4)} = 0, \ \alpha \in [0,2),$$

along with some additional technical conditions. In this case, by taking $\alpha = 0$, and small enough Δt, we note that g and W will satisfy condition 3) from our set-up, and hence our comparison result holds true. In either [Sta09] or our set-up, it is required to have control on both the driver g and the noise W for the comparison theorem to hold. Finally we want to mention that while the conditions from [Sta09] and conditions proposed in this work overlap (in some sense), neither one implies the other. With $\alpha = 0$, the condition on the driver from [Sta09] is stronger; while for $\alpha > 0$, the assumption on the driver is weaker but the condition satisfied by W is more restricted.

In this chapter, we will mostly work with drivers that satisfy the comparison principle, and for brevity we will call such drivers regular, with precise definition as follows:

Definition 2.1.2. *Let g be a driver that satisfies Assumption A, and let $Y_T^1, Y_T^2 \in L^2(\mathcal{F}_T)$ be two terminal conditions.*

1. *We say that the comparison result holds true for BSΔE with driver g, if $Y_T^1 \geq Y_T^2$ implies that $Y_t^1 \geq Y_t^2$, $t \in \mathcal{T}$, and the comparison is strict if $\mathbb{1}_A Y_t^1 = \mathbb{1}_A Y_t^2$ for some $t \in \mathcal{T}$, and $A \in \mathcal{F}_t$, then $\mathbb{1}_A Y_s^1 = \mathbb{1}_A Y_s^2$, $s = t, \ldots, T$.*

2. *The driver g is called regular driver if the comparison result holds true for BSΔE with driver g.*

It can be easily shown that a linear driver $g(t,z) = x_t z$ is regular if $1+x_t\Delta W_t > 0$, for every $t \in \mathcal{T}$. Also, Corollary 2.1.1 implies that a general driver g is regular if $|c_t \Delta W_t| < 1$ for every $t \in \mathcal{T}$. Next, we present several examples of regular drivers, where we consider W as a martingale process such that ΔW_t is uniformly bounded. In particular, this means that $\|\Delta W_t\|_\infty$ is $\mathcal{F}_0$ measurable.

Example 2.1.1. *Let the driver $g(t,z) = c_t|z|$, with c such that $\|c_t\|_\infty < \frac{1}{\|\Delta W_t\|_\infty}$, is a regular driver. We will show in the next section that such driver generates a family of coherent dynamic risk measures.*

Example 2.1.2. *Let us put*

$$g(t,z) = \frac{K}{(K+1)\|\Delta W_t\|_\infty} \ln(\frac{1}{3} + \frac{1}{3}e^{-z} + \frac{1}{3}e^{z}),$$

where $K \in \mathbb{R}^+$ is fixed. As such, $g(t,z)$ is predictable, $g(t,0) = 0$ and $g(t,z)$ is Lipschitz due to the fact that its derivative with respect to z takes value in $(-\frac{K}{(K+1)\|\Delta W_t\|_\infty}, \frac{K}{(K+1)\|\Delta W_t\|_\infty})$. Moreover, the Lipschitz coefficient c_t is such that $|c_t \Delta W_t| < 1$, according to the fact that $|\frac{\partial g(t,z)}{\partial z}| \leq \frac{K-1}{K\|\Delta W_t\|_\infty}$. Thus, the driver g is regular, and the corresponding BSDE has the following form

$$Y_t = Y_T + \sum_{t<s\leq T} \frac{K}{(K+1)\|\Delta W_t\|_\infty} \ln(\frac{1}{3}+\frac{1}{3}e^{-Z_s}+\frac{1}{3}e^{Z_s})\Delta\langle W\rangle_t - \sum_{t<s\leq T} Z_s\Delta W_s + M_T - M_t.$$

We will see in the next section, this BSΔE plays an important role in our study, and it is related to so called convex dynamic risk measures.

2.1.2 g-Expectations. In the seminal paper [Pen97], the author introduced a relationship between solutions of BSDEs (in continuous time) and so called nonlinear expectations or g-expectations (see also [CHMP02]). Later, the theory of nonlinear expectations was successfully applied to some problems from mathematical finance in the context of theory of risk measures. For more details we refer the reader to [RG06, BEK07, RGS13] and references therein. Similar approach can be adopted for the case of BSΔEs, which will be the main goal of this section.

Towards this end, we will assume that the driver g is a regular driver, and that the terminal condition $X \in L^2(\mathcal{F}_T)$. As before, we denote by (Y_t, Z_t, M_t), $t \in \mathcal{T}$, the solution of the corresponding BSΔE. Analogous to the existing literature on BSDEs (cf. [Pen97]), we define *the conditional g-expectation* $\mathcal{E}_g\big[X\big|\mathcal{F}_t\big]$ of a random variable X given $\mathcal{F}_t$ as $\mathcal{E}_g\big[X\big|\mathcal{F}_t\big] := Y_t$.

In what follows, it will be convenient to view the space $L^2(\mathcal{F}_T)$ as an $L^\infty(\mathcal{F}_t)$-module, for every fixed $t \in \mathcal{T}$; saying differently, the random variables from $L^\infty(\mathcal{F}_t)$ will play the role of scalars for the linear space $L^2(\mathcal{F}_T)$. This is a special case of considering $L^0(\mathcal{G})$-module, or simply L^0-module, where the scalars are random variables that are measurable with respect to some σ-algebra $\mathcal{G}$. For more details on general theory of L^0-modules, and their relationship to theory of risk and performance measures, we refer the reader to [FKV09, KV09, BCDK15].

Remark 2.1.5. *Throughout the chapter we will use the following result, which follows immediately from the uniqueness of the solutions and backward nature of BSΔEs. Let g be a driver that satisfies Assumption A, and assume that there exists a triplet of processes (Y', Z', M') such that*

$$Y'_u = X + \sum_{u<s\leq T} g(s, Z'_s)\Delta\langle W\rangle_s - \sum_{u<s\leq T} Z'_s\Delta W_s + M'_T - M'_u, \quad u = t, \ldots, T,$$

where $Y'_u \in L^2(\mathcal{F}_u)$, $Z'_u \in L^2(\mathcal{F}_{u-1})$, $M'_u \in L^2(\mathcal{F}_u)$, and $\mathbb{E}[M'_u W_u|\mathcal{F}_{u-1}] = M'_{u-1}W_{u-1}$. Then, $Y'_t = \mathcal{E}_g[X|\mathcal{F}_t]$.

Next proposition provides some fundamental properties of g-expectations, such as monotonicity, tower property, convexity when driver is convex, and homogeneity when driver is homogenous.

Proposition 2.1.1. *For any regular driver g, the conditional g-expectation satisfies the following properties:*

(i) $\mathcal{E}_g[\mu|\mathcal{F}_t] = \mu$, *for any* $\mu \in \mathbb{R}$, $t \in \mathcal{T}$;

(ii) *if* $X^1 \geq X^2$, $X^1, X^2 \in L^2(\mathcal{F}_T)$, *then* $\mathcal{E}_g\big[X^1\big|\mathcal{F}_t\big] \geq \mathcal{E}_g\big[X^2\big|\mathcal{F}_t\big]$, *for any* $t \in \mathcal{T}$. *Moreover, if* $\mathbb{1}_A\mathcal{E}_g\big[X^1\big|\mathcal{F}_t\big] = \mathbb{1}_A\mathcal{E}_g\big[X^2\big|\mathcal{F}_t\big]$, *for some* $t \in \mathcal{T}$, *and* $A \in \mathcal{F}_t$, *then* $\mathbb{1}_A X^1 = \mathbb{1}_A X^2$;

(iii) $\mathcal{E}_g\big[\mathcal{E}_g\big[X\big|\mathcal{F}_s\big]\big|\mathcal{F}_t\big] = \mathcal{E}_g\big[X\big|\mathcal{F}_{s\wedge t}\big]$, *for any* $X \in L^2(\mathcal{F}_T)$, $s, t \in \mathcal{T}$;

(iv) $\mathcal{E}_g\big[\mathbb{1}_A X\big|\mathcal{F}_t\big] = \mathbb{1}_A\mathcal{E}_g\big[X\big|\mathcal{F}_t\big]$, *for any* $X \in L^2(\mathcal{F}_T)$, $A \in \mathcal{F}_t$, $t \in \mathcal{T}$;

(v) $\mathcal{E}_g\big[X + m\big|\mathcal{F}_t\big] = \mathcal{E}_g\big[X\big|\mathcal{F}_t\big] + m$, *for any* $X \in L^2(\mathcal{F}_T)$, $m \in L^2(\mathcal{F}_t)$, $t \in \mathcal{T}$;

(vi) *if* $g(u, \omega, \cdot)$ *is convex, for any* $u \in \{t, \ldots, T\}$, $\omega \in \Omega$, z_1, $z_2 \in \mathbb{R}$, *that is*

$$g(u, \omega, \mu z_1 + (1-\mu)z_2) \leq \mu g(u, \omega, z_1) + (1-\mu)g(u, \omega, z_2), \quad \mu \in \mathbb{R},\ 0 \leq \mu \leq 1,$$

then

$$\mathcal{E}_g[\lambda X^1 + (1-\lambda)X^2|\mathcal{F}_t] \leq \lambda\mathcal{E}_g[X^1|\mathcal{F}_t] + (1-\lambda)\mathcal{E}_g[X^2|\mathcal{F}_t],$$

for any $X^1, X^2 \in L^2(\mathcal{F}_T)$, $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$.

(vii) *if* $g(u, \omega, \cdot)$ *is homogeneous, for any* $u \in \{t, \ldots, T\}$, $\omega \in \Omega$, *that is*

$$g(u, \omega, \mu z) = \mu g(u, \omega, z), \quad z,\ \mu \in \mathbb{R},$$

then

$$\mathcal{E}_g[\lambda X|\mathcal{F}_t] = \lambda\mathcal{E}_g[X|\mathcal{F}_t], \quad X \in L^2(\mathcal{F}_T),\ \lambda \in L^\infty(\mathcal{F}_t).$$

Proof. (i) If $Y_T = \mu \in \mathbb{R}$, then according to (2.4), we get that

$$Z_T = \frac{\mathbb{E}[Y_T \Delta W_T | \mathcal{F}_{T-1}]}{\Delta \langle W \rangle_T} = \mu \frac{\mathbb{E}[\Delta W_T | \mathcal{F}_{T-1}]}{\Delta \langle W \rangle_T} = 0.$$

Hence, by (2.5) and Assumption A3, we have that

$$Y_{T-1} = \mathbb{E}[Y_T | \mathcal{F}_{T-1}] + g(T, Z_T) \Delta \langle W \rangle_T = \mathbb{E}[Y_T | \mathcal{F}_{T-1}] = \mu.$$

Inductively, in view of (2.9) and (2.10), we have that $Y_t = \mu$ for all $t \in \mathcal{T}$. Hence, $\mathcal{E}_g[\mu | \mathcal{F}_t] = \mu$, $t \in \mathcal{T}$.

(ii) It follows imediatly from Theorem 2.1.2.

(iii) Assume that $t \leq s$, and let (Y, Z, M) be the solution of BSΔE with terminal condition X. Then,

$$Y_u = X + \sum_{u < r \leq T} g(r, Z_r) \Delta \langle W \rangle_r - \sum_{u < r \leq T} Z_r \Delta W_r + M_T - M_u, \quad u \in \mathcal{T}.$$

By considering $u = t$ and $u = s$, we immediately get

$$Y_t = Y_s + \sum_{t < r \leq s} g(r, Z_r) \Delta \langle W \rangle_r - \sum_{t < r \leq s} Z_r \Delta W_r + M_s - M_t. \tag{2.14}$$

Next, we consider the BSΔE with driver g, terminal time s and terminal condition Y_s. By (2.14) and definition of g-expectation, we have that $Y_t = \mathcal{E}_g[Y_s | \mathcal{F}_t]$, which implies that $\mathcal{E}_g[X | \mathcal{F}_t] = \mathcal{E}_g[\mathcal{E}_g[X | \mathcal{F}_s] | \mathcal{F}_t]$.

Now, let us assume that $t > s$. For $t = T$, then property follows from the definition of g-expectation For $t < T$, we us consider the BSΔE with driver g and terminal condition (at time T) $\mathcal{E}_g[X | \mathcal{F}_s]$. In view of (2.4), we have that

$$Z_T = \frac{\mathbb{E}[Y_T \Delta W_T | \mathcal{F}_{T-1}]}{\Delta \langle W \rangle_T} = \mathcal{E}_g[X | \mathcal{F}_s] \frac{\mathbb{E}[\Delta W_T | \mathcal{F}_{T-1}]}{\Delta \langle W \rangle_T} = 0.$$

Hence, by (2.5) and Assumption A3, it is true that

$$Y_{T-1} = \mathbb{E}[Y_T | \mathcal{F}_{T-1}] + g(T, Z_T) \Delta \langle W \rangle_T = \mathbb{E}[\mathcal{E}_g[X | \mathcal{F}_s] | \mathcal{F}_{T-1}] = \mathcal{E}_g[X | \mathcal{F}_s].$$

Inductively, using (2.9) and (2.10), we conclude that $Y_t = \mathcal{E}_g[X|\mathcal{F}_s]$. Consequently, by the definition of g-expectation, we finally get that $\mathcal{E}_g[\mathcal{E}_g[X|\mathcal{F}_s]|\mathcal{F}_t] = \mathcal{E}_g[X|\mathcal{F}_s]$.

(iv) Fix $t \in \mathcal{T}$, $X \in L^2(\mathcal{F}_T)$, $A \in \mathcal{F}_t$, and let (Y, Z, M) be the solution of BSΔE with terminal condition X. Note that $\mathbb{1}_A g(u, Z_u) = g(u, \mathbb{1}_A Z_u)$, $u = t+1, \dots, T$, and thus

$$\mathbb{1}_A Y_u = \mathbb{1}_A X + \sum_{u<s\leq T} g(s, \mathbb{1}_A Z_s)\Delta\langle W\rangle_s - \sum_{u<s\leq T} \mathbb{1}_A Z_s \Delta W_s + \mathbb{1}_A M_T - \mathbb{1}_A M_u, \quad (2.15)$$

for every $u = t, \dots, T$. Also note that since $A \in \mathcal{F}_t$, and $u \geq t+1$, we have that $\mathbb{1}_A Y_u \in L^2(\mathcal{F}_u)$, $\mathbb{1}_A Z_u \in L^2(\mathcal{F}_{u-1})$, and $\mathbb{1}_A M_u \in L^2(\mathcal{F}_u)$, for any $u = t+1, \dots, T$. Due to that fact that M and W are orthogonal, we note that

$$\mathbb{E}[\mathbb{1}_A M_u W_u|\mathcal{F}_{u-1}] = \mathbb{1}_A \mathbb{E}[M_u W_u|\mathcal{F}_u] = \mathbb{1}_A M_{u-1} W_{u-1}.$$

Therefore, in view of (2.15), and Lemma 2.1.5, we have that $1_A Y_t = \mathcal{E}_g[\mathbb{1}_A X|\mathcal{F}_t]$, which implies that $\mathbb{1}_A \mathcal{E}_g[X|\mathcal{F}_t] = \mathcal{E}_g[\mathbb{1}_A X|\mathcal{F}_t]$. (v) If $(Y_u, Z_u, M_u), u \in \mathcal{T}$, be the solution of BS$\Delta$E with terminal condition X, then

$$\begin{aligned} Y_u + m = X + m &+ \sum_{u<s\leq T} g(s, Z_s)\Delta\langle W\rangle_s \\ &- \sum_{u<s\leq T} Z_s \Delta W_s + M_T - M_u, \quad u = t, \dots, T, \end{aligned} \quad (2.16)$$

for any $m \in L^2(\mathcal{F}_t)$. Clearly $Y_u + m \in L^2(\mathcal{F}_u)$, and hence, by Lemma 2.1.5 and (2.16), we conclude that $Y_t + m = \mathcal{E}_g[Y_T + m|\mathcal{F}_t]$, which implies that

$$\mathcal{E}_g[X|\mathcal{F}_t] + m = X + \beta = \mathcal{E}_g[X + m|\mathcal{F}_t].$$

(vi) Let (Y_u^1, Z_u^1, M_u^1), and respectively Y_u^2, Z_u^2, M_u^2, $u \in \mathcal{T}$, be the solution of BSΔE with terminal condition X^1, and respectively X^2. Assuming that $g(u, \omega, \cdot)$ is convex, and using identity (2.1), we have that

$$
\begin{aligned}
\lambda Y_t^1 + (1-\lambda)Y_t^2 =& \lambda X^1 + (1-\lambda)X^2 + \lambda \sum_{t<u\leq T} g(u,\omega,Z_u^1)\Delta\langle W\rangle_u \\
& + (1-\lambda)\sum_{t<u\leq T} g(u,\omega,Z_u^2)\Delta\langle W\rangle_u - \lambda \sum_{t<u\leq T} Z_u^1 \Delta W_u \\
& - (1-\lambda)\sum_{t<u\leq T} Z_u^2\Delta W_u + \lambda(M_T^1 - M_t^1) + (1-\lambda)(M_T^2 - M_t^2) \\
\geq & \lambda X^1 + (1-\lambda)X^2 + \sum_{t<u\leq T} g(u,\omega,\lambda Z_u^1 + (1-\lambda)Z_u^2)\Delta\langle W\rangle_u \\
& - \sum_{t<u\leq T}(\lambda Z_u^1 + (1-\lambda)Z_u^2)\Delta\langle W\rangle_u + \lambda M_T^1 + (1-\lambda)M_T^2 \\
& - \lambda M_t^1 - (1-\lambda)M_t^2, \qquad (2.17)
\end{aligned}
$$

where $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$. Next we consider the process Y' defined as follows

$$
\begin{aligned}
Y'_u =& \lambda X^1 + (1-\lambda)X^2 + \sum_{u<s\leq T} g(s,\lambda Z_s^1 + (1-\lambda)Z_s^2)\Delta\langle W\rangle_s \\
& - \sum_{u<s\leq T}(\lambda Z_s^1 + (1-\lambda)Z_s^2)\Delta\langle W\rangle_s + \lambda M_T^1 + (1-\lambda)M_T^2 \\
& - \lambda M_u^1 - (1-\lambda)M_u^2, \quad u = t,\dots,T.
\end{aligned}
$$

Clearly $Y'_u \in L^2(\mathcal{F}_u)$, $\lambda Z_u^1 + (1-\lambda)Z_u^2 \in L^2(\mathcal{F}_{u-1})$, $\lambda M_u^1 + (1-\lambda)M_u^2 \in L^2(\mathcal{F}_u)$, and $\mathbb{E}[(\lambda M_u^1 + (1-\lambda)M_u^2)W_u|\mathcal{F}_{u-1}] = (\lambda M_{u-1}^1 + (1-\lambda)M_{u-1}^2)W_{u-1}$, for any $u = t+1,\dots,T$. Therefore, by Lemma 2.1.5, we have that $Y'_t = \mathcal{E}_g[\lambda Y_T^1 + (1-\lambda)Y_T^2|\mathcal{F}_t]$, combined with (2.17) concludes the proof.

(vii) The proof is similar to the proof of (vi) and we omit it here.

□

In what follows, we will call a driver g convex, if $g(t,\omega,\cdot)$ is convex, and g positive homogeneous, if $g(t,\omega,\cdot)$ is positive homogeneous, for any $t \in \mathcal{T}$, $\omega \in \Omega$. Also, we will simply say that $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is convex (rather than convex in L^∞-module sense) if

$$\mathcal{E}_g[\lambda X^1 + (1-\lambda)X^2|\mathcal{F}_t] \leq \lambda\mathcal{E}_g[X^1|\mathcal{F}_t] + (1-\lambda)\mathcal{E}_g[X^2|\mathcal{F}_t],$$

for any $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$.

Proposition 2.1.1 shows that g-expectation (or nonlinear expectation) shares many properties with usual conditional expectation. However, as name suggests, generally speaking it is not linear. The next two results show that the g-expectation is linear if and only if the driver is regular and linear.

Proposition 2.1.2. *Assume that g is a regular linear driver. Then $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is linear. Moreover, there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $\mathbb{E}_\mathbb{Q}[X|\mathcal{F}_t] = \mathcal{E}_g[X|\mathcal{F}_t]$ for all $X \in L^2(\mathcal{F}_T)$.*

Proof. Since $g(t,z)$ is regular, then $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is well defined and $\mathcal{E}_g[0|\mathcal{F}_t] = 0$. Assuming that (Y^i, Z^i, M^i), $i = 1,2$, are the solutions of BSΔE (2.1) with terminal condition X^i, $i = 1,2$, for any $a, b \in \mathcal{F}_t$, we get

$$\begin{aligned} aY_t^1 + bY_t^2 =& aY_T^1 + bY_T^2 + \sum_{t<s\leq T} x_s(aZ_s^1 + bZ_s^2)\Delta\langle W\rangle_s - \sum_{t<s\leq T}(aZ_s^1 + bZ_s^2)\Delta W_s \\ &+ (aM_T^1 + bM_T^2) - (aM_t^1 + bM_t^2). \end{aligned}$$

Moreover, $aM_t^1 + bM_t^2$ is orthogonal to W_t, and therefore, $aY_t^1 + bY_t^2$ is the solution of BSΔE with terminal condition $aY_T^1 + bY_T^2$. Hence, linearity of $\mathcal{E}_g[\cdot|\mathcal{F}_t]$ follows.

For any $A \in \mathcal{F}$ we define $\mathbb{Q}(A) := \mathcal{E}_g[\mathbb{1}_A]$. First, we will verify that $\mathbb{Q}$ is a probability measure. Since $1 + x_t\Delta W_t > 0$, then according to Lemma 2.1.1 we have that $\mathbb{Q}(A) = \mathcal{E}_g[\mathbb{1}_A] \geq 0$. If $A = \emptyset$, then $\mathbb{1}_A = 0$ almost surely and therefore $\mathbb{Q}(A) = \mathcal{E}_g[0] = 0$. If $A = \Omega$, then $\mathbb{1}_A = 1$ and hence $\mathbb{Q}(A) = \mathcal{E}_g[1] = 1$. Let $(A_i)_{i=1}^\infty$ be such that $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, and $A_i \bigcap A_j = 0$, for $i \neq j$. Let (Y, Z, M) be the solution of BSΔE with terminal condition $\sum_{i=1}^{\infty} \mathbb{1}_{A_i}$. Then,

$$\mathbb{Q}(\bigcup_{i=1}^{\infty} A_i) = \mathcal{E}_g[\sum_{i=1}^{\infty} \mathbb{1}_{A_i}] = \sum_{i=1}^{\infty} \mathbb{1}_{A_i} + \sum_{t<s\leq T} x(s) Z_s \Delta\langle W\rangle_s - \sum_{t<s\leq T} Z_s \Delta W_s + M_T - M_t.$$

By (2.4), we also have that

$$Z_T = \frac{\mathbb{E}[\sum_{i=1}^{\infty} \mathbb{1}_{A_i} \Delta W_T | \mathcal{F}_{T-1}]}{\Delta\langle W\rangle_T}.$$

From here, by Dominated Convergence Theorem, and again by (2.4), we continue

$$Z_T = \sum_{i=1}^{\infty} \frac{\mathbb{E}[\mathbb{1}_{A_i} \Delta W_T | \mathcal{F}_{T-1}]}{\Delta\langle W\rangle_T} = \sum_{i=1}^{\infty} Z_T^i,$$

where Z_T^i is the part of the solution corresponding to terminal condition $\mathbb{1}_{A_i}$, $i \in \mathbb{N}$. Similarly, we have that $Y_{T-1} = \sum_{i=1}^{\infty} Y_{T-1}^i$. Inductively, we have that $Y_0 = \sum_{i=1}^{\infty} Y_0^i$, or $\mathbb{Q}[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} \mathbb{Q}[A_i]$. Therefore, $\mathbb{Q}$ is a probability measure.

Next, we will show that $\mathbb{Q}$ is equivalent to $\mathbb{P}$. If $A \in \mathcal{F}$, and $\mathbb{P}(A) = 0$, then $\mathbb{1}_A = 0$, $\mathbb{P}$-a.s., and thus $\mathbb{Q}(A) = \mathcal{E}_g[\mathbb{1}_A] = 0$. Conversely, if $\mathbb{Q}(A) = 0$, then, by Lemma 2.1.1, $\mathbb{1}_A = 0$ $\mathbb{P}$-a.s., and hence $\mathbb{P}(A) = 0$. Thus, $\mathbb{Q}$ is equivalent to $\mathbb{P}$.

We are left to show that $\mathbb{E}_{\mathbb{Q}}[X|\mathcal{F}_t] = \mathcal{E}_g[X|\mathcal{F}_t]$, for any $X \in L^2(\mathcal{F})$, $t \in \mathcal{T}$. First, we prove the statement for $t = 0$. For a simple random variable $X = \sum_{i=1}^{n} \mathbb{1}_{A_i} a_i \in L^2(\mathcal{F})$, with $n \in \mathbb{N}$, $A_i \in \mathcal{F}$, $a_i \in \mathbb{R}$, using linearity of g-expectations proved above, we have that $\mathbb{E}_{\mathbb{Q}}[X] = \mathcal{E}_g[X]$. For a general random variable we use the standard approximation procedure. If $X \in L^2_+(\mathcal{F})$, then there exists an increasing, and positive sequence of simple random variable $0 \leq X_1 \leq X_2 \leq \ldots \leq X_n \leq \ldots \leq X$ such that $\lim_{n\to\infty} X_n = X$, and

$$\mathbb{E}_{\mathbb{Q}}[X] = \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}}[X_n] = \lim_{n\to\infty} \mathcal{E}_g[X_n].$$

Using similar procedure as in the first part of the proof, one can show that $Y_0^{X_n} \xrightarrow{a.s.} Y_0^X$, and thus $\mathcal{E}_g[X] = \mathbb{E}_{\mathbb{Q}}[X]$. Finally for the general case, $X \in L^2(\mathcal{F})$, it is enough to split X into its positive and negative part.

From here, and Proposition 2.1.1, for any $t \in \{1, \dots, T\}$, $A \in \mathcal{F}_t$, and $X \in L^2(\mathcal{F})$, we also have

$$\mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A \mathcal{E}_g[X|\mathcal{F}_t]] = \mathcal{E}_g[\mathbb{1}_A \mathcal{E}_g[X|\mathcal{F}_t]] = \mathcal{E}_g[\mathcal{E}_g[\mathbb{1}_A X|\mathcal{F}_t]] = \mathcal{E}_g[\mathbb{1}_A X] = \mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A X],$$

which implies that $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is equal to the usual conditional expectation.

This concludes the proof.

□

Proposition 2.1.3. *Assume that g is a regular driver, and $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is linear for any $t \in \mathcal{T}$. Then, there exists a process x_t, such that x_t is predictable, $|x_t| \le c_t$, where c_t is the Lipschitz coefficient of $g(t, \cdot)$, and $g(t, z) = x_t z$, $t \in \mathcal{T}$.*

Proof. Fix $t \in \mathcal{T}$. Let us consider a BSΔE from time $t-1$ to terminal time t, and with driver g. For any fixed z^1, $z^2 \in \mathbb{R}$, it is straightforward to verify that $Z_t^1 = az^1$, $M_t^1 = 0$, $Y_{t-1}^1 = g(t, az^1)\Delta\langle W\rangle_t$, and respectively $Z_t^2 = bz^2$, $M_t^2 = 0$, $Y_{t-1}^2 = g(t, bz^2)\Delta\langle W\rangle_t$, is the solution to this BSΔE with terminal values $az^1\Delta W_t$, and respectively $bz^2\Delta W_t$, where a, b are arbitrary real numbers. Consequently, $Z_t^0 = az^1 + bz^2$, $M_t^0 = 0$, $Y_{t-1}^0 = g(t, az^1 + bz^2)\Delta\langle W\rangle_t$ is the solution to the BSΔE with terminal condition $(az^1 + bz^2)\Delta W_t$.

By the definition of g-expectation, and using its linearity, we have

$$\begin{aligned} g(t, az^1 + bz^2)\Delta\langle W\rangle_t &= \mathcal{E}_g[(az^1 + bz^2)\Delta W_t|\mathcal{F}_{t-1}] \\ &= a\mathcal{E}_g[z^1\Delta W_t|\mathcal{F}_{t-1}] + b\mathcal{E}_g[z^2\Delta W_t|\mathcal{F}_{t-1}] \\ &= ag(t, z^1)\Delta\langle W\rangle_t + bg(t, z^2)\Delta\langle W\rangle_t. \end{aligned}$$

From here, since $\Delta\langle W\rangle_t \neq 0$, and since z^1 and z^2 were arbitrarily chosen, we have that

$$g(t, \omega, az^1 + bz^2) = ag(t, \omega, z^1) + bg(t, \omega, z^2), \quad \omega \in \Omega,\ z^1,\ z^2,\ a,\ b \in \mathbb{R}.$$

Note that since $g(t, \omega, \cdot)$ is Lipschitz, hence continuous, there exists a random variable $x_t(\omega)$ such that $g(t, \omega, z) = x_t(\omega)z$, $\omega \in \Omega$, $z \in \mathbb{R}$.

Finally, recall that g is predictable, and thus x_t is predictable too. Moreover, by Assumption A2,

$$|x_t(\omega)z_1 - x_t(\omega)z_2| = |g(t, \omega, z_1) - g(t, \omega, z_2)| \leq c_t(\omega)|z_1 - z_2|$$

for any $\omega \in \Omega$, $z_1, z_2 \in \mathbb{R}$, $t \in \mathcal{T}$. Therefore, $|x_t| \leq c_t$, and this completes the proof.

□

2.2 Dynamic Convex Risk Measures and Dynamic Acceptability Indices via g-Expectation

In this section we will explore the connections between g-Expectation and Dynamic Convex Risk Measures (DCRMs), and subsequently the relationship between DCRMs and Dynamic Acceptability Indices (DAIs).

In the seminal paper [ADEH99], the authors proposed an axiomatic approach to defining risk measures that are meant to give a numerical value of the riskiness of a given financial contract or portfolio. Since then, an extensive body of work was devoted to exploration of the axiomatic approach to risk measures, and it is beyond the scope of this thesis to list all the relevant literature on this subject. We refer the reader to [DK13] for an excellent overview of the static (one period of time) risk measures, as well as to the survey paper [AP11] on dynamic risk measures. The values of risk measures can be interpreted as the capital requirement for the purpose of regulating the risk assumed by market participants (typically, by banks). In particular, the risk measures are aimed at quantifying risk, similarly to what was the primary objective of the well-known Value-At-Risk (V@R). Following a similar axiomatic approach, Cherny and Madan [CM09] introduced the notion of coherent acceptability index – function defined on a set of random variables that takes positive values and that is monotone, quasi-concave, and scale invariant. Coherent acceptability indices can be viewed as generalizations of performance measures such as Sharpe Ratio, Gain to Loss Ratio, Risk Adjusted Return on Capital. Coherent acceptability indices appear to be a tool very well tailored to assessing both risk and reward of a given cash flow. The dynamic version of coherent acceptability indices was introduced and studied in [BCZ14].

For a robust representations of general dynamic quasi-concave, monotone and local maps see, for instance, [BCDK15].

As it was shown in [RG06] there is a direct connection between convex risk measures and nonlinear expectations, and consequently there exists a direct link between convex risk measures and BSDEs. These connections were further studied in [CE10], [ESC15], and [Sta09] for the case of discrete time set-ups, thus establishing a relationship between BSΔEs and DRM for terminal cash flows (random variables).

In [CM09, BCZ14] the authors proved that any (dynamic) coherent acceptability index can be associated with a family of (dynamic) coherent risk measures. In [RGS13] the authors study the relationship between dynamic *sub-scale* invariant performance measures and dynamic *convex* risk measures and their connections to BSΔEs. The aim of this section is to develop a unified framework for assessing the risk and performance of cash-flows in a dynamic, discrete time set-up. It is well known that one of the key properties of dynamic risk and performance measures is their time consistency, and in this chapter, we also pay special attention to this property. We refer the reader to [BCP14, BCP15] for a thorough discussion of various forms of time consistency of risk/performance measures.

Besides the usual applications of these measures to risk management, and assessment of portfolio's performance, we will show in next sections that dynamic sub-scale invariant acceptability indices, nonlinear expectations and theory of BSΔEs can be successfully applied to build a general, arbitrage free, nonlinear pricing methodology in complex derivative markets.

For the sake of consistency, we will follow the set-up from [BCIR13, BCZ14, BCR15] adapted to a general probability space. A cash-flow, also called a dividend process, denoted as $D = \{D_t\}_{t=0}^T$, is any real valued, square integrable, stochastic process adapted to filtration $\mathbb{F}$. The set of all cash-flows is denoted by $\mathcal{D}$, that is

$$\mathcal{D} := \big\{(D_t)_{t=0}^T : (D_t)_{t=0}^T \text{ is an adapted process, } D_t \in L^2(\mathcal{F}_t),\ t \in \mathcal{T}\big\}.$$

From financial point of view, an element $D \in \mathcal{D}$ should be interpreted as a cash-flow associated with a portfolio, or with a general financial instrument, such that the amount D_t is received/paid by the holder at time $t \in \mathcal{T}$. For more details, and for more general set-up, please see next chapter.

For any $D \in \mathcal{D}$, $\lambda \in L^\infty(\mathcal{F}_t)$, we define the following multiplicative operator

$$\lambda \cdot_t D := (0, \dots, 0, \lambda D_t, \dots, \lambda D_T).$$

Note that for any $t \in \mathcal{T}$, the set $\mathcal{D}$ is closed under the multiplication $\cdot_t$. We also define an order $\succeq_t$ on $\mathcal{D}$, and say that $D^1 \succeq_t D^2$, whenever $\sum_{s=t}^T D_s^1 \geq \sum_{s=t}^T D_s^2$, $t \in \mathcal{T}$. Hence, $D^1 \succeq_t D^2$ if the future cumulative cash-flow is larger.

Remark 2.2.1. *In what follows, we will simply write λD instead of $\lambda \cdot_t D$. That means for process $D \in \mathcal{D}$, $\lambda D = \lambda \cdot_t D$, and if X is a random variable, then λX is understood as multiplication of random variables λ and X.*

2.2.1 Dynamic Convex Risk Measures via g-Expectation. We start by recalling the definition of Dynamic Convex/Coherent Risk Measures.

Definition 2.2.1 (Dynamic Convex Risk Measure)**.** *A dynamic convex risk measure is a function* $\rho : \mathcal{T} \times \mathcal{D} \times \Omega \to \mathbb{R}$ *that satisfies the following properties:*

R1. ***Adapted.*** $\rho_t(D)$ *is* $\mathcal{F}_t$*-measurable, for any* $t \in \mathcal{T}, D \in \mathcal{D}$.

R2. ***Local.*** $\mathbb{1}_A \rho_t(D) = \mathbb{1}_A \rho_t(\mathbb{1}_A D)$, *for any* $t \in \mathcal{T}, A \in \mathcal{F}_t, D \in \mathcal{D}$.

R3. ***Convex.*** $\rho_t(\lambda D + (1-\lambda)D') \leq \lambda \rho_t(D) + (1-\lambda)\rho_t(D')$, *for any* $t \in \mathcal{T}$, $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$, $D, D' \in \mathcal{D}$.

R4. ***Monotone.*** *If* $D \succeq_t D'$ *for some* $t \in \mathcal{T}$, $D, D' \in \mathcal{D}$, *then* $\rho_t(D) \leq \rho_t(D')$.

R5. ***Cash-additive.*** $\rho_t(D + m\mathbb{1}_{\{s\}}) = \rho_t(D) - m$ *for any* $t \in \mathcal{T}$, $D \in \mathcal{D}$, *and* $\mathcal{F}_t$*-measurable random variable* m *and* $s \geq t$.

R6. ***Time consistent.*** $\rho_t(D) = \rho_t(-\rho_{t+1}(D)\mathbb{1}_{\{t+1\}}) - D_t$ *for any* $t = 0, 1, \ldots, T-1$ *and* $D \in \mathcal{D}$.

If ρ furthermore is

R7. **Positive-homogeneous.** $\rho_t(\lambda D) = \lambda \rho_t(D)$, $\lambda \in L^\infty_+(\mathcal{F}_t)$,

then $\rho : \mathcal{T} \times \mathcal{D} \times \Omega \to \mathbb{R}$ is called a *dynamic coherent risk measure.*

Properties R1-R7 have a clear financial interpellation: adaptiveness means that the measurements are consistent with the flow of information; the locality property essentially means that the values of the risk measure restricted to a set $A \in \mathcal{F}$ remain invariant with respect to the values of the arguments outside of the same set $A \in \mathcal{F}$, and in particular, the events that will not happen in the future do not change the value of the measure today; convexity implies that diversification reduces the risk; monotonicity implies that a cash flow with higher payoffs bears less risk.

Cash-additivity means that adding \$$m$ to a portfolio at any time in the future reduces the overall risk by the same amount m. From the regulatory perspective, the value of a risk measure is typically interpreted as the minimal capital requirement for a bank. There exists various forms of time consistency of risk and performance measures and we refer the reader to recent paper [BCP15] where the authors give a systematic approach to time consistency of LM-measures (local and monotone functions). The time consistency R6 for DCRM considered here is known in the existing literature as strong time consistency. It can be shown that property R6, combined with R4, is equivalent to the following property: if $D_t = D'_t$ and $\rho_{t+1}(D) = \rho_{t+1}(D')$, then $\rho_t(D) = \rho_t(D')$, i.e. if two cash flows bear the same risk tomorrow, and they pay the same dividend today, then today these two cash flows are assessed at the same risk level. Saying differently, the risk is measured consistently in time. Finally, positive-homogeneity means the risk of a rescaled cash flow is rescaled by the same factor.

Similar to, [CE10, ESC15, Sta09] that address discrete time case, and [RG06, RGS13] that consider the continuous time set-up, we will show that the solution of BSΔEs (2.1) with convex drivers, more precisely the corresponding g-Expectation, generates a DCRM. In the sequel, for any regular and convex driver g, we will denote by ρ^g the function defined as follows

$$\rho^g_t(D) = \mathcal{E}_g\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big], \quad t \in \mathcal{T},\ D \in \mathcal{D}. \tag{2.18}$$

Theorem 2.2.1. *Assume that g is a convex and regular driver. Then, the function ρ^g defined as in* (2.18) *is a DCRM.*

Proof. Invoking Proposition 2.1.1, it is straightforward to show that ρ^g_t satisfies Properties R1–R5.

We will show that ρ^g is time consistent. By Proposition 2.1.1 (iii) and (v), we immediately have that

$$\begin{aligned}\rho_t^g(-\rho_{t+1}^g(D)\mathbb{1}_{\{t+1\}}) - D_t =& \mathcal{E}_g[\rho_{t+1}^g(D)|\mathcal{F}_t] - D_t = \mathcal{E}_g\Big[\mathcal{E}_g\Big[-\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_{t+1}\Big]\Big|\mathcal{F}_t\Big] - D_t \\ =& \mathcal{E}_g\Big[\mathcal{E}_g\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_{t+1}\Big]\Big|\mathcal{F}_t\Big] = \mathcal{E}_g\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big] \\ =& \rho_t^g(D),\end{aligned}$$

for any $t \in \mathcal{T}$, $D \in \mathcal{D}$. This concludes the proof. □

As an immediate consequence of Theorem 2.2.1, if additionally the driver g is positive homogeneous in z, then ρ^g is a time consistent coherent risk measure.

Corollary 2.2.1. *Assume that $g(t, z)$ is a convex and regular driver, such that $g(t, \cdot)$ is positive homogeneous. Then, ρ_t^g is a dynamic coherent risk measure.*

Proof. In view of Theorem 2.2.1, we have that $\rho_t^g(D)$ is a DCRM. Moreover, for any fixed $t \in \mathcal{T}$, due to Proposition 2.1.1 (vii), we have that

$$\rho_t^g(\lambda D) = \mathcal{E}_g\Big[-\lambda\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big] = \lambda\mathcal{E}_g\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big] = \lambda\rho_t^g(D),$$

for any $\lambda \in L_+^\infty(\mathcal{F}_t)$, and $D \in \mathcal{D}$. Hence, ρ_t^g is a dynamic coherent risk measure. □

Next, we give some examples of DCRMs generated by various drivers via g-expectation. Similar to Example 2.1.1 and 2.1.2, in the following two examples, we will assume that W is a martingale process such that ΔW_t is uniformly bounded.

Example 2.2.1 (Coherent Case)**.** *We consider the driver $g(t, z) = c|z|$, for some fixed $c \in [0, 1)$. It is easy to see that $g(t, z)$ is convex and positive homogeneous with respect to z. Then, by Corollary 2.2.1, $\rho_t^g(D) = \mathcal{E}_g[-\sum_{s=t}^T D_s|\mathcal{F}_t]$ is a dynamic coherent risk measure.*

Example 2.2.2 (Convex Case). *Let us put*

$$g(t,z) = \frac{K}{(K+1)\|\Delta W_t\|_\infty} \ln\Big(\frac{1}{3} + \frac{1}{3}e^{-z} + \frac{1}{3}e^{z}\Big),$$

where $K \in \mathbb{R}^+$ is fixed. Note that $\frac{\partial g}{\partial z}$ is an increasing function with respect to z, and hence $g(t,\cdot)$ is a convex driver. By Theorem 2.2.1, we have that $\rho_t^g(D) = \mathcal{E}_g[-\sum_{s=t}^T D_s|\mathcal{F}_t]$ is a time consistent DCRM.

By Theorem 2.2.1 any given convex regular driver g generates DCRM. Next result shows that the converse implication also holds true if W is the symmetric random walk and the filtration is generated by W.

Proposition 2.2.1. *Assume that W is a symmetric random walk, $\mathcal{F} = \mathcal{F}^W$, and $\rho_t : L^\infty(\mathcal{F}_T) \to L^\infty(\mathcal{F}_t)$ is a DCRM. Also, let $g(t,z) = \frac{\rho_{t-1}(z\Delta W_t)}{\Delta\langle W\rangle_t}$. Then, $\rho_t(X) = \rho_t^g(X)$, for any $X \in L^\infty(\mathcal{F}_T)$.*

Proof. Fix an $X \in L^\infty(\mathcal{F}_T)$. Since $\rho_t(X) \in L^\infty(\mathcal{F}_t) \subset L^2(\mathcal{F}_t)$, for any $t \in \mathcal{T}$, then, since W has the martingale representation property, there exists $Z_t \in \mathcal{F}_{t-1}$ such that $\rho_t(X) = \mathbb{E}[\rho_t(X)|\mathcal{F}_{t-1}] + Z_t\Delta W_t$. According to time-consistency property of ρ, we have that

$$\begin{aligned}
\rho_t(X) - \rho_{t-1}(X) =& \rho_t(X) - \rho_{t-1}(-\rho_t(X)) \\
=& \mathbb{E}[\rho_t(X)|\mathcal{F}_{t-1}] + Z_t\Delta W_t - \mathbb{E}[\rho_t(X)|\mathcal{F}_{t-1}] - \rho_{t-1}(Z_t\Delta W_t) \\
=& -\rho_{t-1}(Z_t\Delta W_t) + Z_t\Delta W_t.
\end{aligned}$$

Therefore,

$$\rho_{t-1}(X) = \rho_t(X) + \frac{\rho_{t-1}(Z_t\Delta W_t)}{\Delta\langle W\rangle_t}\Delta\langle W\rangle_t - Z_t\Delta W_t$$

and hence, $\rho_t(X) = \rho_t^g(X)$. □

The next example is an application of Proposition 2.2.1.

Example 2.2.3 (Dynamic Entropic Risk Measure)**.** *The dynamic entropic risk measure takes the following form*

$$\rho_t^\gamma(X) = \gamma \ln \Big(\mathbb{E}\Big[\exp(-\frac{X}{\gamma}) \Big| \mathcal{F}_t \Big] \Big) \tag{2.19}$$

where $\gamma > 0$*. So the driver corresponding to entropic risk measure will be*

$$g(t,z) = \frac{\gamma}{\Delta \langle W \rangle_t} \ln \Big(\mathbb{E}\Big[\exp(-\frac{z\Delta W_t}{\gamma}) \Big| \mathcal{F}_{t-1} \Big] \Big) = \frac{\gamma}{\Delta \langle W \rangle_t} \ln \Big(\frac{1}{2} e^{-\frac{z}{\gamma}} + \frac{1}{2} e^{\frac{z}{\gamma}} \Big).$$

Similarly as in Example 2.2.2, we have that $g(t,z)$ *is a convex regular driver.*

Remark 2.2.2. *It is worth observing that, if we take the time step to be equal to* Δt*, and the martingale* W *to be a scaled symmetric random walk* $\mathbb{P}(W_{t+\Delta t} = \pm\sqrt{\Delta t} \,|\, \mathcal{F}_t)$*, with* $W_0 = 0$*, and* $\mathcal{F} = \mathcal{F}^W$*, the driver corresponding to the entropic risk measure* (2.19) *takes the form*

$$g(t,z) = \frac{\gamma}{\Delta t} \ln \Big(\frac{1}{2} e^{-\frac{z}{\gamma}\sqrt{\Delta t}} + \frac{1}{2} e^{\frac{z}{\gamma}\sqrt{\Delta t}} \Big).$$

By direct computations, we have that $\lim_{\Delta t \to 0} g(t,z) = z^2/(2\gamma)$*. Formally, this means that in continuous time framework the BSDE that corresponds to the dynamic entropic risk measure has a quadratic driver, which is consistent with the results in the existing literature [BEK07, Proposition 6.4].*

2.2.2 Dynamic Acceptability Indices via g-Expectation. In [CM09] the authors introduced the notion of coherent acceptability index, meant to measure the performance of a given terminal cash flow, as a monotone, scale-invariant, and quasi-concave function defined on the set of all bounded random variables. The extension of these measures to a dynamic set-up was studied in [BCZ14], where the appropriate notion of time consistency for dynamic coherent acceptability indices was introduced. In [BCZ14], the author follow a discrete time market set-up on a finite probability space. Recently, Biagini and Bion-Nadal [BBN14] studied this type of dynamic performance measure of terminal cash-flows on a general probability space, and discrete time market set-up. For robust representations of general dynamic quasi-concave performance measures we refer the reader to [BCDK15]. The aim of this section is to study dynamic performance measures that are sub-scale invariance. Following [RGS13], who argued that the scale invariance fails to capture all risks in an illiquid market, and, accordingly, postulated the sub-scale invariance instead, we replace the scale invariance condition used in [BCZ14], with a weaker assumption of sub-scale invariance.

Definition 2.2.2 (Dynamic Acceptability Index)**.** *A dynamic acceptability index is a function $\alpha : \mathcal{T} \times \mathcal{D} \times \Omega \to [0, \infty]$ that satisfies the following properties:*

I1. ***Adapted.*** *$\alpha_t(D)$ is $\mathcal{F}_t$-measurable, for any $t \in \mathcal{T}, D \in \mathcal{D}$.*

I2. ***Local.*** *$\mathbb{1}_A \alpha_t(D) = \mathbb{1}_A \alpha_t(\mathbb{1}_A D)$, for any $t \in \mathcal{T}, A \in \mathcal{F}_t, D \in \mathcal{D}$.*

I3. ***Quasi-concave.*** *If $\alpha_t(D) \geq m$ and $\alpha_t(D') \geq m$ for some positive $\mathcal{F}_t$-measurable random variable m, and $D, D' \in \mathcal{D}$, then $\alpha_t(\lambda D + (1-\lambda)D') \geq m$, for any $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$.*

I4. ***Monotone.*** *If $D \succeq_t D'$ for some $t \in \mathcal{T}$, and $D, D' \in \mathcal{D}$, then $\alpha_t(D) \geq \alpha_t(D')$.*

I5. ***Sub-scale Invariant.*** *$\alpha_t(\lambda D) \geq \alpha_t(D)$ for any $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$, $D \in \mathcal{D}$, or, equivalently, $\alpha_t(\lambda D) \leq \alpha_t(D)$ for any $\lambda \in L^\infty(\mathcal{F}_t)$, $\lambda \geq 1$, $D \in \mathcal{D}$.*

I6. ***Time Consistent.*** *For any $t \in \mathcal{T}$, $D, D' \in \mathcal{D}$, the following implication holds true*

$$\alpha_{t+1}(D) \geq m \geq \alpha_{t+1}(D') \quad \Rightarrow \quad \alpha_t(D) \geq m \geq \alpha_t(D'),$$

whenever $D_t \geq 0 \geq D'_t$, and m being an non-negative $\mathcal{F}_t$-measurable random variable.

Remark 2.2.3. *(i) We note that property I5 is weaker than the scale invariance property*

*I5'. **Scale Invariant.** $\alpha_t(\lambda D) = \alpha_t(D)$ for any $\lambda \in L^\infty(\mathcal{F}_t)$, $\lambda > 0$, $D \in \mathcal{D}$,*

(ii) If property I5 from the definition of DAI is replaced with I5', then α is called dynamic coherent acceptability index.

Analogously to [BCZ14, CM09], we will show that a DAI can be generated by a family of DCRMs, and hence by a family of BSΔEs. For this, we will consider families of drivers indexed by positive real numbers that satisfy the following assumptions:

Assumption G:

G1. $g_{x_2} \geq g_{x_1}$ for $x_2 \geq x_1 > 0$;

G2. g_x is a convex regular driver for any $x > 0$;

G3. $g_x = g_{x-}$ for any $(t, \omega, z) \in \mathbb{R}^+ \times \Omega \times \mathbb{R}$.

With slight abuse of notations, we will denote by g the family of drivers $(g_x)_{x>0}$.

In what follows, for any family of drivers g that satisfy assumption G, we denote by α^g the following function

$$\alpha_t^g(D)(\omega) := \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big](\omega) \leq 0\Big\}, \quad \omega \in \Omega, t \in \mathcal{T},\ D \in \mathcal{D}. \tag{2.20}$$

Then we have the following theorem.

Theorem 2.2.2. *Assume that the family of drivers $g = (g_x)_{x>0}$ satisfies Assumption G. Then, α^g is a dynamic acceptability index.*

The proof will follow from a series of lemmas proved below.

Lemma 2.2.1. *Assume that the family of drivers $g = (g_x)_{x>0}$ satisfies Assumption G. Then, α^g satisfies properties I1-I5.*

Proof. We will show that α^g satisfies properties I1-I5.

I1. Let us consider the set $A_\gamma = \{\omega \in \Omega : \alpha_t^g(D)(\omega) \geq \gamma\}$, where $\gamma \in \mathbb{R}$, $t \in \mathcal{T}$ and $D \in \mathcal{D}$ are fixed. We want to show that $A_\gamma \in \mathcal{F}_t$. If $\gamma \leq 0$, then it is clear that $A_\gamma = \Omega \in \mathcal{F}_t$. For $\gamma > 0$, we will prove that

$$\alpha_t^g(D)(\omega) \geq \gamma, \quad \omega \in \Omega,$$

is equivalent to

$$\mathcal{E}_{g_\gamma}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq 0, \quad \omega \in \Omega. \tag{2.21}$$

According to definition (2.20), any $\omega \in \Omega$ such that $\alpha_t^g(D)(\omega) \geq \gamma$ satisfies that

$$\sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq 0\Big\} \geq \gamma,$$

which, by assumption G1, implies that

$$\lim_{x\uparrow\gamma} \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq 0.$$

Therefore, in order to show (2.21) holds for such ω, we only need to verify that

$$\lim_{x\uparrow\gamma} \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) = \mathcal{E}_{g_\gamma}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \quad \omega \in A_\gamma. \tag{2.22}$$

Let (Y^x, Z^x, M^x), $0 < x \leq \gamma$, be the solutions of BSΔEs with drivers g_x and the terminal condition $Y_T^x = -\sum_{s=t}^T D_s$. Then, (2.22) is implied by

$$\lim_{x\uparrow\gamma} Y_t^x = Y_t^\gamma, \tag{2.23}$$

which holds clearly for $t = T$.

Suppose that (2.23) is true for some $t < u \leq T$. According to (2.9), we have the representation:

$$Z_u^x = \frac{\mathbb{E}[Y_u^x \Delta W_u | \mathcal{F}_{u-1}]}{\Delta\langle W\rangle_u}.$$

Notice that $\{Y_u^x\}_{0<x<\gamma}$ is increasing with respect to x because of Theorem 2.1.2. Hence, by dominated convergence theorem, we get that

$$\lim_{x\uparrow\gamma} Z_u^x = \lim_{x\uparrow\gamma} \frac{\mathbb{E}[Y_u^x \Delta W_u|\mathcal{F}_{u-1}]}{\Delta\langle W\rangle_u} = \frac{\mathbb{E}[Y_u^\gamma \Delta W_u|\mathcal{F}_{u-1}]}{\Delta\langle W\rangle_u} = Z_u^\gamma.$$

Next, by (2.10), Y_{u-1}^x can be represented by

$$Y_{u-1}^x = \mathbb{E}[Y_u^x|\mathcal{F}_{u-1}] + g_x(u, Z_u^x)\Delta\langle W\rangle_u,$$

where the following equality

$$\lim_{x\uparrow\gamma} \mathbb{E}[Y_u^x|\mathcal{F}_{u-1}] = \mathbb{E}[Y_u^\gamma|\mathcal{F}_{u-1}]$$

holds true due to dominated convergence theorem. Moreover, we have that

$$\begin{aligned}
&|g_x(u,\omega,Z_u^x(\omega)) - g_\gamma(u,\omega,Z_u^\gamma(\omega))| \\
\leq &|g_x(u,\omega,Z_u^x(\omega)) - g_x(u,\omega,Z_u^\gamma(\omega))| + |g_x(u,\omega,Z_u^\gamma(\omega)) - g_\gamma(u,\omega,Z_u^\gamma(\omega))| \\
\leq &c_t^x(\omega)|Z_u^x(\omega) - Z_u^\gamma(\omega)| + |g_x(u,\omega,Z_u^\gamma(\omega)) - g_\gamma(u,\omega,Z_u^\gamma(\omega))|,
\end{aligned}$$

for almost all $\omega \in \Omega$. Recall that $c_t^x(\omega)$ is defined as the smallest Lipschitz constant of g_x. Since that the family of drivers satisfy assumption G1 and G2, then we have that $c_t^x \leq c_t^\gamma$. The following equality follows immediately:

$$\lim_{x\uparrow\gamma} g_x(u, Z_u^x) = g_\gamma(u, Z_u^\gamma)$$

Thus, we have proved

$$\lim_{x\uparrow\gamma} Y_{u-1}^x = Y_{u-1}^\gamma,$$

and (2.23) is true by induction.

On the other hand, for any $\omega \in \Omega$ such that $\mathcal{E}_{g_\gamma}\big[-\sum_{s=t}^T D_s\big|\mathcal{F}_t\big](\omega) \leq 0$. We get that

$$\alpha_t^g(D)(\omega) = \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\sum_{s=t}^T D_s\Big|\mathcal{F}_t\Big](\omega) \leq 0\Big\} \geq \gamma,$$

which implies that $\omega \in A_\gamma$. In view of the above, we conclude that

$$A_\gamma = \Big\{\omega \in \Omega : \alpha_t^g(D)(\omega) \geq \gamma\Big\} = \Big\{\omega \in \Omega : \mathcal{E}_{g_\gamma}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq 0\Big\}.$$

Since that $\mathcal{E}_{g_\gamma}\big[-\sum_{s=t}^{T} D_s \big| \mathcal{F}_t\big]$ is $\mathcal{F}_t$-measurable, then $A_\gamma \in \mathcal{F}_t$. This completes the proof of showing that α^g is adapted.

I2. Let us fix $t \in \mathcal{T}$, $D \in \mathcal{D}$, and $A \in \mathcal{F}_t$. Then, for almost all $\omega \in \Omega$, we have that

$$\begin{aligned}\mathbb{1}_A(\omega)\alpha_t^g(D)(\omega) =& \mathbb{1}_A(\omega) \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq 0\Big\} \\ =& \mathbb{1}_A(\omega) \sup\Big\{x \in \mathbb{R}, x > 0 : \mathbb{1}_A(\omega)\mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq 0\Big\}.\end{aligned}$$

By (v) in Proposition 2.1.1, we deduce that

$$\mathbb{1}_A \mathcal{E}_{g_x}[-\sum_{s=t}^{T} D_s | \mathcal{F}_t] = \mathcal{E}_{g_x}[-\mathbb{1}_A \sum_{s=t}^{T} D_s | \mathcal{F}_t],$$

and locality of α^g follows immediately.

I3. Let us fix $t \in \mathcal{T}$, D, $D' \in \mathcal{D}$. Also, let $\gamma > 0$ be some $\mathcal{F}_t$-measurable random variable such that $\alpha_t^g(D)(\omega) \geq \gamma(\omega)$ and $\alpha_t^g(D')(\omega) \geq \gamma(\omega)$ hold for almost all ω. Fix one such ω and denote $\gamma^* = \gamma(\omega)$. Similar to the proof of adaptiveness, we have that $\alpha_t^g(D)(\omega) \geq \gamma^*$ and $\alpha_t^g(D')(\omega) \geq \gamma^*$ will imply that

$$\begin{aligned}\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) = \mathcal{E}_{g_{\gamma^*}}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq 0, \\ \mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D'_s \Big| \mathcal{F}_t\Big](\omega) = \mathcal{E}_{g_{\gamma^*}}\Big[-\sum_{s=t}^{T} D'_s \Big| \mathcal{F}_t\Big](\omega) \leq 0,\end{aligned}$$

respectively. Then according to (vi) in Proposition 2.1.1, we get that

$$\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} \Big(\lambda D_s + (1-\lambda)D'_s\Big) \Big| \mathcal{F}_t\Big](\omega) \leq 0,$$

holds for almost all $\omega \in \Omega$, and it implies that

$$\begin{aligned}\alpha_t^g(\lambda D + (1-\lambda)D')(\omega) = \sup\Big\{& x \in \mathbb{R}, x > 0 : \\ & \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} \Big(\lambda D_s + (1-\lambda)D'_s\Big) \Big| \mathcal{F}_t\Big](\omega) \leq 0\Big\} \geq \gamma(\omega).\end{aligned}$$

Therefore, the Quasi-concavity of α^g holds.

I4. Fix $t \in \mathcal{T}$. Let $D,\ D' \in \mathcal{D}$, and suppose that $D \succeq_t D'$. For any fixed $\mathcal{F}_t$-measurable random variable $\gamma > 0$ such that $\alpha_t^g(D') \geq \gamma$, we have that

$$\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D'_s \Big| \mathcal{F}_t\Big](\omega) \leq 0$$

holds for almost every $\omega \in \Omega$. Due to the assumption $D \succeq_t D'$, which implies that $\sum_{s=t}^{T} D_s \geq \sum_{s=t}^{T} D'_s$, we get by (ii) in Proposition 2.1.1 that

$$\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \leq \mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D'_s \Big| \mathcal{F}_t\Big](\omega) \leq 0,$$

and the statement $\alpha_t^g(D) \geq \gamma$ follows. Hence, α^g is monotone.

I5. Fix $t \in \mathcal{T}$, $D \in \mathcal{D}$. For any fixed $\mathcal{F}_t$-measurable random variable $\gamma > 0$ such that $\alpha_t^g(D) \geq \gamma$, we have that

$$\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D'_s \Big| \mathcal{F}_t\Big](\omega) \leq 0$$

holds for almost every $\omega \in \Omega$. By convexity of g-expectation, it is true that

$$\begin{aligned}\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\lambda\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big] &\leq \lambda\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big] + (1-\lambda)\mathcal{E}_{g_{\gamma(\omega)}}\Big[0\Big|\mathcal{F}_t\Big] \\ &= \lambda\mathcal{E}_{g_{\gamma(\omega)}}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big] \leq 0,\end{aligned}$$

for any $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$, and almost all $\omega \in \Omega$. Therefore, we conclude that $\alpha_t^g(\lambda D) \geq \gamma$, which implies that $\alpha_t^g(\lambda D) \geq \alpha_t^g(D)$.

□

Remark 2.2.4. *In Assumption G, G1 and G2 are natural conditions for constructing DAI via g-expectation. The reason to assume G3 comes from duality between DCRM and DAI. If a DAI is given, then DCRM can be defined as*

$$\rho_t^\gamma(D) = \operatorname{ess\,inf}\{c \in \mathcal{F}_t : \alpha_t^g(D + c\mathbb{1}_{\{t\}}) \geq \gamma\},$$

which is equivalent to

$$\rho_t^{\gamma}(D) = \operatorname{ess\,inf}\Big\{c \in \mathcal{F}_t : \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s - c\mathbb{1}_{\{t\}}\Big|\mathcal{F}_t\Big] \leq 0\Big\} \geq \gamma\Big\}.$$

So we have that,

$$\rho_t^{\gamma}(D) = \operatorname{ess\,inf}\Big\{c \in \mathcal{F}_t : \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big] \leq c\Big\} \geq \gamma\Big\}.$$

Random variable $\rho_t^{\gamma}(D)$ can be represented as

$$\rho_t^{\gamma}(D) = \operatorname{ess\,inf}\Big\{c \in \mathcal{F}_t : \forall\varepsilon > 0, \mathcal{E}_{g_{\gamma-\varepsilon}}\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big] \leq c\Big\},$$

Hence, due to assumption G1, we get that $\rho_t^{\gamma}(D) = \sup_{\varepsilon>0}\{\mathcal{E}_{g_{\gamma-\varepsilon}}[-\sum_{s=t}^{T} D_s|\mathcal{F}_t]\} = \mathcal{E}_{g_{\gamma-}}[-\sum_{s=t}^{T} D_s|\mathcal{F}_t]$. In order for duality to hold, it is required to have that $\rho_t^{g_\gamma}(D) = \rho_t^{\gamma}(D)$ which is equivalent to

$$\mathcal{E}_{g_{\gamma-}}\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big] = \mathcal{E}_{g_\gamma}\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big].$$

Finally, according to comparison theorem, we have that $g_\gamma = g_{\gamma-}$.

Lemma 2.2.2. *Assume that the family of drivers $g = (g_x)_{x>0}$ satisfies Assumption G. Also suppose that g_x is positive homogeneous for any $x \in \mathbb{R}^+$. Then, α^g is scale invariant.*

Proof. We need to show that α^g is scale invariant.

Let $t \in \mathcal{T}$, $D \in \mathcal{D}$, and $\lambda \in L^\infty(\mathcal{F}_t)$, $\lambda > 0$. By definition, we have that

$$\alpha_t^g(\lambda D)(\omega) = \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\lambda\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big](\omega) \leq 0\Big\},$$

holds for almost all $\omega \in \Omega$. In view of the fact that $g_x(t, \cdot)$ is positive homogeneous for any $x > 0$, Proposition 2.1.1 (vii) implies that

$$\alpha_t^g(\lambda D)(\omega) = \sup\Big\{x \in \mathbb{R}, x > 0 : \lambda\mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s\Big|\mathcal{F}_t\Big](\omega) \leq 0\Big\},$$

for almost all $\omega \in \Omega$. Since $\lambda > 0$, then we get that $\lambda\mathcal{E}_{g_x}[-\sum_{s=t}^T D_s|\mathcal{F}_t] \leq 0$ is equivalent to $\mathcal{E}_{g_x}[-\sum_{s=t}^T D_s|\mathcal{F}_t] \leq 0$. Hence, $\alpha_t^g(\lambda D)(\omega) = \sup\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}[-\sum_{s=t}^T D_s|\mathcal{F}_t](\omega) \leq 0\} = \alpha_t^g(D)(\omega)$, for almost all $\omega \in \Omega$. This concludes that α^g is a dynamic coherent acceptability index. □

Remark 2.2.5. *In view of I4 and I5 in Definition 2.2.2, for any $D \succeq_t 0$ and $\lambda \in L^\infty(\mathcal{F}_t)$ a DAI satisfies that $\alpha_t(\lambda D) = \alpha_t(D)$. As matter of fact, for a DAI generated by a family of DCRMs: $\alpha_t(D)(\omega) = \sup\{x \in \mathbb{R}, x > 0 : \rho^x(D)(\omega) \leq 0\}$, $\omega \in \Omega$, $t \in \mathcal{T}$, $D \in \mathcal{D}$, since $\rho_t^x(D) \leq 0$ whenever $D \succeq_t 0$, so that $\alpha_t(\lambda D) = \alpha_t(D) = \infty$ if $D \succeq_t 0$. The financial interpretation is clear: a positive cash flow does not bear any risk, including liquidity risk, therefore such cash flow is accepted at any level and is scale invariant.*

By using the following lemma, we will prove that α^g is time consistent.

Lemma 2.2.3. *Assume that the family of drivers $g = (g_x)_{x>0}$ satisfies Assumption G. Also suppose that $D_t \geq 0 \geq D'_t$ for some $t \in \mathcal{T}$, $D, D' \in \mathcal{D}$, and there exists $c \in \mathbb{R}^+, A \in \mathcal{F}_t$ such that $\mathbb{1}_A\alpha_{t+1}^g(D) \geq \mathbb{1}_A c \geq \mathbb{1}_A\alpha_{t+1}^g(D')$. Then $\mathbb{1}_A\alpha_t^g(D) \geq \mathbb{1}_A c \geq \mathbb{1}_A\alpha_t^g(D')$.*

Proof. Suppose that $\mathbb{1}_A\alpha_t^g(D) \geq \mathbb{1}_A c \geq \mathbb{1}_A\alpha_t^g(D')$ for some $t \in \mathcal{T}$, $c \in \mathbb{R}^+$, $A \in \mathcal{F}_t$, and $D, D' \in \mathcal{D}$ such that $D_t \geq 0 \geq D'_t$. Then, for any $c' < c$, we have that $\mathcal{E}_{g_{c'}}[-\mathbb{1}_A\sum_{s=t+1}^T D_s|\mathcal{F}_{t+1}] \leq 0$. Since $D_t \geq 0$, then according to Proposition 2.1.1 (ii) and (iii), we get that

$$\mathcal{E}_{g_{c'}}\Big[-\mathbb{1}_A\sum_{s=t}^T D_s\Big|\mathcal{F}_t\Big] = -\mathbb{1}_A D_t + \mathcal{E}_{g_{c'}}\Big[\mathcal{E}_{g_{c'}}\Big[-\mathbb{1}_A\sum_{s=t+1}^T D_s\Big|\mathcal{F}_{t+1}\Big]\mathcal{F}_t\Big] \leq 0.$$

Due to the fact that c' is arbitrary, the following is true for almost all $\omega \in \Omega$:

$$\mathbb{1}_A(\omega)\alpha_t^g(D)(\omega) = \mathbb{1}_A(\omega) \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \le 0\Big\}$$
$$= \mathbb{1}_A(\omega) \sup\Big\{x \in \mathbb{R}, x > 0 : \mathcal{E}_{g_x}\Big[-\mathbb{1}_A \sum_{s=t}^{T} D_s \Big| \mathcal{F}_t\Big](\omega) \le 0\Big\}$$
$$\ge \mathbb{1}_A(\omega)c.$$

On the other hand, by knowing that $\mathbb{1}_A\alpha_{t+1}^g(D') \le \mathbb{1}_A c$, we will prove $\mathbb{1}_A\alpha_t^g(D') \le \mathbb{1}_A c$ by contradiction. Assume that there exists some $A' \subset A$, $\mathbb{P}(A') > 0$ such that $\alpha_t^g(D') > c$ on A'. Then there exists a $c' > c$ and $A'' \subset A$, $\mathbb{P}(A'') > 0$ such that $\alpha_t^g(D) > c'$ on A". Hence, we have that

$$\mathcal{E}_{g_{c'}}\Big[-\mathbb{1}_{A''}\sum_{s=t}^{T} D'_s \Big| \mathcal{F}_t\Big] \le 0.$$

However, since $\alpha_{t+1}^g(D') \le c$ on A, then $\alpha_{t+1}^g(D') < c'$ on A''. In view of the fact that $D'_t \le 0$, we get for $\omega \in A''$ that

$$\mathcal{E}_{g_{c'}}\Big[-\mathbb{1}_{A''}\sum_{s=t}^{T} D'_s \Big| \mathcal{F}_t\Big](\omega) = -\mathbb{1}_{A''}(\omega)D'_t(\omega) + \mathcal{E}_{g_{c'}}\Big[\mathcal{E}_{g_{c'}}\Big[-1_{A''}\sum_{s=t+1}^{T} D'_s \Big| \mathcal{F}_{t+1}\Big| \mathcal{F}_t\Big](\omega) > 0.$$

Hence, there is a contradiction and such result implies that $\mathbb{1}_A\alpha_t^g(D') \le \mathbb{1}_A c$. $\square$

Now we are ready to finish proving Theorem 2.2.2, which is to verify that α^g is time consistent.

Lemma 2.2.4. *Assume that the family of drivers $g = (g_x)_{x>0}$ satisfies Assumption G. Then, α^g satisfies I6.*

Proof. For any $t \in \mathcal{T}$, $D \in \mathcal{D}$, if there exists a positive $\mathcal{F}_t$-measurable random variable m such that $\alpha_{t+1}^g(D) \ge m$, then there exists a sequence of simple random variables $\phi_n = \sum_{i=1}^{k} \mathbb{1}_{A_i^n} a_i^n$ where $A_i^n \in \mathcal{F}_t$, $a_i^n \in \mathbb{R}^+$, $n \in \{1, 2, \ldots\}$, such that $\phi_n \le \phi_{n+1}$ and

$\lim_{n\to\infty}\phi_n = m$. Hence, we have that $\alpha^g_{t+1}(D) \geq \phi_n$ for any n. Since that $D_t \geq 0$, then according to locality of α^g and Lemma 2.2.3, The following is true

$$\alpha^g_t(D) \geq \phi_n, \qquad n \in \{1, 2, \ldots\}.$$

Therefore, we conclude that $\alpha^g_t(D) \geq m$.

Assume that there exists a positive $\mathcal{F}_t$-measurable random variable m such that $\alpha^g_{t+1}(D') \leq m$. Fix $N \in \mathbb{R}^+$, and for any $\omega \in \Omega$, define

$$m_N(\omega) = \begin{cases} m(\omega) & \text{if } m(\omega) < N, \\ N & \text{if } m(\omega) \geq N. \end{cases}$$

It is clear that $(\alpha^g_{t+1})_N(D') \leq m_N$ and m_N is bounded. Hence, there exists a sequence of simple random variables ϕ_n such that $\phi_n \geq m_N$, and $\lim_{n\to\infty}\phi_n = m_N$. Therefore $(\alpha^g_{t+1})_N(D') \leq \phi_n$ for all n. Since $D'_t \leq 0$, then by locality of α^g and Lemma 2.2.3, we have that $(\alpha^g_t)_N(D') \leq \phi_n$ for any $n \in \{1, 2, \ldots, \}$. Moreover, it implies that $(\alpha^g_t)_N(D') \leq m_N$. Let N go to infinity, we conclude that $\alpha^g_t(D') \leq m$. □

Before proceed to the next section, we give several examples of DAIs generated by g-expectations. While discussing these examples, we take W as a symmetric random walk.

Example 2.2.4. *Let $g = (g_x)_{x>0}$ be in the form of $g_x(t, z) = \frac{x}{(x+1)}\ln(\frac{1}{3} + \frac{1}{3}e^{-z} + \frac{1}{3}e^{z})$. Similar to Example 2.2.2, we have that each $g_x(t, z)$ a convex regular driver. For fixed t, z, $g_x(t, z)$ is increasing with respect to x. Therefore, g satisfies Assumption G. Thus, $\alpha^g_t(D)$ is a DAI.*

Example 2.2.5 (Coherent DAI)**.** *Let $g = (g_x)_{x>0}$ be in the form of $g_x(t, z) = \frac{x}{x+1}|z|$. Then g satisfies Assumption G. Hence, $g_x(t, \cdot)$ is positive homogeneous. According to Lemma 2.2.2, $\alpha^g_t(D)$ is a dynamic coherent acceptability index.*

Example 2.2.6 (Entropic DAI). *Let $g = (g_x)_{x>0}$ be such that*

$$g_x(t,z) = \frac{1}{x\Delta\langle W\rangle_t} \ln\Big(\frac{1}{2}e^{-xz} + \frac{1}{2}e^{xz}\Big).$$

Then, due to Example 2.2.3, such family of drivers satisfies Assumption G, and α^g is a DAI, to which we refer as entropic DAI.

2.3 Dynamic Conic Finance as Market Model

Cherny and Madan [CM10] proposed the conic finance framework for pricing non-dividend paying securities using static acceptability indices. In [BCIR13], the authors generalized such technique to a dynamic framework that allows cash flows to pay dividends and be subjected to transaction costs by using dynamic coherent acceptability indices that were obtained in [BCZ14]. Nevertheless, in [AS08] and [RGS13], the authors presented a systematic criticism to the positive homogeneity and sublinearity assumptions frequently adopted in the framework of coherent risk measures, and Bion-Nadal [BN09] introduced a dynamic approach to bid and ask prices taking into account both transaction costs and liquidity risk, based on Time Consistent Pricing Procedures. In this section, we will build a market model in which the securities are priced by an acceptability method. This set-up accounts for transaction costs and liquidity risk, by using the dynamic quasi-concave acceptability indices via g-expectations we developed earlier.

2.3.1 Market Set-up. In this section we retain the same probabilistic framework and the same notations as in the previous sections. In particular, we consider an underlying probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$, and we assume that all processes considered below are $\mathbb{F}$-adapted and appropriately integrable. On this probability space we consider a market consisting of a banking account (or money market account) and K securities. Throughout, we pick the T as the largest time horizon. We also adopt the convention that all security prices are already discounted with the banking account. Our market is further characterized as follows:

(M1) The process $D^{\text{ask},i} := (D_t^{\text{ask},i})_{t\in\mathcal{T}} \in \mathcal{D}$, with $D_0^{\text{ask},i} = 0$, represents the dividend process associated with *holding a long position of 1 share* of the ith security, $i = 1,\dots,K$. Correspondingly, $D^{\text{bid},i} := (D_t^{\text{bid},i})_{t\in\mathcal{T}} \in \mathcal{D}$, with $D_0^{\text{bid},i} = 0$, is the dividend process associated with *holding a short position of 1 share* of security $i = 1,\dots,K$. We stress that processes $D^{\text{ask/bid},i}$ represent bullet dividend cash flows, rather than cumulative dividend.

(M2) $P_t^{\text{ask}}(\varphi, D^{\text{ask/bid},i})$, respectively $P_t^{\text{bid}}(\varphi, D^{\text{ask/bid},i})$, denote the ex-dividend prices of purchasing, respectively selling, $\varphi \in L_+^\infty(\mathcal{F}_t)$ shares of cash flows $D^{\text{ask},i}$ or $D^{\text{bid},i}$ that are associated with security $i \in \{1,\dots,K\}$ at time $t \in \mathcal{T}$. We also assume that the pricing operators $P_t^{\text{ask}}, P_t^{\text{bid}} : L_+^\infty(\mathcal{F}_t) \times \mathcal{D} \to L^2(\mathcal{F}_t)$ are such that $P_t^{\text{ask}}(0,\cdot) = P_t^{\text{bid}}(0,\cdot) = 0$, $t \in \mathcal{T}$. We refer to Definition 2.3.2 for better understanding of the roles of $P_t^{\text{ask}}(\varphi, D^{\text{ask/bid},i})$ and $P_t^{\text{bid}}(\varphi, D^{\text{ask/bid},i})$ play in our theory.

(M3) The dividend process associated with holding 1 unit of the banking account is given by processes $D^0 = (0,\dots,0,1) \in \mathcal{D}$. Moreover, $\varphi \in L_+^\infty(\mathcal{F}_t)$ units of the banking account are purchased/sold at time t for price $P_t^{\text{ask}}(\varphi, D^0) = P_t^{\text{bid}}(\varphi, D^0) = \varphi \cdot 1 = \varphi$. In particular, $P_t^{\text{ask}}(1, D^0) = P_t^{\text{bid}}(1, D^0) = 1$.[2]

Let $\mathcal{M}$ be the set of all dividend processes in this market, i.e.

$$\mathcal{M} = \{D^0, D^{\text{ask/bid},i}, i = 1,\dots,K\}.$$

We will use the notation $(\mathcal{M}, P^{\text{ask}}, P^{\text{bid}})$ to denote our market model.

[2]This is consistent with our convention that all prices are discounted by the banking account, so that the price of one unit of the banking account is 1 at any time $t \in \mathcal{T}$.

The market model $(\mathcal{M}, P^{\text{ask}}, P^{\text{bid}})$ such that $D^{\text{ask},i} = D^{\text{bid},i}$, $i = 1, \ldots, K$, and $P_t^{\text{ask}}(\varphi, D) = P_t^{\text{bid}}(\varphi, D)$, for any $D \in \mathcal{M}$, is called frictionless market model.

Remark 2.3.1. *In accordance with our framework, it is generally assumed that* $D^{ask,i} \neq D^{bid,i}$, $i = 1, \ldots, K$, *and* $P_t^{ask}(\varphi, D) \neq P_t^{bid}(\varphi, D)$, $D \in \mathcal{M}$. *We also remark that in this thesis we do not postulate that the prices* $P_t^{ask}(\varphi, D)$ *and* $P_t^{bid}(\varphi, D)$ *are homogeneous (of degree one) in* φ. *In other words, we acknowledge the fact that in practice the unit price of a security typically depends on the size of the position in the security (cf. Example 2.3.3 below). This is due, for the most part, to market liquidity considerations. As we shall see later, the bid/ask prices generated by our acceptability method are not necessarily homogeneous.*

We now illustrate the processes introduced above in the context of dividend paying stock and Credit Default Swap (CDS) contract.

Example 2.3.1. *Denote by* S_T *the fundamental value associated to 1 share of a dividend paying stock after dividend payment at time* T. *The dividend paid by 1 share of the stock at each time* $t = 1, \ldots, T$, *is denoted by* D_t, *regardless of what position the investor is in. Therefore, the dividend process associated with 1 share of the stock is*

$$D^{ask} = D^{bid} = \{0, D_1, \ldots, D_{T-1}, D_T + S_T\}.$$

In this case, the ex-dividend ask and bid price process P^{ask} *and* P^{bid} *are the market quoted prices for selling, respectively buying stock* S; *see also Example 2.3.3.*

Example 2.3.2. *A CDS contract is an agreement between the protection buyer and the protection seller, in which the protection buyer pays a regular fixed premium up to occurrence of a pre-specified credit event, in return, the seller promises a compensation to the buyer. Typically, CDS contracts are traded on over-the-counter markets in which dealers quote CDS spreads to investors. Consider a CDS contract that is* *initiated at* $t = 0$*, expires at* $t = T$ *with nominal value* $\mathcal{N}$*, the protection buyer pays a spread* κ^{ask} *to the dealer at each time node in exchange for a compensation* δ *at default time* τ*; the protection seller receives a spread* κ^{bid} *from the dealer at each time node and he needs to pay* δ *to the dealer at* τ*. Please note that* κ^{ask}*,* κ^{bid} *and* δ *all depends on* $\mathcal{N}$*, and such dependence is not necessarily linear.*

The dividend processes D^{ask} *and* D^{bid} *associated to buying and selling the CDS with specifications above, respectively, satisfy*

$$\sum_{s=0}^{t} D_s^{ask} := \mathbb{1}_{\{\tau \le t\}}\delta - \kappa^{ask}\sum_{s=1}^{t}\mathbb{1}_{\{s<\tau\}}, \ \sum_{s=0}^{t} D_t^{bid} := \mathbb{1}_{\{\tau \le t\}}\delta - \kappa^{bid}\sum_{s=1}^{t}\mathbb{1}_{\{s<\tau\}},$$

for $t = 1, \ldots, T$*. In this case, the ex-dividend ask and bid price process* P^{ask} *and* P^{bid} *specify the mark-to-market values of the CDS. In general,* P^{ask} *and* P^{bid} *are also not positively homogeneous with respect to* $\mathcal{N}$*; a property confirmed in a personal communication with practitioners trading CDS contracts.*

We close this subsection by illustrating the inhomogeneity of prices in a order-driven market.

Example 2.3.3. *In an order-driven market, orders to buy and sell are centralized in a limit order book available to market participants and orders are executed against the best available offers in the limit order book.*

Table 2.1. Order Book of AAPL (Yahoo Finance 10:46AM EST 12/04/2014)

Bid Price	Bid Size	Ask Price	Ask Size
116.59	400	116.61	200
116.58	400	116.62	700
116.57	800	116.63	543
116.56	500	116.64	643
116.55	543	116.65	343

Table 2.1 is the limit order book of Apple Inc (AAPL) publicly traded stock. As we can see, there are up to 200 shares available for purchase at a price of \$116.61 per share. Hence, $P^{ask}(1) = 116.61$, *and for* $0 \leq \varphi \leq 200$, $P^{ask}(\varphi) = \varphi P^{ask}(1)$. *Similarly, for* $200 < \varphi \leq 900$, *we have that* $P^{ask}(\varphi) = 200 \cdot 116.61 + (\varphi - 200) \cdot 116.62 > \varphi P^{ask}(1)$. *Thus, the ask price* $P^{ask}(\cdot)$ *is not homogenous in number of shares traded. Moreover, it is easy to note that* $P^{ask}(\cdot)$ *is a convex function. Similarly, the function* $P^{bid}(\cdot)$ *is not homogenous and it is concave.*

We need to stress that it is not our goal to build a stylized model for the time evolution of a limit order book. The above example serves to show that, generally speaking, the market prices are non-homogeneous functions of the order size. This feature is one of the stylized market features that our model captures. We believe that this feature is invariant of any specific time resolution at which trading is done, so our discrete time model, which does not refer to any specific time scale is well placed to model this feature. In particular, the model is meant to deal with valuation of complex financial products that are not traded via high frequency trading.

2.3.2 Self-financing Trading Strategies and Arbitrage. Due to nonlinearity of the prices and presence of transaction costs, the classical definition of self-financing trading strategy and arbitrage are not suitable for the market model proposed above. In this section we define the notion of self-financing trading strategy by using the general concept that a self-financing trading strategy is a trading strategy that does not allow injection or substraction of money during trading periods. Similarly, the notion of arbitrage is build on the idea that a self-financing trading strategy can not yield a riskless profit. Moreover, following market practice, we will allow that an investor can simultaneously have both long and short positions of the same security at the same time, i.e. the long and short positions in the same security at the same time are not necessarily netted out.

Definition 2.3.1. *A trading strategy is a predictable process* $\phi := \big\{(\phi_t^0, \phi_t^{l,1}, \phi_t^{s,1}, \ldots, \phi_t^{l,K}, \phi_t^{s,K})\big\}_{t=1}^T$*, where* $\phi_t^0 \in L^\infty(\mathcal{F}_{t-1})$ *is the number of units of banking account held from time* $t-1$ *to* t*;* $\phi_t^{l,i} \in L_+^\infty(\mathcal{F}_{t-1})$ *is the number of shares in long position of security* i *held from time* $t-1$ *to* t*; and* $\phi_t^{s,i} \in L_+^\infty(\mathcal{F}_{t-1})$ *is the number of shares in short position of security* i *held from time* $t-1$ *to* t*. Sometimes, we will use the notation* $\phi_t^i = (\phi_t^{l,i}, \phi_t^{s,i})$*, for* $i = 1, \ldots, K$*.*

Definition 2.3.2. *Let* ϕ *be a trading strategy.*

(V1) The set-up cost process $\widetilde{V}(\phi)$ *associated with* ϕ *is defined as*

$$\widetilde{V}_t(\phi) = \phi_{t+1}^0 + \sum_{i=1}^K \Big(P_t^{ask}(\phi_{t+1}^{l,i}, D^{ask,i}) - P_t^{bid}(\phi_{t+1}^{s,i}, D^{bid,i})\Big), \qquad t = 0, \ldots, T-1.$$

(V2) The liquidation value process $V(\phi)$ *associated with* ϕ *is defined as*

$$\begin{aligned} V_t(\phi) = \phi_t^0 &+ \sum_{i=1}^K \Big(P_t^{bid}(\phi_t^{l,i}, D^{ask,i}) - P_t^{ask}(\phi_t^{s,i}, D^{bid,i})\Big) \\ &+ \sum_{i=1}^K \Big(\phi_t^{l,i} D_t^{ask,i} - \phi_t^{s,i} D_t^{bid,i}\Big), \qquad t = 1, \ldots, T. \end{aligned}$$

For each $t \in \mathcal{T}$, an investor could have both long and short positions of a security at the same time. The process $\widetilde{V}(\phi)$ represents the cost of setting up the portfolio ϕ, and $V_t(\phi)$ is interpreted as the liquidation value of the portfolio at time t (before any time t transactions), including any dividends acquired from time $t-1$ to time t. Note that, generally speaking, $V_t(\phi) \neq \widetilde{V}_t(\phi)$ due to transaction costs and non-homogeneity of P^{ask} and P^{bid}.

Definition 2.3.3. *A trading strategy* ϕ *is self-financing if*

$$\begin{aligned} &\Delta\phi_{t+1}^0 + \sum_{i=1}^K \Big(\mathbb{1}_{\Delta\phi_{t+1}^{l,i} \geq 0} P_t^{ask}(\Delta\phi_{t+1}^{l,i}, D^{ask,i}) - \mathbb{1}_{\Delta\phi_{t+1}^{l,i} < 0} P_t^{bid}(-\Delta\phi_{t+1}^{l,i}, D^{ask,i}) \qquad (2.24) \\ &- \mathbb{1}_{\Delta\phi_{t+1}^{s,i} \geq 0} P_t^{bid}(\Delta\phi_{t+1}^{s,i}, D^{bid}) + \mathbb{1}_{\Delta\phi_{t+1}^{s,i} < 0} P_t^{ask}(-\Delta\phi_{t+1}^{s,i}, D^{bid})\Big) \\ &= \sum_{i=1}^K (\phi_t^{l,i} D_t^{ask,i} - \phi_t^{s,i} D_t^{bid,i}). \qquad (2.25) \end{aligned}$$

for $t = 0, \ldots, T-1$.

Definition 2.3.3 provides a natural interpretation of self-financing condition in market with friction. The cash flows that are being bought or sold should depend on the both positions before and after re-balance at each time t. All the money that is used for getting to the new position is equal to the dividends acquired from time $t-1$ to time t. Therefore, no money flows in or out of the portfolio.

Next, we will introduce the concept of arbitrage that is relevant for our theory.

Definition 2.3.4. *An arbitrage opportunity at time t, $t \in \mathcal{T}$, is a self-financing trading strategy, such that $V_T(\phi) - \widetilde{V}_t(\phi) \geq 0$ and $\mathbb{P}(V_T(\phi) - \widetilde{V}_t(\phi) > 0) > 0$. We call a market arbitrage free at time t if there exists no arbitrage opportunity in the model at time t.*

It needs to be observed that in the above definition we consider the difference between the liquidation value of the portfolio at maturity and the set-up cost of the portfolio at time t (recall that the interest rates are taken to be zero, so there is no time value of money). Thus, the above definition regards the realized net change in trader's wealth.

In what follows, we will provide two other characterizations of arbitrage opportunities, which are useful in our work. Towards this end, we first define the following sets,

$$\mathcal{S}(t) := \begin{cases} \{\phi : \phi \text{ is self-financing, } \widetilde{V}_0(\phi) = 0\}, & t = 0, \\ \{\phi : \phi \text{ is self-financing, } \phi_s = 0 \text{ for all } s \leq t\}, & t = 1, \ldots, T. \end{cases}$$

Note that for $\phi \in \mathcal{S}(t)$, we have $\phi_t = 0$. Due to our assumption that $P_t^{\text{bid}}(0, \cdot) = P_t^{\text{ask}}(0, \cdot) = 0$ and ϕ is self-financing, we have that $\widetilde{V}_t(\phi) = V_t(\phi) = 0$.

Next, we introduce the set of cash flows generated by strategies in $\mathcal{S}(t)$:

$$\mathcal{H}^0(t) = \Big\{ \Big(\underbrace{0, \dots, 0}_{t}, \Delta V_{t+1}(\phi), \dots, \Delta V_T(\phi) \Big) : \phi \in \mathcal{S}(t) \Big\}, \qquad t \in \mathcal{T}. \tag{2.26}$$

Since $P_t^{\text{ask}}, P_t^{\text{bid}} : L_+^{\infty}(\mathcal{F}_t) \times \mathcal{D} \to \mathcal{D}$, we have that $\widetilde{V}_t(\phi)$, $V_t(\phi) \in L^2(\mathcal{F}_t)$ for any self-financing trading strategy ϕ, and therefore $\mathcal{H}^0(t) \subset \mathcal{D}$.

Next result gives two characterization of arbitrage.

Proposition 2.3.1. *The following statements are equivalent:*

(1) There exists an arbitrage opportunity at time t.

(2) There exists a strategy $\xi \in \mathcal{S}(t)$, such that $V_T(\xi) \geq 0$ and $\mathbb{P}(V_T(\xi) > 0) > 0$.

(3) There exists a cash flow $(0, \dots, 0, H_{t+1}, \dots, H_T) \in \mathcal{H}^0(t)$, such that $\sum_{s=t+1}^{T} H_s \geq 0$ and $\mathbb{P}(\sum_{s=t+1}^{T} H_s > 0) > 0$.

Proof. For a fixed $t \in \mathcal{T}$, we will show that (1) is equivalent to (2), and (2) is equivalent to (3).

(2) $\Rightarrow$ (1) Assume that there exists $\xi \in \mathcal{S}(t)$ such that $V_T(\xi) \geq 0$ and $\mathbb{P}(V_T(\xi) > 0) > 0$. Since $\xi \in \mathcal{S}(t)$, then $\widetilde{V}_t(\xi) = 0$, and (1) follows immediately.

(1) $\Rightarrow$ (2) Assume that ϕ is an arbitrage opportunity at time t. We define a trading strategy ξ as follows

$$\begin{aligned} \xi_u^0 = \psi_u^{l,i} = \xi_u^{s,i} = 0, &\qquad u = 0, \dots, t, \ i = 0, \dots, K, \\ \xi_u^0 = \phi_u^0 - \widetilde{V}_t(\phi), &\qquad u = t+1, \dots, T, \\ \xi_u^{l/s,i} = \phi_u^{l/s,i}, &\qquad u = t+1, \dots, T, \ i = 1, \dots, K. \end{aligned}$$

It is straightforward to see that $\widetilde{V}_0(\xi) = 0$, and $\xi_u = 0$ for $u \leq t$. To show that ξ is self-financing, first notice that

$$\begin{aligned}
\Delta\xi^0_{t+1} + \sum_{i=1}^{K} \Big(& \mathbb{1}_{\Delta\xi^{l,i}_{t+1}\geq 0} P^{\text{ask}}_t(\Delta\xi^{l,i}_{t+1}, D^{\text{ask},i}) - 1_{\Delta\xi^{l,i}_{t+1}<0} P^{\text{bid}}_t(-\Delta\xi^{l,i}_{t+1}, D^{\text{ask},i}) \\
& - \mathbb{1}_{\Delta\xi^{s,i}_{t+1}\geq 0} P^{\text{bid}}_t(\Delta\xi^{s,i}_{t+1}, D^{\text{bid},i}) + \mathbb{1}_{\Delta\xi^{s,i}_{t+1}<0} P^{\text{ask}}_t(-\Delta\xi^{s,i}_{t+1}, D^{\text{bid},i}) \Big) \\
=& \phi^0_{t+1} - \widetilde{V}_t(\phi) + \sum_{i=1}^{K} \Big(P^{\text{ask}}_t(\phi^{l,i}_{t+1}, D^{\text{ask},i}) - P^{\text{bid}}_t(\phi^{s,i}_{t+1}, D^{\text{bid},i}) \Big) \\
=& \widetilde{V}_t(\phi) - \widetilde{V}_t(\phi) = 0 \\
=& \sum_{i=1}^{K} (\xi^{l,i}_t D^{\text{ask},i}_t - \xi^{i,s}_t D^{\text{bid},i}_t).
\end{aligned} \tag{2.27}$$

Since $\Delta\xi_u = \Delta\phi_u$ for any $u \geq t+1$, and ϕ is a self-financing trading strategy, then also in view of (2.27), we conclude that ξ is a self-financing strategy. Hence, $\xi \in \mathcal{S}(t)$, and $V_T(\xi) = V_T(\phi) - V_t(\phi)$ which implies that $V_T(\xi) \geq 0$, $\mathbb{P}(V_T(\xi) > 0) > 0$.

(2) $\Rightarrow$ (3) Assume that $\xi \in \mathcal{S}(t)$, and $V_T(\xi) \geq 0$, $\mathbb{P}(V_T(\xi) > 0) > 0$. Then, we define the process H as follows

$$H_s := \begin{cases} 0, & s = 0, \ldots, t, \\ \Delta V_s(\xi), & s = t+1, \ldots, T. \end{cases}$$

Then $H \in \mathcal{H}^0(t)$, $\sum_{s=t+1}^{T} H_s = \sum_{s=0}^{T} \Delta V_s(\xi) = V_T(\xi) \geq 0$, and thus $\mathbb{P}(\sum_{s=t+1}^{T} H_s > 0) = \mathbb{P}(V_T(\xi) > 0) > 0$.

(3) $\Rightarrow$ (2) Now, suppose that there exists a cash flow $\big(0, \ldots, 0, \widehat{H}_{t+1}, \ldots, \widehat{H}_T\big) \in \mathcal{H}^0(t)$ such that $\sum_{s=t+1}^{T} \widehat{H}_s \geq 0$ and $\mathbb{P}(\sum_{s=t+1}^{T} \widehat{H}_s > 0) > 0$. Then, by definition of $\mathcal{H}^0(t)$ there exists a $\xi \in \mathcal{S}(t)$ such that $V_t(\xi) = 0$, $\Delta V_s(\xi) = \widehat{H}_s$, $s \in \{t+1, \ldots, T\}$ and

$$V_T(\xi) = \sum_{s=0}^{T} \Delta V_s(\xi) = \sum_{s=t+1}^{T} \widehat{H}_s \geq 0.$$

Moreover, $\mathbb{P}(V_T(\xi) > 0) = \mathbb{P}(\sum_{s=t+1}^{T} \widehat{H}_s > 0) > 0$. Thus, (3) holds true.

This concludes the proof. □

Remark 2.3.2. *With the results from Proposition 2.3.1 at hand, we say that the no-arbitrage condition (NA) for $\mathcal{H}^0(t)$ holds if (3) does not hold, and throughout this section, we will characterize arbitrage opportunities by properties (2) or (3) as in Proposition 2.3.1.*

Clearly, since $\mathcal{S}(t+1) \subset \mathcal{S}(t)$, absence of arbitrage at time $t \in \{0, 1, \ldots, T-1\}$ implies absence of arbitrage at any future times s, $s = t+1, \ldots, T-1$. In particular, if a market is arbitrage free at time 0, then such market is arbitrage free at any time $t \in \mathcal{T}$. Hence, to show that there is no arbitrage opportunity in the market, it is enough to show that there is no arbitrage at time 0. Accordingly, we have the following definition.

Definition 2.3.5. *Our market model is called arbitrage free if there exists no arbitrage opportunity in the model at time 0.*

It is important to observe though, that, contrary to the classical frictionless market model, absence of arbitrage at time $t = 1, \ldots, T-1$, in models considered here does not (in general) imply absence of arbitrage at time s, where $s < t$. This will be illustrated in the following example.

Example 2.3.4. *Let $T = 2$, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, and we consider a market with one banking account and one security paying no dividend. Assume that the pricing operator is homogeneous with respect to the number shares traded, and the price process of the security is given in Table 2.2.*

Consider the trading strategy $\xi_1 = (-10, 1)$, $\xi_2(\omega_1, \omega_2) = (1, 0)$ and $\xi_2(\omega_3, \omega_4) = (0, 0)$. Thus, ξ is a self-financing strategy such that $\widetilde{V}_0 = 0$, $V_1(\omega_1, \omega_2) = V_2(\omega_1, \omega_2) = 1$ and $V_1(\omega_3, \omega_4) = V_2(\omega_3, \omega_4) = 0$. According to Definition 2.3.4, it is an arbitrage opportunity at time 0. However, it is not hard to observe that for any $\varphi \in \mathcal{S}(1)$, φ could not be an arbitrage opportunity at time 1.

Table 2.2. Stock Price Dynamics, Example 2.3.4.

	$P^{\text{ask}}(\omega_1)/P^{\text{bid}}(\omega_1)$	$P^{\text{ask}}(\omega_2)/P^{\text{bid}}(\omega_2)$	$P^{\text{ask}}(\omega_3)/P^{\text{bid}}(\omega_3)$	$P^{\text{ask}}(\omega_4)/P^{\text{bid}}(\omega_4)$
$T=0$	10/10	10/10	10/10	10/10
$T=1$	12/11	12/11	11/10	11/10
$T=2$	13/12	11/10	12/11	10/9

2.3.3 Pricing Operators. In this section we will introduce some pricing operators for cash flows $D \in \mathcal{D}$ through an acceptability method and show some properties of these prices. Then, in the next section, we will show that a market model, in which the fundamental assets are priced according to our pricing operators, satisfies the properties postulated in (M2) and (M3), and that this market is arbitrage free in the sense of Definition 2.3.4.

For any random variable a that is $\mathcal{F}_t$-measurable, and for any process $X = (X_t)_{t\in\mathcal{T}}$, we will use the following notation:

$$\delta_t(a) = \mathbb{1}_{\{t\}}a := \{0, \ldots, 0, a, 0, \ldots, 0\},$$
$$\delta_t^+(X) = \{0, \ldots, 0, X_{t+1}, \ldots, X_T\}.$$

For any $D \in \mathcal{D}$, $\delta_t^+(D)$ represents the the future cash flow, at time $t \in \mathcal{T}$, of the dividend stream D. We are going to evaluate this future dividend cash flow, in other words, to calculate the ex-dividend prices of D, by using an acceptability based method.

Assume that an investor wants to buy φ shares of cash flow $D \in \mathcal{D}$ at time $t \in \mathcal{T}$, where $\varphi \in \mathcal{F}_t$, $\varphi \geq 0$, then the market as the counterparty will charge $P_t^{\text{ask}}(\varphi, D)$ at time t and promises to deliver $\delta_t^+(D)$ to the buyer. Thus, the corresponding cash-flow from the perspective of the market is $\{0, \ldots, 0, P_t^{\text{ask}}(\varphi, D), -\varphi D_{t+1}, \ldots, -\varphi D_T\}$. To decide the proper price $P_t^{\text{ask}}(\varphi, D)$, the market will choose the smallest $P_t^{\text{ask},D}(\varphi)$ such that

$\{0,\ldots,0,P_t^{\text{ask}}(\varphi,D),-\varphi D_{t+1},\ldots,-\varphi D_T\}$ is acceptable with respect to some acceptability index α^g at some level γ. Similarly, the market will choose the largest $P_t^{\text{bid}}(\varphi,D)$ such that the cash flow
$\{0,\ldots,0,-P_t^{\text{bid}}(\varphi,D),\varphi D_{t+1},\ldots,\varphi D_T\}$ is acceptable at level γ, when an investor is selling $\varphi \geq 0$ shares of $D \in \mathcal{D}$ at time $t \in \mathcal{T}$.

Throughout this section, we will always consider the family of drivers $g = (g_x)_{x>0}$ that satisfy Assumption G. We proceed by defining the acceptability ask price $a_t^{g,\gamma}$ and the acceptability bid price $b_t^{g,\gamma}$.

Definition 2.3.6. *Let $g = (g_x)_{x>0}$ be a family of drivers. The acceptability ask price of $\varphi \in L_+^\infty(\mathcal{F}_t)$ shares of the cash flow $D \in \mathcal{D}$, at level γ, at time $t \in \mathcal{T}$ is defined as*

$$a_t^{g,\gamma}(\varphi,D) = \operatorname{ess\,inf}\{a \in L^2(\mathcal{F}_t) : \alpha_t^g(\delta_t(a) - \delta_t^+(\varphi D)) \geq \gamma\}; \tag{2.28}$$

and the acceptability bid price of $\varphi \in L_+^\infty(\mathcal{F}_t)$ shares of $D \in \mathcal{D}$, at level γ, at time $t \in \mathcal{T}$ is defined as

$$b_t^{g,\gamma}(\varphi,D) = \operatorname{ess\,sup}\{b \in L^2(\mathcal{F}_t) : \alpha_t^g(\delta_t^+(\varphi D) - \delta_t(b)) \geq \gamma\}. \tag{2.29}$$

For $\varphi \in L_+^\infty(\mathcal{F}_t)$ and $D \in \mathcal{D}$, $a_t^{g,\gamma}(\varphi,D)$ and $b_t^{g,\gamma}(\varphi,D)$ are the ex-dividend prices at time t, therefore, they do not account for $D_0,\ldots,D_t$.

Remark 2.3.3. *Note that in Definition 2.3.6, φ is a $\mathcal{F}_t$-measurable random variable, thus by applying the pricing operators $a_t^{g,\gamma}$ and $b_t^{g,\gamma}$ to cash flows generated by any trading strategy ϕ, we will get well-defined set-up cost process $\widetilde{V}_t(\phi)$ and liquidation value process $V_t(\phi)$. Also by observing the fact that $a_t^{g,\gamma}(\varphi,D) = a_t^{g,\gamma}(1,\varphi D)$ and $b_t^{g,\gamma}(\varphi,D) = b_t^{g,\gamma}(1,\varphi D)$, we will prove most results for $a_t^{g,\gamma}(1,D)$ and $b_t^{g,\gamma}(1,D)$. Then such results are also true for $a_t^{g,\gamma}(\varphi,D)$ and $b_t^{g,\gamma}(\varphi,D)$.*

Remark 2.3.4. *We call $a_t^{g,\gamma}(1,D)$ the time t acceptability ask price of D at level γ, and $b_t^{g,\gamma}(1,D)$ the time t acceptability bid price of D at level γ. For simplicity, we will use the notation $a_t^{g,\gamma}(D) = a_t^{g,\gamma}(1,D)$ and $b_t^{g,\gamma}(D) = b_t^{g,\gamma}(1,D)$.*

Remark 2.3.5. *We observe that in Definition 2.3.6 φ is non-negative. This is consistent with the term "buy/sell φ shares of some security" which is used in practice.*

Next, we provide some important properties of acceptability ask and bid prices.

Theorem 2.3.1. *The acceptability ask and bid prices of $D \in \mathcal{D}$, at level $\gamma > 0$, at time $t \in \mathcal{T}$, satisfy the following properties:*

P1. Representation:

$$a_t^{g,\gamma}(D) = \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big],$$

$$b_t^{g,\gamma}(D) = -\mathcal{E}_{g_\gamma}\Big[-\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big].$$

P2. Non-negative Spread:

$$a_t^{g,\gamma}(D) \geq b_t^{g,\gamma}(D).$$

P3. Convexity and Concavity:

$$a_t^{g,\gamma}(\lambda D^1 + (1-\lambda)D^2) \leq \lambda a_t^{g,\gamma}(D^1) + (1-\lambda)a_t^{g,\gamma}(D^2),$$

$$b_t^{g,\gamma}(\lambda D^1 + (1-\lambda)D^2) \geq \lambda b_t^{g,\gamma}(D^1) + (1-\lambda)b_t^{g,\gamma}(D^2),$$

for $D^1, D^2 \in \mathcal{D}$, $\lambda \in L_+^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$.

P4. Market Impact:

$$a_t^{g,\gamma}(\lambda\varphi, D) \leq \lambda a_t^{g,\gamma}(\varphi, D),\ b_t^{g,\gamma}(\lambda\varphi, D) \geq \lambda b_t^{g,\gamma}(\varphi, D),\ \lambda, \varphi \in L_+^\infty(\mathcal{F}_t),\ 0 \leq \lambda \leq 1;$$

$$a_t^{g,\gamma}(\lambda\varphi, D) \geq \lambda a_t^{g,\gamma}(\varphi, D),\ b_t^{g,\gamma}(\lambda\varphi, D) \leq \lambda b_t^{g,\gamma}(\varphi, D),\ \lambda, \varphi \in L_+^\infty(\mathcal{F}_t), \quad \lambda \geq 1.$$

P5. Time Consistency:

$$a_t^{g,\gamma}(D) = a_t^{g,\gamma}(\delta_{t+1}(D_{t+1} + a_{t+1}^{g,\gamma}(D))),$$

$$b_t^{g,\gamma}(D) = b_t^{g,\gamma}(\delta_{t+1}(D_{t+1} + b_{t+1}^{g,\gamma}(D))).$$

P6. Linearity (if the drivers are linear): If $g_\gamma(t,z) = x(t)z$, $t \in \mathcal{T}$. Then, there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that

$$a_t^{g,\gamma}(\varphi, D) = b_t^{g,\gamma}(\varphi, D) = \varphi \mathbb{E}_{\mathbb{Q}}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big],$$

for $\varphi \in L_+^\infty(\mathcal{F}_t)$.

Proof. In P1, P3, P4 and P5, we will prove the results only for acceptability ask prices; the case of acceptability bid prices is treated similarly.

P1. Due to G3 of Assumption G, by similar arguments as in Theorem 2.2.2, we get that $\alpha_t^g(X) \geq \gamma$ for $\gamma > 0$ is equivalent to the fact that $\rho_t^{g_\gamma}(X) \leq 0$. Also, in view of the definition of acceptability ask price, we have

$$\begin{aligned}
a_t^{g,\gamma}(D) &= \operatorname{ess\,inf}\{a \in L^2(\mathcal{F}_t) : \alpha_t^g(\delta_t(a) - \delta_t^+(D)) \geq \gamma\} \\
&= \operatorname{ess\,inf}\left\{a \in L^2(\mathcal{F}_t) : \mathcal{E}_{g_\gamma}\Big[-a + \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] \leq 0\right\} \\
&= \operatorname{ess\,inf}\left\{a \in L^2(\mathcal{F}_t) : a \geq \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big]\right\} \\
&= \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big].
\end{aligned}$$

P2. By convexity of g-expectation, we have that

$$\frac{1}{2}\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] + \frac{1}{2}\mathcal{E}_{g_\gamma}\Big[-\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] \geq \mathcal{E}_{g_\gamma}\Big[\frac{1}{2}\Big(\sum_{s=t+1}^{T} D_s - \sum_{s=t+1}^{T} D_s\Big) \Big| \mathcal{F}_t\Big] = 0.$$

Hence, due to property P1, we get $a_t^{g,\gamma}(D) \geq b_t^{g,\gamma}(D)$.

P3. Property P3 follows from convexity of $g_\gamma(t,\cdot)$, convexity of the g-expectation, and from P1.

P4. By taking $D^1 = D$, and $D^2 = 0$, for $\lambda, \varphi \in L^\infty(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$, we have $a_t^{g,\gamma}(\lambda\varphi D) \leq \lambda a_t^{g,\gamma}(\varphi D)$. Since $a_t^{g,\gamma}(\lambda\varphi D) = a_t^{g,\gamma}(\lambda\varphi, D)$ and $a_t^{g,\gamma}(\varphi D) = a_t^{g,\gamma}(\varphi, D)$, we immediately get that $a_t^{g,\gamma}(\lambda\varphi, D) \leq \lambda a_t^{g,\gamma}(\varphi, D)$.

For $\lambda \in L^\infty(\mathcal{F}_t)$, $\lambda \geq 1$, $D \in \mathcal{D}$, we have that $a_t^{g,\gamma}(\varphi, \frac{D}{\lambda}) = a_t^{g,\gamma}(\frac{\varphi}{\lambda}, D) \leq \frac{1}{\lambda} a_t^{g,\gamma}(\varphi, D)$.

P5. According to P1 and Proposition 2.1.1 (iii), we have that

$$\begin{aligned} a_t^{g,\gamma}(\delta_{t+1}(D_{t+1} + a_{t+1}^{g,\gamma}(D))) &= \mathcal{E}_{g_\gamma}\Big[D_{t+1} + \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+2}^{T} D_s \Big| \mathcal{F}_{t+1}\Big]\Big|\mathcal{F}_t\Big] \\ &= \mathcal{E}_{g_\gamma}\Big[\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_{t+1}\Big]\Big|\mathcal{F}_t\Big] = \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] \\ &= a_t^{g,\gamma}(D). \end{aligned}$$

P6. By Proposition 2.1.2, there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$, such that $\mathcal{E}_{g_\gamma}[X|\mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}[X|\mathcal{F}_t]$ for all $X \in L^2(\mathcal{F}_T)$. Then, using P1, for $\varphi \in \mathcal{F}_t$, we obtain

$$\begin{aligned} a_t^{g,\gamma}(\varphi, D) =& a_t^{g,\gamma}(1, \varphi D) = \mathcal{E}_{g_\gamma}\Big[\varphi \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] \\ =& \mathbb{E}_{\mathbb{Q}}\Big[\varphi \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] = \varphi \mathbb{E}_{\mathbb{Q}}\Big[\sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big]. \end{aligned}$$

Hence, $a_t^{g,\gamma}(\varphi, D) = \varphi \alpha_t^{g,\gamma}(1, D)$.

This concludes the proof.

□

Remark 2.3.6. *By Property P4 in Theorem 2.3.1 we have that $a_t^{g,\gamma}(\lambda, D) \geq \lambda a_t^{g,\gamma}(D)$, for any $\lambda \geq 1$. This indicates that if an investor is buying a cash flow, the price moves up against the buyer (the effect of market impact). Converse property holds for the bid price. Such property of acceptability ask and bid prices is consistent with real market quotes, as it was shown for equity markets in Example 2.3.3.*

In case of classical risk-neutral pricing the discounted cumulative dividend prices of cash flows are martingales under an equivalent martingale measure $\mathbb{Q}$. In our pricing framework, a similar 'martingale property' also holds true as shown in the next result. For a given dividend stream $D \in \mathcal{D}$, we define the *acceptability cumulative dividend prices* at time t as follows

$$a_t^{\mathrm{cld},g,\gamma}(D) := \sum_{s=0}^{t} D_s + a_t^{g,\gamma}(D),$$
$$b_t^{\mathrm{cld},g,\gamma}(D) := \sum_{s=0}^{t} D_s + b_t^{g,\gamma}(D).$$

As an immediate consequence of Theorem 2.3.1.P5, we obtain.

Corollary 2.3.1. *The acceptability ask and bid cumulative dividend prices of a cash-flow D, at level $\gamma > 0$ satisfy*

$$a_t^{cld,g,\gamma}(D) = a_t^{g,\gamma}(\delta_{t+1}(a_{t+1}^{cld,g,\gamma}(D))),$$
$$b_t^{cld,g,\gamma}(D) = b_t^{g,\gamma}(\delta_{t+1}(b_{t+1}^{cld,g,\gamma}(D))).$$

Remark 2.3.7. *This corollary is a counterpart of martingale property in case of linear pricing. The time t cumulative dividend price of D is equal to evaluating time $t+1$ cumulative dividend price at time t. We call such property the time consistency of acceptability ask/bid prices.*

So far, we have showed that if the market picks the same level $\gamma > 0$ for the given family of drivers g on both buying and selling side, then we have some nice properties of the ask and bid prices. On the other hand, in general, market participants will choose different acceptability levels or different acceptability indices (different family of drivers) for buying and/or selling. In the rest of this section, we will provide some results regarding such possibilities.

Proposition 2.3.2. *Let g^1 and g^2 be two families of drivers. Then, $a_t^{g^1,\gamma_1}(D) \geq b_t^{g^2,\gamma_2}(D)$, for any $D \in \mathcal{D}$, $t \in \mathcal{T}$, $\gamma_1, \gamma_2 > 0$.*

Proof. We will prove the statement recursively, backward in time component.

Let $A_t := \{\omega \in \Omega : g^1_{\gamma_1}(\omega, t, z) \geq g^2_{\gamma_2}(\omega, t, z), z \in \mathbb{R}\}$, $t \in \mathcal{T} \setminus \{0\}$. since the drivers are predictable, both A_t and A_t^c are $\mathcal{F}_{t-1}$ measurable, .

By definition of A_T, we have that $\mathbb{1}_{A_T} g^1_{\gamma_1}(T, z) \geq \mathbb{1}_{A_T} g^2_{\gamma_2}(T, z)$ and $\mathbb{1}_{A_T^c} g^1_{\gamma_1}(T, z) \leq \mathbb{1}_{A_T^c} g^2_{\gamma_2}(T, z)$, for all $z \in \mathbb{R}$. Hence, in view of Theorem 2.1.2 and Theorem 2.3.1, we get that

$$\mathbb{1}_{A_T} \mathcal{E}_{g^1_{\gamma_1}}[D_T|\mathcal{F}_{T-1}] \geq \mathbb{1}_{A_T} \mathcal{E}_{g^2_{\gamma_2}}[D_T|\mathcal{F}_{T-1}] \geq -1_{A_T} \mathcal{E}_{g^2_{\gamma_2}}[-D_T|\mathcal{F}_{T-1}],$$
$$\mathbb{1}_{A_T^c} \mathcal{E}_{g^1_{\gamma_1}}[D_T|\mathcal{F}_{T-1}] \geq -\mathbb{1}_{A_T^c} \mathcal{E}_{g^1_{\gamma_1}}[-D_T|\mathcal{F}_{T-1}] \geq -1_{A_T^c} \mathcal{E}_{g^2_{\gamma_2}}[-D_T|\mathcal{F}_{T-1}].$$

Therefore, $\mathbb{1}_{A_T} a^{g^1,\gamma_1}_{T-1}(D) \geq \mathbb{1}_{A_T} b^{g^2,\gamma_2}_{T-1}(D)$ and $\mathbb{1}_{A_T^c} a^{g^1,\gamma_1}_{T-1}(D) \geq \mathbb{1}_{A_T^c} b^{g^2,\gamma_2}_{T-1}(D)$, and thus the statement holds true for $t = T$.

Note that by definition of acceptability cumulative dividend prices, we have that

$$a^{g^1,\gamma_1}_{T-2}(D) + \sum_{s=1}^{T-2} D_s = a^{\text{cld},g^1,\gamma_1}_{T-2}(D),$$
$$b^{g^2,\gamma_2}_{T-2}(D) + \sum_{s=1}^{T-2} D_s = b^{\text{cld},g^2,\gamma_2}_{T-2}(D).$$

In view of Proposition 2.3.1, we also have that that

$$\begin{aligned} a^{\text{cld},g^1,\gamma_1}_{T-2}(D) =& \mathcal{E}_{g^1_{\gamma_1}}\left[a^{g^1,\gamma_1}_{T-1}(D) + \sum_{s=1}^{T-1} D_s \Big| \mathcal{F}_{T-2}\right] \\ =& \mathcal{E}_{g^1_{\gamma_1}}\left[a^{g^1,\gamma_1}_{T-1}(D) + D_{T-1} \Big| \mathcal{F}_{T-2}\right] + \sum_{s=1}^{T-2} D_s, \\ b^{\text{cld},g^2,\gamma_2}_{T-2}(D) =& -\mathcal{E}_{g^2_{\gamma_2}}\left[-b^{g^2,\gamma_2}_{T-1}(D) - \sum_{s=1}^{T-1} D_s \Big| \mathcal{F}_{T-2}\right] \\ =& -\mathcal{E}_{g^2_{\gamma_2}}\left[-b^{g^2,\gamma_2}_{T-1}(D) - D_{T-1} \Big| \mathcal{F}_{T-2}\right] + \sum_{s=1}^{T-2} D_s. \end{aligned}$$

Thus,

$$a^{g^1,\gamma_1}_{T-2}(D) = \mathcal{E}_{g^1_{\gamma_1}}[a^{g^1,\gamma_1}_{T-1}(D) + D_{T-1}|\mathcal{F}_{T-2}]$$

and

$$b^{g^2,\gamma_2}_{T-2}(D) = \mathcal{E}_{g^2_{\gamma_2}}[b^{g^2,\gamma_2}_{T-1}(D) + D_{T-1}|\mathcal{F}_{T-2}].$$

In view of above, we have $a^{g^1,\gamma_1}_{T-1}(D) + D_{T-1} \geq b^{g^2,\gamma_2}_{T-1}(D) + D_{T-1}$. Consequently, applying again the comparison Theorem 2.1.2 and Theorem 2.3.1, we deduce

$$\begin{aligned}\mathbb{1}_{A_{T-1}}\mathcal{E}_{g^1_{\gamma_1}}[a^{g^1,\gamma_1}_{T-1}(D) + D_{T-1}|\mathcal{F}_{T-2}] &\geq \mathbb{1}_{A_{T-1}}\mathcal{E}_{g^2_{\gamma_2}}[b^{g^2,\gamma_2}_{T-1}(D) + D_{T-1}|\mathcal{F}_{T-2}] \\ &\geq -\mathbb{1}_{A_{T-1}}\mathcal{E}_{g^2_{\gamma_2}}[-b^{g^2,\gamma_2}_{T-1}(D) - D_{T-1}|\mathcal{F}_{T-2}],\end{aligned}$$

and

$$\begin{aligned}\mathbb{1}_{A^c_{T-1}}\mathcal{E}_{g^1_{\gamma_1}}[a^{g^1,\gamma_1}_{T-1}(D) + D_{T-1}|\mathcal{F}_{T-2}] &\geq -\mathbb{1}_{A^c_{T-1}}\mathcal{E}_{g^1_{\gamma_1}}[-a^{g^1,\gamma_1}_{T-1}(D) - D_{T-1}|\mathcal{F}_{T-2}] \\ &\geq -\mathbb{1}_{A^c_{T-1}}\mathcal{E}_{g^1_{\gamma_1}}[-b^{g^2,\gamma_2}_{T-1}(D) - D_{T-1}|\mathcal{F}_{T-2}].\end{aligned}$$

Therefore, $a^{g^1,\gamma_1}_{T-2}(D) \geq b^{g^2,\gamma_2}_{T-2}(D)$. We continue this backward procedure for any finite number of steps till $t = 0$.

The proof is complete. □

Proposition 2.3.2 shows that regardless of what drivers and what level of acceptability one chooses for buying and selling side, the ask price will be greater than the bid price. In particular, when the same family of drivers g is chosen for both trading sides, then $a^{g,\gamma_1}_t(D) \geq b^{g,\gamma_2}_t(D)$, for any $D \in \mathcal{D}$, $t \in \mathcal{T}$, $\gamma_1, \gamma_2 > 0$.

The next result shows that the bid-ask spread increases when acceptability level is increased.

Proposition 2.3.3. *For any $\gamma_2 \geq \gamma_1 > 0$, $t \in \mathcal{T}$, and any $D \in \mathcal{D}$,*

$$a^{g,\gamma_1}_t(D) \leq a^{g,\gamma_2}_t(D), \quad b^{g,\gamma_1}_t(D) \geq b^{g,\gamma_2}_t(D).$$

Proof. It is sufficient to note that

$$\{a \in \mathcal{F}_t : \alpha^g_t(\delta_t(a) - \delta^+_t(D)) \geq \gamma_2\} \subseteq \{a \in \mathcal{F}_t : \alpha^g_t(\delta_t(a) - \delta^+_t(D)) \geq \gamma_1\},$$

for any $\gamma_2 \geq \gamma_1$. Using the definition of acceptability ask and bid prices, the result follows at once.

□

Suppose that two counterparties A and B are looking to make a trade on a cash flow D at time t, such as, A is willing to sell $\varphi \in L^{\infty}_{+}(\mathcal{F}_t)$ shares of D, and B wants to buy φ shares of D. Assume that both parties are using acceptability pricing theory. Namely, party A will use the family of drivers g^1, and level γ_1, to calculate his ask price, and party B will use g^2, and level γ_2, to calculate her bid price. Clearly the trade will happen only if B's bid price meets A's ask price. Note that Proposition 2.3.2 guarantees only that $a_t^{g^1,\gamma_1}(\varphi, D) \geq b_t^{g^2,\gamma_2}(\varphi, D)$, and hence, it is important to investigate under which conditions $a_t^{g^1,\gamma_1}(\varphi, D) = b_t^{g^2,\gamma_2}(\varphi, D)$. Not to our surprise, this question has a close connection to linear pricing theory. As shown in the next result, in order for bid and ask prices to coincide, the drivers (and hence the prices) have to be locally linear.

Proposition 2.3.4. *Let* $A \in \mathcal{F}_t$, g^1 *and* g^2 *be two families of drivers,* γ_1, $\gamma_2 > 0$. *Also, fix* $t \in \mathcal{T}$, $D \in \mathcal{D}$. *Then,*

$$\mathbb{1}_A a_t^{g^1,\gamma_1}(\varphi, D) = \mathbb{1}_A b_t^{g^2,\gamma_2}(\varphi, D)$$

if and only if there exists a driver $\tilde{g}_{t,D,A,\varphi}$ *such that* $\tilde{g}_{t,D,A,\varphi}(s,\cdot)$, $s = t+1,\dots,T$, *is linear, and*

$$\begin{aligned} \mathbb{1}_A \mathcal{E}_{g^1_{\gamma_1}}\Big[\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] &= \mathbb{1}_A \mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big]; \\ \mathbb{1}_A \mathcal{E}_{g^2_{\gamma_2}}\Big[-\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] &= \mathbb{1}_A \mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[-\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big], \end{aligned} \tag{2.30}$$

for any $0 \leq \lambda \leq \varphi$.

Proof. ($\Longleftarrow$) If there exists a driver $\tilde{g}_{t,D,A,\varphi}$ such that $\tilde{g}_{t,D,A,\varphi}(s,\cdot)$, $s = t+1,\dots,T$, is linear, and

$$\mathbb{1}_A \mathcal{E}_{g^1_{\gamma_1}}\Big[\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] = \mathbb{1}_A \mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big],$$

$$\mathbb{1}_A \mathcal{E}_{g^2_{\gamma_2}}\Big[-\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] = \mathbb{1}_A \mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[-\lambda \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big],$$

for $0 \leq \lambda \leq \varphi$, then, by taking $\lambda = \varphi$, we have

$$\mathbb{1}_A a_t^{g^1,\gamma_1}(\varphi, D) = \mathbb{1}_A \mathcal{E}_{g^1_{\gamma_1}}\Big[\varphi \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] = \mathbb{1}_A \mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[\varphi \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big];$$

$$\mathbb{1}_A b_t^{g^2,\gamma_2}(\varphi, D) = -\mathbb{1}_A \mathcal{E}_{g^2_{\gamma_2}}\Big[-\varphi \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big] = \mathbb{1}_A \mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[\varphi \sum_{s=t+1}^{T} D_s \Big| \mathcal{F}_t\Big],$$

and hence, $\mathbb{1}_A a_t^{g^1,\gamma_1}(\varphi, D) = \mathbb{1}_A b_t^{g^2,\gamma_2}(\varphi, D)$.

($\Longrightarrow$) Assume that $\mathbb{1}_A a_t^{g^1,\gamma_1}(\varphi, D) = \mathbb{1}_A b_t^{g^2,\gamma_2}(\varphi, D)$. Then, for $0 \leq \lambda \leq \varphi$, there exists $0 \leq \lambda' \leq 1$ such that $\lambda = \lambda'\varphi$. By Theorem 2.3.1.P4, we get

$$\mathbb{1}_A a_t^{g^1,\gamma_1}(\lambda, D) \leq \mathbb{1}_A \lambda' a_t^{g^1,\gamma_1}(\varphi, D) = \mathbb{1}_A \lambda' b_t^{g^2,\gamma_2}(\varphi, D) \leq \mathbb{1}_A b_t^{g^2,\gamma_2}(\lambda, D),$$

however, in view of Proposition 2.3.2, we have that $\mathbb{1}_A a_t^{g^1,\gamma_1}(\lambda, D) \geq \mathbb{1}_A b_t^{g^2,\gamma_2}(\lambda, D)$. Therefore,

$$\mathbb{1}_A a_t^{g^1,\gamma_1}(\lambda, D) = \mathbb{1}_A \lambda' a_t^{g^1,\gamma_1}(\varphi, D) = \mathbb{1}_A \lambda' b_t^{g^2,\gamma_2}(\varphi, D) = \mathbb{1}_A b_t^{g^2,\gamma_2}(\lambda, D), \tag{2.31}$$

Let $(\mathcal{E}_{g^1_{\gamma_1}}[\mathbb{1}_A \varphi \sum_{u=t+1}^{T} D_u | \mathcal{F}_s], \widetilde{Z}_s, \widetilde{M}_s)$, $t \leq s \leq T$, be the solution of BSΔE corresponding to driver $g^1_{\gamma_1}$ and terminal condition $\mathbb{1}_A \varphi \sum_{s=t+1}^{T} D_s$. Let us define $x_s^{t,D,A,\varphi}$ for $t+1 \leq s \leq T$ as

$$x_s^{t,D,A,\varphi} = \begin{cases} \frac{g^1_{\gamma_1}(s,\widetilde{Z}_s)}{\widetilde{Z}_s} & \text{if } \widetilde{Z}_s \neq 0 \\ 0 & \text{if } \widetilde{Z}_s = 0. \end{cases} \tag{2.32}$$

Next, we define $\tilde{g}_{t,D,A,\varphi}(s, z) = x_s^{t,D,A,\varphi} z$ for $t+1 \leq s \leq T$, and $z \in \mathbb{R}$. We will show that $\tilde{g}_{t,D,A,\varphi}$ is the desired driver.

First, let us show that $\tilde{g}_{t,D,A,\varphi}$ satisfies Assumption A. First note that $x_s^{t,D,A,\varphi}$ defined in (2.32) is $\mathcal{F}_{s-1}$-measurable, and thus $\tilde{g}_{t,D,A,\varphi}(s, z)$ is $\mathcal{F}_{s-1}$-measurable for any $z \in \mathbb{R}$, and so it satisfies A1. Since $g^1_{\gamma_1}$ satisfies assumption A2, then $|x_s^{t,D,A,\varphi}| =$

$\frac{|g^1_{\gamma_1}(s,\widetilde{Z}_s)|}{|\widetilde{Z}_s|} \leq c^1_{\gamma_1}(s)$ on the set $\{\widetilde{Z}_s \neq 0\}$, where $c^1_{\gamma_1}(s)$ is the Lipschitz coefficient of $g^1_{\gamma_1}$, for $s \in \{t+1,\ldots,T\}$. Of course, $|x^{t,D,A,\varphi}_s| = 0 \leq c^1_{\gamma_1}(s)$ on the complement of $\{\widetilde{Z}_s \neq 0\}$, for $s \in \{t+1,\ldots,T\}$. Thus, $\tilde{g}_{t,D,A,\varphi}$ satisfies A2. Clearly, $\tilde{g}$ satisfies A3, and thus it satisfies Assumption A.

Next, we will show that the identities (2.30) are fulfilled. By the construction of $\tilde{g}_{t,D,A,\varphi}$, we have that $\mathbb{1}_A\mathcal{E}_{g^1_{\gamma_1}}[\varphi\sum_{s=t+1}^T D_s|\mathcal{F}_t] = \mathbb{1}_A\mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}[\varphi\sum_{s=t+1}^T D_s|\mathcal{F}_t]$, and thus, for $0 \leq \lambda \leq \varphi$, with $\lambda = \lambda'\varphi$, we get

$$\mathbb{1}_A\mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[\lambda\sum_{s=t+1}^T D_s\Big|\mathcal{F}_t\Big] = \mathbb{1}_A\lambda'\mathcal{E}_{\tilde{g}_{t,D,A,\varphi}}\Big[\varphi\sum_{s=t+1}^T D_s\Big|\mathcal{F}_t\Big] = \mathbb{1}_A\lambda'\mathcal{E}_{g^1_{\gamma_1}}\Big[\varphi\sum_{s=t+1}^T D_s\Big|\mathcal{F}_t\Big]$$
$$= \mathbb{1}_A\mathcal{E}_{g^1_{\gamma_1}}\Big[\lambda\sum_{s=t+1}^T D_s\Big|\mathcal{F}_t\Big],$$

where the last equality holds because of (2.31). Second identity in (2.30) is proved similarly.

This concludes the proof.

□

Remark 2.3.8. *Proposition 2.3.4 implies that* $a^{g^1,\gamma_1}_t(\varphi,D) = b^{g^2,\gamma_2}_t(\varphi,D)$ *if and only if* $a^{g^1,\gamma_1}_t(\lambda,D) = b^{g^2,\gamma_2}_t(\lambda,D)$*, for any* $0 \leq \lambda \leq \varphi$*. In other words, if two counterparties agree on the prices for* φ *shares, then they will also agree on prices for any smaller (positive) number of shares* $\lambda\varphi$*.*

To conclude this section, we show that if $a^{g^1,\gamma_1}_t(D) = b^{g^2,\gamma_2}_t(D)$ for all $D \in \mathcal{D}$ and $t \in \mathcal{T}$, then $g^1_{\gamma_1}$ and $g^2_{\gamma_2}$ have to be equal and linear. This is one reason why the results in Proposition 2.3.4 hold true only locally.

Proposition 2.3.5. *Let* g^1 *and* g^2 *be two families of drivers, and* $\gamma_1,\gamma_2 > 0$*.*

(i) Assume that $a^{g^1,\gamma_1}_t(D) = b^{g^2,\gamma_2}_t(D)$ *for any* $D \in \mathcal{D}$*, and for a fixed* $t \in \mathcal{T}$*. Then,*

$\mathcal{E}_{g^1_{\gamma_1}}[\ \cdot\ |\mathcal{F}_t]=\mathcal{E}_{g^2_{\gamma_2}}[\ \cdot\ |\mathcal{F}_t]$. *Moreover, in this case the functional* $\mathcal{E}_{g^1_{\gamma_1}}[\ \cdot\ |\mathcal{F}_t]$ *is linear.*

(ii) Assume that $a_t^{g^1,\gamma_1}(D) = b_t^{g^2,\gamma_2}(D)$ *for any* $D \in \mathcal{D}$, *and any* $t \in \mathcal{T}$. *Then, there exists a driver* $\widetilde{g}(t,z)$ *such that* $\widetilde{g}(t,\cdot)$ *is linear, and* $g^1_{\gamma_1}(t,z) = g^2_{\gamma_2}(t,z) = \widetilde{g}(t,z)$ *for any* $t \in \mathcal{T}$, $z \in \mathbb{R}$.

Proof. Due to the assumption and the representations of bid/ask prices, we have

$$\mathcal{E}_{g^1_{\gamma_1}}\Big[\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big] = -\mathcal{E}_{g^2_{\gamma_2}}\Big[-\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big], \quad D \in \mathcal{D}. \tag{2.33}$$

Clearly, (2.33) is also true for $-D$ and therefore

$$\begin{aligned}\mathcal{E}_{g^2_{\gamma_2}}\Big[\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big] =& -\mathcal{E}_{g^1_{\gamma_1}}\Big[-\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big]\\ &\leq \mathcal{E}_{g^1_{\gamma_1}}\Big[\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big] = -\mathcal{E}_{g^2_{\gamma_2}}\Big[-\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big].\end{aligned}$$

Since $-\mathcal{E}_{g^2_{\gamma_2}}[-\sum_{s=t+1}^{T} D_s|\mathcal{F}_t] \leq \mathcal{E}_{g^2_{\gamma_2}}[\sum_{s=t+1}^{T} D_s|\mathcal{F}_t]$, then

$$\begin{aligned}\mathcal{E}_{g^2_{\gamma_2}}\Big[\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big] =& -\mathcal{E}_{g^1_{\gamma_1}}\Big[-\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big]\\ &= \mathcal{E}_{g^1_{\gamma_1}}\Big[\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big] = -\mathcal{E}_{g^2_{\gamma_2}}\Big[-\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big]\end{aligned} \tag{2.34}$$

for any $D \in \mathcal{D}$.

In view of (2.34), we have that $a_t^{g^i,\gamma_i}(\varphi, D) = a_t^{g^i,\gamma_i}(1, \varphi D) = b_t^{g^i,\gamma_i}(1, \varphi D) = b_t^{g^i,\gamma_i}(\varphi, D)$, $i = 1, 2$, and thus, by Proposition 2.3.4, we obtain that

$$\mathcal{E}_{g^i_{\gamma_i}}\Big[\varphi \sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big] = \varphi\mathcal{E}_{g^i_{\gamma_i}}\Big[\sum_{s=t+1}^{T} D_s\Big|\mathcal{F}_t\Big], \quad i = 1, 2, \tag{2.35}$$

for any $\varphi \in L^\infty_+(\mathcal{F}_t)$, $D \in \mathcal{D}$. Moreover, by (2.34) again, equation (2.35) is also true for $\varphi \in L^\infty_-(\mathcal{F}_t)$. Hence, (2.35) is true for any $\varphi \in L^\infty(\mathcal{F}_t)$.

To proceed, let $D^1, D^2 \in \mathcal{D}$, $a, b \in L^\infty(\mathcal{F}_t)$. Then by convexity of g-expectation and (2.35), we have that

$$\begin{aligned}\mathcal{E}_{g^i_{\gamma_i}}\Big[a\sum_{s=t+1}^T D^1_s + b\sum_{s=t+1}^T D^2_s\Big|\mathcal{F}_t\Big] &\leq \frac{1}{2}\mathcal{E}_{g^i_{\gamma_i}}\Big[2a\sum_{s=t+1}^T D^1_s\Big|\mathcal{F}_t\Big] + \frac{1}{2}\mathcal{E}_{g^i_{\gamma_i}}\Big[2b\sum_{s=t+1}^T D^2_s\Big|\mathcal{F}_t\Big] \\ &= a\mathcal{E}_{g^i_{\gamma_i}}\Big[\sum_{s=t+1}^T D^1_s\Big|\mathcal{F}_t\Big] + b\mathcal{E}_{g^i_{\gamma_i}}\Big[\sum_{s=t+1}^T D^2_s\Big|\mathcal{F}_t\Big].\end{aligned}$$

Due to (2.34), it is also true that

$$\mathcal{E}_{g^i_{\gamma_i}}\Big[a\sum_{s=t+1}^T D^1_s + b\sum_{s=t+1}^T D^2_s\Big|\mathcal{F}_t\Big] \geq a\mathcal{E}_{g^i_{\gamma_i}}\Big[\sum_{s=t+1}^T D^1_s\Big|\mathcal{F}_t\Big] + b\mathcal{E}_{g^i_{\gamma_i}}\Big[\sum_{s=t+1}^T D^2_s\Big|\mathcal{F}_t\Big],$$

and consequently, in view of the above, we deduce that

$$\mathcal{E}_{g^i_{\gamma_i}}\Big[a\sum_{s=t+1}^T D^1_s + b\sum_{s=t+1}^T D^2_s\Big|\mathcal{F}_t\Big] = a\mathcal{E}_{g^i_{\gamma_i}}\Big[\sum_{s=t+1}^T D^1_s\Big|\mathcal{F}_t\Big] + b\mathcal{E}_{g^i_{\gamma_i}}\Big[\sum_{s=t+1}^T D^2_s\Big|\mathcal{F}_t\Big]. \quad (2.36)$$

Thus, (2.34), (2.35) and (2.36) imply that $\mathcal{E}_{g^1_{\gamma_1}}[\,\cdot\,|\mathcal{F}_t] = \mathcal{E}_{g^2_{\gamma_2}}[\,\cdot\,|\mathcal{F}_t]$, and that they are linear.

If $a^{g^1,\gamma_1}_t(D) = b^{g^2,\gamma_2}_t(D)$ for any $D \in \mathcal{D}$, $t \in \mathcal{T}$, then $\mathcal{E}_{g^1_{\gamma_1}}[\,\cdot\,|\mathcal{F}_t] = \mathcal{E}_{g^2_{\gamma_2}}[\,\cdot\,|\mathcal{F}_t]$ and they are linear for any $t \in \mathcal{T}$. By Proposition 2.1.3, there exists a linear driver $\widetilde{g}(t,z)$ such that $g^1_{\gamma_1}(t,z) = g^2_{\gamma_2}(t,z) = \widetilde{g}(t,z)$, $t \in \mathcal{T}$, $z \in \mathbb{R}$.

This concludes the proof. □

2.3.4 Market Models. In this section, we will consider some market models that follow the set-up introduced in Section 2.3.1. Recall that a market is denoted by a triple $(\mathcal{M}, P^{\text{ask}}, P^{\text{bid}})$, where $\mathcal{M}$ is the subspace of $\mathcal{D}$ that consists of processes $D^{\text{ask/bid},i}$, $i = 1, \ldots, K$, as in (M1), and of $D^0 = (0, \ldots, 0, 1)$, which is the dividend process of the banking account. The functionals P^{ask} and P^{bid} allow to compute the ex-dividend prices of the cash flow $\widetilde{D} \in \mathcal{M}$. We will define P^{ask} and P^{bid} by using pricing operators introduced in Section 2.3.3.

Ask and Bid Prices Computed at the Same Acceptability Level

Let g be a family of drivers that satisfies Assumption G, and let $\gamma > 0$. We consider the market model $(\mathcal{M}, a^{g,\gamma}, b^{g,\gamma})$. Namely, we put

$$P_t^{\text{ask}}(\varphi, \widetilde{D}) = a_t^{g,\gamma}(\varphi, \widetilde{D}), \quad P_t^{\text{bid}}(\varphi, \widetilde{D}) = b_t^{g,\gamma}(\varphi, \widetilde{D}),$$

for any cash flow $\widetilde{D} \in \mathcal{M}$, and $\varphi \in L_+^\infty(\mathcal{F}_t)$.

Since $a^{g,\gamma}, b^{g,\gamma} : L_+^\infty(\mathcal{F}_t) \times \mathcal{D} \to L^2(\mathcal{F}_t)$, for any $t \in \mathcal{T}$, then such market is well-defined. Moreover, due to Theorem 2.3.1, we also have the following representation

$$P_t^{\text{ask}}(\varphi, \widetilde{D}) = \mathcal{E}_{g_\gamma}\Big[\varphi \sum_{s=t+1}^{T} \widetilde{D}_s \Big| \mathcal{F}_t\Big],$$

$$P_t^{\text{bid}}(\varphi, \widetilde{D}) = -\mathcal{E}_{g_\gamma}\Big[-\varphi \sum_{s=t+1}^{T} \widetilde{D}_s \Big| \mathcal{F}_t\Big].$$

According to Proposition 2.1.1, we have that $P_t^{\text{ask}}(0, \widetilde{D}) = \mathcal{E}_{g_\gamma}[0|\mathcal{F}_t] = 0$, and $P_t^{\text{bid}}(0, \widetilde{D}) = -\mathcal{E}_{g_\gamma}[0|\mathcal{F}_t] = 0$. Hence, the market satisfies assumption (M2). In particular, for $D^0 = (0, \dots, 0, 1)$, it is clear that

$$P_t^{\text{ask}}(\varphi, D^0) = a_t^{g,\gamma}(\varphi, D^0) = \mathcal{E}_{g_\gamma}[\varphi|\mathcal{F}_t] = \varphi,$$

$$P_t^{\text{bid}}(\varphi, D^0) = b_t^{g,\gamma}(\varphi, D^0) = -\mathcal{E}_{g_\gamma}[-\varphi|\mathcal{F}_t] = \varphi,$$

for any $\varphi \in L_+^\infty(\mathcal{F}_t)$, which implies that the market satisfies (M3).

Next, we will show that our market satisfies the following important properties, proved over the course of two theorems:

(M4) The market is arbitrage-free.

(M5) For any $\widetilde{D} \in \mathcal{M}$, $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \le \lambda \le 1$, $t \in \mathcal{T}$,

$$P_t^{\text{ask}}(\lambda\varphi^1 + (1-\lambda)\varphi^2, \widetilde{D}) \le \lambda P_t^{\text{ask}}(\varphi^1, \widetilde{D}) + (1-\lambda) P_t^{\text{ask}}(\varphi^2, \widetilde{D}),$$

$$P_t^{\text{bid}}(\lambda\varphi^1 + (1-\lambda)\varphi^2, \widetilde{D}) \ge \lambda P_t^{\text{bid}}(\varphi^1, \widetilde{D}) + (1-\lambda) P_t^{\text{bid}}(\varphi^2, \widetilde{D}).$$

(M6) For any $\widetilde{D} \in \mathcal{M}$, $\varphi^1, \lambda \in L^\infty_+(\mathcal{F}_t)$, $\varphi^2 \in L^\infty_-(\mathcal{F}_t)$, $0 \leq \lambda \leq 1$, $i = 1, \ldots, K$, and $t \in \mathcal{T}$,

$$\lambda P_t^{\text{ask}}(\varphi^1, \widetilde{D}) - (1-\lambda) P_t^{\text{bid}}(-\varphi^2, \widetilde{D}) \geq \mathbb{1}_{\vartheta \geq 0} P_t^{\text{ask}}(\vartheta, \widetilde{D}) - \mathbb{1}_{\vartheta < 0} P_t^{\text{bid}}(-\vartheta, \widetilde{D}),$$

where $\vartheta = \lambda \varphi^1 + (1-\lambda)\varphi^2$.

Properties (M5) and (M6) imply that diversification in the trading is favored. If an investor nets his purchasing and selling in a convex way, then the cost of trading will be reduced.

Now we proceed by showing that such market is arbitrage free.

Theorem 2.3.2. *The market model $(\mathcal{M}, a^{g,\gamma}, b^{g,\gamma})$, $\gamma > 0$, is arbitrage free.*

Proof. Assume, the market admits an arbitrage, so that, according to Proposition 2.3.1, there is a trading strategy $\phi \in \mathcal{S}(0,\gamma) := \mathcal{S}(0, a^{g,\gamma}, b^{g,\gamma})$, such that $V_T(\phi) \geq 0$ and $\mathbb{P}(V_T(\phi) > 0) > 0$. We will show that this leads to a contradiction, by showing that for a self-financing portfolio ϕ inequality (2.40) below is satisfied, leading to a contradictory inequality that $\mathbb{P}(0 > 0) > 0$.

Note that

$$\begin{aligned} V_T(\phi) =& \phi_T^0 + \sum_{i=1}^K \left(\phi_T^{l,i} D_T^{\text{ask},i} - \phi_T^{s,i} D_T^{\text{bid},i} \right) \\ =& \phi_1^0 + \sum_{s=2}^T \Delta \phi_s^0 + \sum_{i=1}^K \left(\phi_T^{l,i} D_T^{\text{ask},i} - \phi_T^{s,i} D_T^{\text{bid},i} \right). \end{aligned}$$

Since ϕ is self-financing, then we have

$$\begin{aligned} \Delta \phi_u^0 =& \sum_{i=1}^K \left(\phi_{u-1}^{l,i} D_{u-1}^{\text{ask},i} - \phi_{u-1}^{s,i} D_{u-1}^{\text{bid},i} \right) - \sum_{i=1}^K \Big(\mathbb{1}_{\Delta \phi_u^{l,i} \geq 0} a_{u-1}^{g,\gamma}(\Delta \phi_u^{l,i}, D^{\text{ask},i}) \\ & - \mathbb{1}_{\Delta \phi_u^{l,i} < 0} b_{u-1}^{g,\gamma}(-\Delta \phi_u^{l,i}, D^{\text{ask},i}) - \mathbb{1}_{\Delta \phi_u^{s,i} \geq 0} b_{u-1}^{g,\gamma}(\Delta \phi_u^{s,i}, D^{\text{bid},i}) \\ & + \mathbb{1}_{\Delta \phi_u^{s,i} < 0} a_{u-1}^{g,\gamma}(\Delta \phi_u^{s,i}, D^{\text{bid},i}) \Big), \qquad u = 2, \ldots, T. \end{aligned}$$

For convenience, we use the following notations:

$$\begin{aligned}
&\xi_t^0 = \phi_t^0, \quad t = 1, \dots, T, \\
&\xi_t^i = \phi_t^{l,i}, \quad i = 1, \dots, K, \ t = 1, \dots, T, \\
&\xi_t^j = -\phi_t^{s,j-K}, \quad j = K+1, \dots, 2K, \ t = 1, \dots, T.
\end{aligned} \tag{2.37}$$

Also, let us define $\widehat{D}$ as

$$\begin{aligned}
&\widehat{D}^0 = D^0, \\
&\widehat{D}^i = D^{\text{ask},i}, \quad i = 1, \dots, K, \\
&\widehat{D}^j = D^{\text{bid},j-K}, \quad j = K+1, \dots, 2K.
\end{aligned}$$

Then, we have that

$$\begin{aligned}
\Delta\xi_t^0 =& \sum_{i=1}^{2K} \xi_{t-1}^i \widehat{D}_{t-1}^i - \sum_{i=1}^{2K} \mathcal{E}_{g_\gamma}\left[\Delta\xi_t^i \sum_{s=t}^{T} \widehat{D}_s^i \middle| \mathcal{F}_{t-1}\right], \quad t = 2, \dots, T, \\
V_T(\phi) =& \xi_T^0 + \sum_{i=1}^{2K} \xi_T^i \widehat{D}_T^i = \xi_1^0 + \sum_{s=2}^{T} \Delta\xi_s^0 + \sum_{i=1}^{2K} \xi_T^i \widehat{D}_T^i \\
=& \xi_1^0 + \sum_{s=2}^{T} \left(\sum_{i=1}^{2K} \xi_{s-1}^i \widehat{D}_{s-1}^i - \sum_{i=1}^{2K} \mathcal{E}_{g_\gamma}\left[\Delta\xi_s^i \sum_{u=s}^{T} \widehat{D}_u^i \middle| \mathcal{F}_{s-1}\right]\right) + \sum_{i=1}^{2K} \xi_T^i \widehat{D}_T^i \\
=& \xi_1^0 + \sum_{s=2}^{T} \left(\sum_{i=1}^{2K} \xi_{s-1}^i \widehat{D}_{s-1}^i - \sum_{i=1}^{2K} \mathcal{E}_{g_\gamma}\left[\Delta\xi_s^i \sum_{u=s}^{T} \widehat{D}_u^i \middle| \mathcal{F}_{s-1}\right]\right) \\
&+ \sum_{i=1}^{2K} (\xi_1^i + \sum_{s=2}^{T} \Delta\xi_s^i) \widehat{D}_T^i.
\end{aligned}$$

By multiplying both sides by $\frac{1}{2KT}$, and applying the conditional g-expectation

$\mathcal{E}_{g_\gamma}[\,\cdot\,|\mathcal{F}_{T-1}]$ to both sides, in view of Proposition 2.1.1.(v), we deduce

$$\begin{aligned}\mathcal{E}_{g_\gamma}\Big[\frac{1}{2KT}V_T(\phi)\Big|\mathcal{F}_{T-1}\Big] =& \frac{1}{2TK}\xi_1^0 + \frac{1}{2TK}\sum_{s=2}^{T}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i \qquad (2.38)\\ &- \frac{1}{2TK}\sum_{s=2}^{T}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big] + \mathcal{E}_{g_\gamma}\Big[\frac{1}{2KT}\sum_{i=1}^{2K}(\xi_1^i+\sum_{s=2}^{T}\Delta\xi_s^i)\widehat{D}_T^i\Big|\mathcal{F}_{T-1}\Big]\\ \le& \frac{1}{2KT}\xi_1^0 + \frac{1}{2KT}\sum_{s=2}^{T}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i - \frac{1}{2KT}\sum_{s=2}^{T}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\\ &+ \frac{1}{2KT}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\widehat{D}_T^i\Big|\mathcal{F}_{T-1}\Big] + \frac{1}{2KT}\sum_{s=2}^{T}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\widehat{D}_T^i\Big|\mathcal{F}_{T-1}\Big]\\ =& \frac{1}{2KT}\xi_1^0 + \frac{1}{2KT}\sum_{s=2}^{T}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i - \frac{1}{2KT}\sum_{s=2}^{T-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\\ &+ \frac{1}{2KT}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\widehat{D}_T^i\Big|\mathcal{F}_{T-1}\Big] + \frac{1}{2KT}\sum_{s=2}^{T-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\widehat{D}_T^i\Big|\mathcal{F}_{T-1}\Big].\end{aligned}$$

Since $V_T(\phi)\ge 0$, $\mathbb{P}(V_T(\phi)>0)>0$, then by Proposition 2.1.1.(ii) applied to the left hand side of (2.38), we have that

$$\mathcal{E}_{g_\gamma}\Big[\frac{1}{2KT}V_T(\phi)\Big|\mathcal{F}_{T-1}\Big]\ge 0, \quad \text{and} \quad \mathbb{P}\left(\mathcal{E}_{g_\gamma}\Big[\frac{1}{2KT}V_T(\phi)\Big|\mathcal{F}_{T-1}\Big]>0\right)>0.$$

Consequently,

$$\begin{aligned}&\xi_1^0 + \sum_{s=2}^{T-1}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i - \sum_{s=2}^{T-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\\ &+\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{u=T-1}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{T-1}\Big] + \sum_{s=2}^{T-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=T-1}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{T-1}\Big]\ge 0,\end{aligned}$$

and the strict inequality holds on some set with positive probability. Let us use the notation

$$\begin{aligned}\Pi_t :=& \xi_1^0 + \sum_{s=2}^{t}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i - \sum_{s=2}^{t}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\\ &+\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{u=t}^{T}\widehat{D}_s^i\Big|\mathcal{F}_t\Big] + \sum_{s=2}^{t}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=t}^{T}\widehat{D}_u^i\Big|\mathcal{F}_t\Big], \qquad (2.39)\end{aligned}$$

where $t\in\{1,\dots,T-1\}$. We just showed that $\Pi_{T-1}\ge 0$ and $\mathbb{P}(\Pi_{T-1}>0)>0$. Next, we are going to prove that $\Pi_t\ge 0$ and $\mathbb{P}(\Pi_t>0)>0$ will imply that $\Pi_{t-1}\ge 0$, $\mathbb{P}(\Pi_{t-1}>0)>0$, for any $t\in\{2,\dots,T-1\}$.

Using Proposition 2.1.1.(iii)-(iv), we obtain that

$$\begin{aligned}
\mathcal{E}_{g_\gamma}\Big[\frac{1}{2Kt}\Pi_t\Big|\mathcal{F}_{t-1}\Big] = &\frac{1}{2Kt}\Big(\xi_1^0+\sum_{s=2}^{t}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i-\sum_{s=2}^{t}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\Big)\\
&+\mathcal{E}_{g_\gamma}\Big[\frac{1}{2Kt}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{u=t}^{T}\widehat{D}_u^i\Big|\mathcal{F}_t\Big]+\frac{1}{2Kt}\sum_{s=2}^{t}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=t}^{T}\widehat{D}_u^i\Big|\mathcal{F}_t\Big]\Big|\mathcal{F}_{t-1}\Big]\\
\le&\frac{1}{2Kt}\Big(\xi_1^0+\sum_{s=2}^{t}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i-\sum_{s=2}^{t}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\Big)\\
&+\frac{1}{2Kt}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{u=t}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{t-1}\Big]+\frac{1}{2Kt}\sum_{s=2}^{t}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=t}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{t-1}\Big]\\
=&\frac{1}{2Kt}\Big(\xi_1^0+\sum_{s=2}^{t}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i-\sum_{s=2}^{t-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\Big)\\
&+\frac{1}{2Kt}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{u=t}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{t-1}\Big]+\frac{1}{2Kt}\sum_{s=2}^{t-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=t}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{t-1}\Big]\\
=&\frac{1}{2Kt}\Big(\xi_1^0+\sum_{s=2}^{t-1}\sum_{i=1}^{2K}\xi_{s-1}^i\widehat{D}_{s-1}^i-\sum_{s=2}^{t-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=s}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{s-1}\Big]\Big)\\
&+\frac{1}{2Kt}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{u=t-1}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{t-1}\Big]+\frac{1}{2Kt}\sum_{s=2}^{t-1}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\Delta\xi_s^i\sum_{u=t-1}^{T}\widehat{D}_u^i\Big|\mathcal{F}_{t-1}\Big]\\
=&\frac{1}{2Kt}\Pi_{t-1},
\end{aligned}$$

for $t\in\{2,\dots,T-1\}$. Since $\Pi_t\ge 0$ and $\mathbb{P}(\Pi_t>0)>0$, then according to Proposition 2.1.1.(ii), we get that $\mathcal{E}_{g_\gamma}\Big[\frac{1}{2Kt}\Pi_t\Big|\mathcal{F}_{t-1}\Big]$ and $\mathbb{P}(\mathcal{E}_{g_\gamma}\Big[\frac{1}{2Kt}\Pi_t\Big|\mathcal{F}_{t-1}\Big]>0)>0$, and thus $\Pi_{t-1}\ge 0$, $\mathbb{P}(\Pi_{t-1}>0)>0$. Hence, by backward induction, it is true that

$$\Pi_1=\xi_1^0+\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{s=1}^{T}\widehat{D}_s^i\Big|\mathcal{F}_1\Big]\ge 0,$$

and $\mathbb{P}(\Pi_1>0)>0$. Consequently, in view of (2.37) and the representation for $a^{g,\gamma}$ and $b^{g,\gamma}$ (cf. Theorem 2.3.1) we obtain that

$$\begin{aligned}
0\le&\frac{1}{2K}\xi_1^0+\frac{1}{2K}\sum_{i=1}^{2K}\mathcal{E}_{g_\gamma}\Big[\xi_1^i\sum_{s=1}^{T}\widehat{D}_s^i\Big]\\
=&\frac{1}{2K}\phi_1^0+\frac{1}{2K}\sum_{i=1}^{K}\Big(a_0^{g,\gamma}(\phi_1^{l,i},D^{\mathrm{ask},i})-b_0^{g,\gamma}(-\phi_1^{s,i},D^{\mathrm{bid},i})\Big),
\end{aligned}$$

and

$$\mathbb{P}\Big(\phi_1^0 + \sum_{i=1}^{K} \Big(a_0^{g,\gamma}(\phi_1^{l,i}, D^{\text{ask},i}) - b_0^{g,\gamma}(-\phi_1^{s,i}, D^{\text{bid},i})\Big) > 0\Big) > 0. \tag{2.40}$$

Since ϕ is a self-financing, then

$$\xi_1^0 = -\sum_{i=1}^{K} \Big(a_0^{g,\gamma}(\phi_1^{l,i}, D^{\text{ask},i}) - b_0^{g,\gamma}(-\phi_1^{s,i}, D^{\text{bid},i})\Big).$$

Thus, (2.40) means that $\mathbb{P}(0 > 0) > 0$, which is a contradiction. □

Proposition 2.3.6. *Assume that* $(\mathcal{M}, P^{ask}, P^{bid}) = (\mathcal{M}, a^{g,\gamma}, b^{g,\gamma})$, $\gamma > 0$. *Then, properties (M5) and (M6) hold true.*

Proof. Due to Theorem 2.3.1.P3, we have that

$$P_t^{\text{ask}}(\lambda\varphi^1 + (1-\lambda)\varphi^2, \widetilde{D}) \le \lambda P_t^{\text{ask}}(\varphi^1, \widetilde{D}) + (1-\lambda)P_t^{\text{ask}}(\varphi^2, \widetilde{D}),$$

and

$$P_t^{\text{bid}}(\lambda\varphi^1 + (1-\lambda)\varphi^2, \widetilde{D}) \ge \lambda P_t^{\text{bid}}(\varphi^1, \widetilde{D}) + (1-\lambda)P_t^{\text{bid}}(\varphi^2, \widetilde{D}),$$

for any $\widetilde{D} \in \mathcal{M}$, $\lambda \in L^\infty(\mathcal{F}_t)$, $0 \le \lambda \le 1$, $t \in \mathcal{T}$, which implies that condition (M5) is satisfied.

We are left to show that (M6) holds. In view of Theorem 2.3.1.P1, we have that

$$\begin{aligned}
&\lambda P_t^{\text{ask}}(\varphi^1, \widetilde{D}) - (1-\lambda)P_t^{\text{bid}}(-\varphi^2, \widetilde{D}) \\
&\quad = \lambda \mathcal{E}_{g_\gamma}\Big[\varphi^1 \sum_{s=t+1}^{T} \widetilde{D}_s \Big| \mathcal{F}_t\Big] + (1-\lambda)\mathcal{E}_{g_\gamma}\Big[\varphi^2 \sum_{s=t+1}^{T} \widetilde{D}_s \Big| \mathcal{F}_t\Big] \\
&\quad \ge \mathcal{E}_{g_\gamma}\Big[\vartheta \sum_{s=t+1}^{T} \widetilde{D}_s \Big| \mathcal{F}_t\Big] = \mathbb{1}_{\vartheta \ge 0} P_t^{\text{ask}}(\vartheta, \widetilde{D}) - \mathbb{1}_{\vartheta < 0} P_t^{\text{bid}}(-\vartheta, \widetilde{D}).
\end{aligned}$$

This concludes the proof. □

In Theorem 2.3.2, we showed that if we take a family of drivers $g = (g_x)_{x>0}$, and if we choose the same acceptability level $\gamma > 0$ to define the ask and bid prices, then the market model using such prices is arbitrage free. In next section, using the results from the current section, we will prove that market model is still arbitrage free even we choose different acceptability level for the two trading sides.

Ask and Bid Prices Computed at Different Acceptability Levels

Let g be a family of drivers that satisfies Assumption G. Also, let $\gamma_1, \gamma_2 > 0$. We consider the market model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_2})$. That is $P_t^{\text{ask}}(\varphi, \widetilde{D}) = a_t^{g,\gamma_1}(\varphi, \widetilde{D})$ and $P_t^{\text{bid}}(\varphi, \widetilde{D}) = b_t^{g,\gamma_2}(\varphi, \widetilde{D})$, for $\widetilde{D} \in \mathcal{M}$, $\varphi \in L_+^\infty(\mathcal{F}_t)$.

Similarly as in Section 2.3.4, it is not hard to verify here that the market model of the present section satisfies properties (M2) and (M3), and we leave the verification of these properties to the reader. At this time we are unable to verify that property (M6) holds for this market. However, we can verify that property (M5) holds.

Theorem 2.3.3. *The market model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_2})$, $\gamma_1, \gamma_2 > 0$, is arbitrage free.*

Proof. First, we consider the case $\gamma_1 \leq \gamma_2$. We will prove the result by contradiction. Namely, we will assume that market model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_2})$ admits an arbitrage, and we will conclude that market model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_1})$ also admits an arbitrage, which is impossible in view of Theorem 2.3.2. Intuitively, the statement and its proof is clear: since $b^{g,\gamma_1} \geq b^{g,\gamma_2}$, one will trade at higher bid prices in the new market $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_1})$, and hence it is enough to trade in this market all assets but banking account according to the arbitrage strategy from $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_2})$. Finally, the banking account is set up such that the trading strategy remains self-financing in $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_1})$.

Let us assume that the market model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_2})$ admits an arbitrage opportunity, which, according to Proposition 2.3.1, means that there is a trading strategy $\phi \in \mathcal{S}(0, a^{g,\gamma_1}, b^{g,\gamma_2})$, such that $V_T(\phi) \geq 0$ and $\mathbb{P}(V_T(\phi) > 0) > 0$. Using ϕ, we will construct an arbitrage strategy $\xi \in \mathcal{S}(0, a^{g,\gamma_1}, b^{g,\gamma_1})$. Specifically, we set

$$\begin{aligned}\xi_1^0 &= -\sum_{i=1}^{K} \Big(a_0^{g,\gamma_1}(\phi_1^{l,i}, D^{\text{ask},i}) - b_0^{g,\gamma_1}(\phi_1^{s,i}, D^{\text{bid},i}) \Big), \\ \xi_1^{l/s,i} &= \phi_1^{l/s,i}, \qquad i = 1, \dots, K,\end{aligned}$$

and

$$\begin{aligned}\xi_t^0 &= \xi_1^0 + \sum_{u=2}^{t} \zeta_u, \qquad t = 2, \dots, T, \\ \xi_t^i &= \phi_t^i, \qquad i = 1, \dots, K, \ t = 2, \dots, T,\end{aligned}$$

where

$$\begin{aligned}\zeta_t^0 =& \sum_{i=1}^{K} \Big(\phi_{t-1}^{l,i} D_{t-1}^{\text{ask},i} - \phi_{t-1}^{s,i} D_{t-1}^{\text{bid},i} \Big) - \sum_{i=1}^{K} \Big(\mathbb{1}_{\Delta\phi_t^{l,i} \geq 0} a_{t-1}^{g,\gamma_1}(\Delta\phi_t^{l,i}, D^{\text{ask},i}) \\ & - \mathbb{1}_{\Delta\phi_t^{l,i} < 0} b_{t-1}^{g,\gamma_1}(-\Delta\phi_t^{l,i}, D^{\text{ask},i}) - \mathbb{1}_{\Delta\phi_t^{s,i} \geq 0} b_{t-1}^{g,\gamma_1}(\Delta\phi_t^{s,i}, D^{\text{bid},i}) \\ & + 1_{\Delta\phi_t^{s,i} < 0} a_{t-1}^{g,\gamma_1}(-\Delta\phi_t^{s,i}, D^{\text{bid},i}) \Big), \qquad t = 2, \dots, T.\end{aligned}$$

First, we will show that $\xi \in \mathcal{S}(0, a^{g,\gamma_1}, b^{g,\gamma_1})$. We have that[3]

$$\begin{aligned}V_0^{\gamma_1}(\xi) =& \Delta\xi_1^0 + \sum_{i=1}^{K} \Big(\mathbb{1}_{\Delta\xi_1^{l,i} \geq 0} P_0^{\text{ask}}(\Delta\xi_1^{l,i}, D^{\text{ask},i}) - \mathbb{1}_{\Delta\xi_1^{l,i} < 0} P_0^{\text{bid}}(-\Delta\xi_1^{l,i}, D^{\text{ask},i}) \\ & - \mathbb{1}_{\Delta\xi_1^{s,i} \geq 0} P_0^{\text{bid}}(\Delta\xi_1^{s,i}, D^{\text{bid}}) + \mathbb{1}_{\Delta\xi_1^{s,i} < 0} P_0^{\text{ask}}(-\Delta\xi_1^{s,i}, D^{\text{bid}}) \Big) \\ =& -\sum_{i=1}^{K} \Big(a_0^{g,\gamma_1}(\phi_1^{l,i}, D^{\text{ask},i}) - b_0^{g,\gamma_1}(\phi_1^{s,i}, D^{\text{bid},i}) \Big) \\ & + \sum_{i=1}^{K} \Big(a_0^{g,\gamma_1}(\phi_1^{l,i}, D^{\text{ask},i}) - b_0^{g,\gamma_1}(\phi_1^{s,i}, D^{\text{bid},i}) \Big) \\ =& 0 = \sum_{i=1}^{K} \Big(\xi_1^{l,i} D_0^{\text{ask},i} - \xi_1^{s,i} D_0^{\text{bid},i} \Big),\end{aligned}$$

---

[3]Here, we are using the convention that V^{γ_1} is computed relative to the model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_1})$, and that V^{γ_1,γ_2} is computed relative to the model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_2})$.

and

$$
\begin{aligned}
\Delta\xi^0_{t+1} + \sum_{i=1}^{K} \Big(& \mathbb{1}_{\Delta\xi^{l,i}_{t+1}\geq 0} P^{\text{ask}}_t(\Delta\xi^{l,i}_{t+1}, D^{\text{ask},i}) - \mathbb{1}_{\Delta\xi^{l,i}_{t+1}<0} P^{\text{bid}}_t(-\Delta\xi^{l,i}_{t+1}, D^{\text{ask},i}) \\
& - \mathbb{1}_{\Delta\xi^{s,i}_{t+1}\geq 0} P^{\text{bid}}_t(\Delta\xi^{s,i}_{t+1}, D^{\text{bid}}) + \mathbb{1}_{\Delta\xi^{s,i}_{t+1}<0} P^{\text{ask}}_t(-\Delta\xi^{s,i}_{t+1}, D^{\text{bid}}) \Big) \\
=\zeta^0_{t+1} + \sum_{i=1}^{K} \Big(& \mathbb{1}_{\Delta\phi^{l,i}_s\geq 0} a^{g,\gamma_1}_{s-1}(\Delta\phi^{l,i}_s, D^{\text{ask},i}) - \mathbb{1}_{\Delta\phi^{l,i}_s<0} b^{g,\gamma_1}_{s-1}(-\Delta\phi^{l,i}_s, D^{\text{ask},i}) \\
& - \mathbb{1}_{\Delta\phi^{s,i}_s\geq 0} b^{g,\gamma_1}_{s-1}(\Delta\phi^{s,i}_s, D^{\text{bid},i}) + 1_{\Delta\phi^{s,i}_s<0} a^{g,\gamma_1}_{s-1}(-\Delta\phi^{s,i}_s, D^{\text{bid},i}) \Big) \\
=\sum_{i=1}^{K} & (\xi^{l,i}_t D^{\text{ask},i}_t - \xi^{s,i}_t D^{\text{bid},i}_t), \qquad t = 2, \ldots, T.
\end{aligned}
$$

Hence, ξ is a self-financing trading strategy and $\xi \in \mathcal{S}(0, a^{g,\gamma_1}, b^{g,\gamma_1})$.

Next we will show that $\xi^0_1 \geq \phi^0_1$ and $\Delta\xi^0_t \geq \Delta\phi^0_t$, $t \in \{2, \ldots, T\}$, which will imply that

$$
V^{\gamma_1}_T(\xi) = \xi^0_1 + \sum_{t=2}^{T} \xi^0_t \geq \phi^0_1 + \sum_{t=2}^{T} \phi^0_t = V^{\gamma_1,\gamma_2}_T(\phi).
$$

Consequently, we have that $V^{\gamma_1}_T(\xi) \geq 0$ and $\mathbb{P}(V^{\gamma_1}_T(\xi) > 0) > 0$, and thus ξ is an arbitrage opportunity in market model $(\mathcal{M}, a^{g,\gamma_1}, b^{g,\gamma_1})$, which contradicts Theorem 2.3.2.

Since $\gamma_1 \leq \gamma_2$, by Proposition 2.3.3, we get that $b^{g,\gamma_1}_t(\varphi, D) \geq b^{g,\gamma_2}_t(\varphi, D)$, for any $\varphi \in L^\infty_+(\mathcal{F}_t)$ and $D \in \mathcal{D}$. Hence,

$$
\begin{aligned}
\xi^0_1 &= -\sum_{i=1}^{K} \Big(a^{g,\gamma_1}_0(\phi^{l,i}_1, D^{\text{ask},i}) - b^{g,\gamma_1}_0(\phi^{s,i}_1, D^{\text{bid},i}) \Big) \\
&\geq -\sum_{i=1}^{K} \Big(a^{g,\gamma_1}_0(\phi^{l,i}_1, D^{\text{ask},i}) - b^{g,\gamma_2}_0(\phi^{s,i}_1, D^{\text{bid},i}) \Big) \\
&= \phi^0_1,
\end{aligned}
$$

and it is clear that $\Delta\xi_t^0 = \zeta_t$, $t \in \{2, \dots, T\}$. Moreover,

$$\begin{aligned}\Delta\xi_t^0 =&\sum_{i=1}^{K}\Big(\phi_{t-1}^{l,i}D_{t-1}^{\mathrm{ask},i} - \phi_{t-1}^{s,i}D_{t-1}^{\mathrm{bid},i}\Big) - \sum_{i=1}^{K}\Big(\mathbb{1}_{\Delta\phi_t^{l,i}\geq 0}a_{t-1}^{g,\gamma_1}(\Delta\phi_t^{l,i}, D^{\mathrm{ask},i})\\ &- \mathbb{1}_{\Delta\phi_t^{l,i}<0}b_{t-1}^{g,\gamma_1}(-\Delta\phi_t^{l,i}, D^{\mathrm{ask},i}) - \mathbb{1}_{\Delta\phi_t^{s,i}\geq 0}b_{t-1}^{g,\gamma_1}(\Delta\phi_t^{s,i}, D^{\mathrm{bid},i})\\ &+ 1_{\Delta\phi_t^{s,i}<0}a_{t-1}^{g,\gamma_1}(-\Delta\phi_t^{s,i}, D^{\mathrm{bid},i})\Big)\\ \geq&\sum_{i=1}^{K}\Big(\phi_{t-1}^{l,i}D_{t-1}^{\mathrm{ask},i} - \phi_{t-1}^{s,i}D_{t-1}^{\mathrm{bid},i}\Big) - \sum_{i=1}^{K}\Big(\mathbb{1}_{\Delta\phi_t^{l,i}\geq 0}a_{t-1}^{g,\gamma_1}(\Delta\phi_t^{l,i}, D^{\mathrm{ask},i})\\ &- \mathbb{1}_{\Delta\phi_t^{l,i}<0}b_{t-1}^{g,\gamma_2}(-\Delta\phi_t^{l,i}, D^{\mathrm{ask},i}) - \mathbb{1}_{\Delta\phi_t^{s,i}\geq 0}b_{t-1}^{g,\gamma_2}(\Delta\phi_t^{s,i}, D^{\mathrm{bid},i})\\ &+ 1_{\Delta\phi_t^{s,i}<0}a_{t-1}^{g,\gamma_1}(-\Delta\phi_t^{s,i}, D^{\mathrm{bid},i})\Big)\\ =&\Delta\phi_t^0,\end{aligned}$$

for every $t = 2, \dots, T$. Therefore, we have that $\xi_T^0 = \xi_1^0 + \sum_{t=2}^{T}\Delta\xi_t^0 \geq \phi_1^0 + \sum_{t=2}^{T}\Delta\phi_t^0 = \phi_T^0$.

The proof for $\gamma_1 \geq \gamma_2$ is analogous.

□

2.4 Derivatives Valuation with Dynamic Conic Finance

The aim of this section is to build a pricing methodology for general contingent claims, by using the theory of dynamic acceptability indices. We assume that there exists an underlying market[4] $(\mathcal{M}, P^{\mathrm{ask}}, P^{\mathrm{bid}})$ that satisfies conditions (M1)-(M6). We will use the securities from underlying market as hedging instruments.

2.4.1 Conic Valuation of Derivative Cash flows. We start by introducing the concept of super-hedging in our context. and then an extension of $\mathcal{H}^0(t)$.

Definition 2.4.1. *Let us fix $t \in \mathcal{T}$, and let $V(\phi)$ be the liquidation value process generated by a trading strategy $\phi \in \mathcal{S}(t)$. A process $\overline{V} \in \mathcal{D}$, is said to be super-hedged*

[4]Note that in view of previous section such markets exist.

at zero-cost, at time t, by ϕ, if there exists $Z \in \mathcal{L}_+(t)$, where

$$\mathcal{L}_+(t) := \Big\{ (Z_s)_{s=0}^T : Z_s \in L^2_+(\Omega, \mathcal{F}_s, \mathbb{P}),\ Z_s = 0,\ s \le t \Big\},$$

such that

$$\overline{V}_s = V_s(\phi) - Z_s, \qquad s = 0, \ldots, T, \tag{2.41}$$

If in (2.41), *$Z = 0$, then we say that $\overline{V}$ can be replicated by ϕ.*

We now introduce the set of cash flows that can be super-hedged by strategies $\phi \in \mathcal{S}(t)$ at zero cost:

$$\mathcal{H}(t) := \Big\{ \Big(0, \ldots, 0, \Delta(V_{t+1}(\phi) - Z_{t+1}), \ldots, \Delta(V_T(\phi) - Z_T)\Big) : \phi \in \mathcal{S}(t),\ Z \in \mathcal{L}_+(t) \Big\}. \tag{2.42}$$

We proceed by defining acceptability ask and bid prices for a cash flow $D \in \mathcal{D}$.

Definition 2.4.2. *Let $g = (g_x)_{x>0}$ be a family of drivers that satisfy Assumption G. The acceptability ask price of $\varphi \in L^\infty_+(\mathcal{F}_t)$ shares of the cash flow D, at level γ, at time $t \in \mathcal{T}$, is defined as*

$$\widehat{a}_t^{g,\gamma}(\varphi, D) = \operatorname{ess\,inf}\{a \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t) \ \textit{so that}\ \alpha_t^g(\delta_t(a) + H - \delta_t^+(\varphi D)) \ge \gamma\}, \tag{2.43}$$

and the acceptability bid price of $\varphi \ge 0$, $\varphi \in L^\infty_+(\mathcal{F}_t)$, shares of D, at level γ, at time $t \in \mathcal{T}$ is defined as

$$\widehat{b}_t^{g,\gamma}(\varphi, D) = \operatorname{ess\,sup}\{b \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t) \ \textit{so that}\ \alpha_t^g(\delta_t^+(\varphi D) + H - \delta_t(b)) \ge \gamma\}. \tag{2.44}$$

Remark 2.4.1. *(i) If $\mathcal{H}(t)$ is equal to $\{(0, \ldots, 0)\}$, which means hedging is not admitted, then*

$$\begin{aligned}\widehat{a}_t^{g,\gamma}(\varphi, D) &= \operatorname{ess\,inf}\{a \in L^2(\mathcal{F}_t) : \alpha_t^g(\delta_t(a) - \delta_t^+(\varphi D)) \ge \gamma\} \\ &= a_t^{g,\gamma}(\varphi, D),\end{aligned}$$

and

$$\widehat{b}_t^{g,\gamma}(\varphi, D) = \operatorname{ess\,sup}\{b \in L^2(\mathcal{F}_t) : \alpha_t^g(\delta_t^+(\varphi D) - \delta_t(b)) \geq \gamma\}$$
$$= b_t^{g,\gamma}(\varphi, D).$$

(ii) Clearly, $\widehat{a}_t^{g,\gamma}(\varphi, D) \leq a_t^{g,\gamma}(\varphi, D)$ *and* $\widehat{b}_t^{g,\gamma}(\varphi, D) \geq b_t^{g,\gamma}(\varphi, D)$*, for* $\varphi \in L_+^\infty(\mathcal{F}_t)$*,* $D \in \mathcal{D}$*.*

Similarly as in Section 2.3.3, we note that $\widehat{a}_t^{g,\gamma}(\varphi, D) = \widehat{a}_t^{g,\gamma}(1, \varphi D)$ and $\widehat{b}_t^{g,\gamma}(\varphi, D) = \widehat{b}_t^{g,\gamma}(1, \varphi D)$, and thus, we will prove most of the results for $\widehat{a}_t^{g,\gamma}(1, D)$ and $\widehat{b}_t^{g,\gamma}(1, D)$, from which the general case will follow.

Proposition 2.4.1. *The acceptability ask and bid prices admit the following representations*

$$\begin{aligned} \widehat{a}_t^{g,\gamma}(D) &= \operatorname*{ess\,inf}_{H \in \mathcal{H}(t)} \mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\right], \\ \widehat{b}_t^{g,\gamma}(D) &= \operatorname*{ess\,sup}_{H \in \mathcal{H}(t)} -\mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(-H_s - D_s)\Big|\mathcal{F}_t\right], \end{aligned} \tag{2.45}$$

for $D \in \mathcal{D}$*, at level* $\gamma > 0$*, at time* $t \in \mathcal{T}$*.*

Proof. We show the proof of first equality; the proof for the bid price is similar.

From the definition of $\widehat{a}_t^{g,\gamma}$, we get that

$$\widehat{a}_t^{g,\gamma}(D) = \operatorname{ess\,inf}\left\{a \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t), \mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\right] \leq a\right\},$$

and consequently

$$\begin{aligned} &\operatorname*{ess\,inf}_{H \in \mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big] \\ &\qquad \leq \operatorname{ess\,inf}\Big\{a \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t), \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big] \leq a\Big\}. \end{aligned} \tag{2.46}$$

To prove the converse inequality, we show that strict inequality in (2.46) does not hold true. Assume that on some set $A \in \mathcal{F}_t$, $\mathbb{P}(A) > 0$, we have that

$$\begin{aligned}
&\operatorname*{ess\,inf}_{H\in\mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big] \\
&\qquad < \operatorname{ess\,inf}\Big\{a \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t),\ \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big] \le a\Big\}.
\end{aligned}$$

Then, there exists an $H' \in \mathcal{H}$, such that on A

$$\begin{aligned}
&\mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(D_s - H'_s)\Big|\mathcal{F}_t\right] \\
&\qquad < \operatorname{ess\,inf}\left\{a \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t),\ \mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\right] \le a\right\}.
\end{aligned}$$

Consider $b \in L^2(\mathcal{F}_t)$, such that on set A

$$\begin{aligned}
&\mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(D_s - H'_s)\Big|\mathcal{F}_t\right] < b \\
&\qquad < \operatorname{ess\,inf}\left\{a \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t),\ \mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\right] \le a\right\}.
\end{aligned} \tag{2.47}$$

Then, we have that

$$\mathbb{1}_A b \in \left\{a \in L^2(\mathcal{F}_t) : \exists H \in \mathcal{H}(t),\ \mathbb{1}_A\mathcal{E}_{g_\gamma}\left[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\right] \le a\right\}.$$

Hence, for almost all $\omega \in A$, we get $\widehat{a}_t^{g,\gamma}(D)(\omega) \le b(\omega)$. However, by (2.47), we also have that $\widehat{a}_t^{g,\gamma}(D)(\omega) > b(\omega)$ for such ω's, that leads to a contradiction. This concludes the proof. □

As an immediate consequence of Proposition 2.4.1, a technical result is provided, which shows a "symmetry" between acceptability ask and bid prices, that will be used later.

Corollary 2.4.1. *Let $g = (g_x)_{x>0}$ be a family of drivers that satisfy Assumption G. Then,*

$$\widehat{a}_t^{g,\gamma}(D) = -\widehat{b}_t^{g,\gamma}(-D),$$

for $D \in \mathcal{D}$, $\gamma > 0$, $t \in \mathcal{T}$.

2.4.2 Arbitrage. Here, we will define and study arbitrage with regard to the extended market model consisting of the underlying market $(\mathcal{M}, P^{\text{ask}}, P^{\text{bid}})$ and of the (derivative) cash flows priced according to acceptability bid and ask prices given in Definition 2.4.2.

Towards this end we first introduce some relevant notions and notations. Fix some $t \in \mathcal{T}$. Let $D \in \mathcal{D}$ be a cash flow, and let $\varphi \in L^\infty(\mathcal{F}_t)$ be the number of shares that D is traded. Assume that the ask price for $|\varphi|$ shares of D at time t is $\widehat{P}_t^{\text{ask}}(|\varphi|, D)$, and respectively, the bid price is $\widehat{P}_t^{\text{bid}}(|\varphi|, D)$. We say that $S_t(\varphi, D) : L^\infty(\mathcal{F}_t) \times \mathcal{D} \to L^2(\mathcal{F}_t)$ is the set-up cost for φ shares of D at time $t \in \mathcal{T}$ if

$$S_t(\varphi, D) = \mathbb{1}_{\varphi \geq 0} \widehat{P}_t^{\text{ask}}(\varphi, D) - \mathbb{1}_{\varphi < 0} \widehat{P}_t^{\text{bid}}(-\varphi, D).$$

Accordingly, we denote by $\widehat{\mathcal{D}}(t) = \{(0, \dots, -S_t(\varphi, D), \varphi D_{t+1}, \dots, \varphi D_T) : D \in \mathcal{D}, \varphi \in L^\infty(\mathcal{F}_t)\}$ the set of all derivative cash flows initiated at set-up cost S_t, and time $t \in \mathcal{T}$. It is clear that $\widehat{\mathcal{D}}(t) \subset \mathcal{D}$.

Definition 2.4.3. *An arbitrage opportunity at time $t \in \mathcal{T}$ is a pair $(\widehat{D}, \phi)$ consisting of derivative cash flow $\widehat{D} \in \widehat{\mathcal{D}}(t)$ and a trading strategy $\phi \in \mathcal{S}(t)$, such that $V_T(\phi) + \sum_{s=t}^T \widehat{D}_s \geq 0$ and $\mathbb{P}(V_T(\phi) + \sum_{s=t}^T \widehat{D}_s > 0) > 0$.*

Similarly to Proposition 2.3.1, we will characterize arbitrage opportunities in the derivative market model in terms of cash flows.

Proposition 2.4.2. *Fix $t \in \mathcal{T}$. The following statements are equivalent:*

(1) There exists an arbitrage opportunity at time t.

(2) There exists a derivative cash flow $\widehat{D} \in \widehat{\mathcal{D}}(t)$ and a super-hedging cash flow $H \in \mathcal{H}(t)$, such that $\sum_{s=t}^T (\widehat{D}_s + H_s) \geq 0$ and $\mathbb{P}\big(\sum_{s=t}^T (\widehat{D}_s + H_s) > 0\big) > 0$.

Proof. (1) $\Rightarrow$ (2) Assume that there exists an arbitrage opportunity at time t. Then, according to Definition 2.4.3, there exists $\widehat{D} \in \widehat{\mathcal{D}}(t)$ and $\phi \in \mathcal{S}(t)$, such that

$$\begin{aligned} V_T(\phi) + \sum_{s=t}^{T} \widehat{D}_s \geq 0, \\ \mathbb{P}(V_T(\phi) + \sum_{s=t}^{T} \widehat{D}_s > 0) > 0. \end{aligned} \tag{2.48}$$

By definition of $\mathcal{H}(t)$, we have that $H = (0, \ldots, 0, \Delta V_{t+1}(\phi), \ldots, \Delta V_T(\phi)) \in \mathcal{H}(t)$. Then according to (2.48), it follows that $\sum_{s=t}^{T}(\widehat{D}_s + H_s) \geq 0$ and $\mathbb{P}\big(\sum_{s=t}^{T}(\widehat{D}_s + H_s) > 0\big) > 0$.

(2) $\Rightarrow$ (1) Assume that there exists $\widehat{D} \in \widehat{\mathcal{D}}(t)$, and $H \in \mathcal{H}$, such that $\sum_{s=t}^{T}(\widehat{D}_s + H_s) \geq 0$ and $\mathbb{P}\big(\sum_{s=t}^{T}(\widehat{D}_s + H_s) > 0\big) > 0$. Then, by definition of $\mathcal{H}(t)$, there exists $\phi \in \mathcal{S}(t)$ such that

$$\begin{aligned} & V_T(\phi) + \sum_{s=t}^{T} \widehat{D}_s \geq \sum_{s=t}^{T} H_s + \sum_{s=t}^{T} \widehat{D}_s \geq 0, \\ & \mathbb{P}\Big(V_T(\phi) + \sum_{s=t}^{T} \widehat{D}_s > 0\Big) \geq \mathbb{P}\Big(\sum_{s=t}^{T} H_s + \sum_{s=t}^{T} \widehat{D}_s > 0\Big) > 0. \end{aligned}$$

Hence, $(\widehat{D}, \phi)$ is an arbitrage opportunity. □

2.4.3 Good Deals and No-Good Deals. Next, we introduce the concept of good deals for sets of cash flows, which plays an essential role in the theory of no-arbitrage pricing in the extended market, and in derivation of fundamental properties of the acceptability ask and bid prices. Towards this end, let $(\mu^x)_{x>0}$ be an increasing family of DCRMs (cf. Definition 2.2.1):

Definition 2.4.4. *A good deal for $\mathcal{H}(t)$, $t \in \mathcal{T}$, at level $\gamma > 0$, is a cash flow $H \in \mathcal{H}(t)$, such that $\mu_t^\gamma(H)(\omega) < 0$, for $\omega \in A$, for some $A \in \mathcal{F}_t$, and $\mathbb{P}(A) > 0$. Respectively, we say that the no-good deal condition (NGD) holds true for $\mathcal{H}(t)$, at time $t \in \mathcal{T}$, and at level $\gamma > 0$, if $\mu_t^\gamma(H) \geq 0$ for any $H \in \mathcal{H}(t)$.*

In the rest of this chapter, we take $\mu^x := \rho^{g_x}, x > 0$, where $g = (g_x)_{x>0}$ is a family of drivers that satisfy Assumption G, and $(\rho^{g_x})_{x>0}$ is the family of DCRMs such that for each $x > 0$ the DCRM ρ^{g_x} is given as in (2.18).[5]

We proceed by proving some technical results as preparation for showing the relationship between NGD and no-arbitrage pricing in the extended market. We begin with three technical lemmas.

Lemma 2.4.1. *Let $\phi \in \mathcal{S}(t)$, and let $\mathcal{L}_+(t)$ be defined as*

$$\mathcal{L}_+(t) := \Big\{(Z_s)_{s=0}^T : Z_s \in L^2_+(\Omega, \mathcal{F}_s, \mathbb{P}), Z_s = 0,\ s \leq t\Big\},$$

for $t = 0, \dots, T-1$. Then, for any $t \in \{0, \dots, T-1\}$, the set

$$\mathcal{H}(t) := \Big\{\Big(0, \dots, 0, \Delta(V_{t+1}(\phi) - Z_{t+1}), \dots, \Delta(V_T(\phi) - Z_T)\Big) : \phi \in \mathcal{S}(t), Z \in \mathcal{L}_+(t)\Big\}$$

is a convex set.

Proof. Let H^1, $H^2 \in \mathcal{H}(t)$, and let $\lambda \in L^\infty(\mathcal{F}_t)$ such that $0 \leq \lambda \leq 1$. Then, there exists ϕ, $\xi \in \mathcal{S}(t)$, Z^1, $Z^2 \in \mathcal{L}_+(t)$, such that

$$H^1 = \Big(0, \dots, 0, \Delta(V_{t+1}(\phi) - Z^1_{t+1}), \dots, \Delta(V_T(\phi) - Z^1_T)\Big),$$

and

$$H^2 = \Big(0, \dots, 0, \Delta(V_{t+1}(\xi) - Z^2_{t+1}), \dots, \Delta(V_T(\xi) - Z^2_T)\Big).$$

It can be proved that there exists $\theta \in \mathcal{S}(t)$, such that $\lambda V_s(\phi) + (1-\lambda)V_s(\xi) \leq V_s(\theta)$ for any $s = t+1, \dots, T$. Therefore,

$$\begin{aligned}\lambda \sum_{u=t+1}^{s} H^1_u + (1-\lambda) \sum_{u=t+1}^{s} H^2_u &= \lambda V_s(\phi) + (1-\lambda)V_s(\xi) - \lambda Z^1_s - (1-\lambda)Z^2_s \\ &\leq V_s(\theta) - (\lambda Z^1_s + (1-\lambda)Z^2_s) \leq V_s(\theta),\end{aligned}$$

[5]Note though, that in (2.18) symbol g represents a single driver.

for any $s = t+1, \dots, T$. By Definition 2.4.1, we have that $\lambda H^1 + (1-\lambda)H^2 \in \mathcal{H}(t)$. The proof is complete. □

Lemma 2.4.2. *Fix $t \in \mathcal{T}$, and level $\gamma > 0$. Assume that there exists $B \in \mathcal{F}_t$, such that $\mathbb{1}_B \rho_t^{g_\gamma}(H) \geq 0$ for any $H \in \mathcal{H}(t)$. Then, for any $D \in \mathcal{D}$, $\mathbb{1}_B \widehat{a}_t^{g,\gamma}(D) \geq \mathbb{1}_B \widehat{b}_t^{g,\gamma}(D)$.*

Proof. We prove the statement by contradiction. Assume that there exists some $D \in \mathcal{D}$, $B' \subset B$, such that $\widehat{b}_t^{g,\gamma}(D)(\omega) > \widehat{a}_t^{g,\gamma}(D)(\omega)$, on B'. Then, by Proposition 2.4.1, we have that

$$\operatorname*{ess\,sup}_{H \in \mathcal{H}(t)} -\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(-H_s - D_s)\Big|\mathcal{F}_t\Big](\omega) > \operatorname*{ess\,inf}_{H \in \mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big](\omega),$$

where $\omega \in B'$. Let $M = (\widehat{b}_t^{g,\gamma}(D) + \widehat{a}_t^{g,\gamma}(D))/2$. Then, there exists $H^1, H^2 \in \mathcal{H}(t)$ such that

$$-\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(-H_s^1 - D_s)\Big|\mathcal{F}_t\Big](\omega) > M(\omega) > \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s^2)\Big|\mathcal{F}_t\Big](\omega),$$

for $\omega \in B'$. Hence, we get that

$$\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(-H_s^1 - D_s)\Big|\mathcal{F}_t\Big](\omega) + \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s^2)\Big|\mathcal{F}_t\Big](\omega) < 0, \ \omega \in B'. \tag{2.49}$$

On the other hand, in view of Proposition 2.1.1.(vi), we have

$$\begin{aligned}
&\mathcal{E}_{g_\gamma}\Big[\frac{1}{2}\Big(\sum_{s=t+1}^{T}(-H_s^1 - D_s) + \sum_{s=t+1}^{T}(D_s - H_s^2)\Big)\Big|\mathcal{F}_t\Big] \\
\leq &\frac{1}{2}\Big(\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(-H_s^1 - D_s)\Big|\mathcal{F}_t\Big] + \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s - H_s^2)\Big|\mathcal{F}_t\Big]\Big),
\end{aligned}$$

which combined with (2.49), and using (2.18), we obtain

$$\rho_t^{g_\gamma}(\frac{1}{2}(H^1 + H^2))(\omega) = \mathcal{E}_{g_\gamma}\Big[\frac{1}{2}\sum_{s=t+1}^{T}(-H_s^1 - H_s^2)\Big|\mathcal{F}_t\Big](\omega) < 0,$$

for $\omega \in B'$. However, by Lemma 2.4.1, we have that $\frac{1}{2}(H^1 + H^2) \in \mathcal{H}(t)$, and since $\mathbb{1}_B \rho_t^{g_\gamma}(H) \geq 0$ for any $H \in \mathcal{H}(t)$, we have that $\mathbb{1}_B \rho_t^{g_\gamma}(\frac{1}{2}(H^1 + H^2)) \geq 0$, which leads to a conclusion that $\mathbb{P}(B') = 0$, and hence, $\mathbb{1}_B \widehat{a}_t^{g,\gamma}(D) \geq \mathbb{1}_B \widehat{b}_t^{g,\gamma}(D)$. The proof is complete. □

An application of Lemma 2.4.2, which is also useful to our study of arbitrage, is stated as follows:

Lemma 2.4.3. *Fix $t \in \mathcal{T}$, and level $\gamma > 0$. Assume that there exists $B \in \mathcal{F}_t$, such that $\mathbb{1}_B \rho_t^{g\gamma}(H) \geq 0$ for any $H \in \mathcal{H}(t)$. Then, $\mathbb{1}_B \widehat{a}_t^{g,\gamma}(D) = \mathbb{1}_B \widehat{b}_t^{g,\gamma}(D) = 0$, for any $D \in \mathcal{H}(t)$.*

Proof. Since $D \in \mathcal{H}(t)$, then $D - \delta_t^+(D) = 0$, and thus

$$\widehat{a}_t^{g,\gamma}(D) = \operatorname{ess\,inf}\{a \in \mathcal{F}_t : \exists H \in \mathcal{H}(t), \text{ s.t. } \alpha_t^g(\delta_t(a) + H - \delta_t^+(D)) \geq \gamma\} \leq 0.$$

Similarly, $\widehat{b}_t^{g,\gamma}(D) \geq 0$, and therefore $\widehat{a}_t^{g,\gamma}(D) \leq 0 \leq \widehat{b}_t^{g,\gamma}(D)$. On the other hand, by Lemma 2.4.2, $\mathbb{1}_B \widehat{a}_t^{g,\gamma}(D) \geq \mathbb{1}_B \widehat{b}_t^{g,\gamma}(D)$. Hence, $\mathbb{1}_B \widehat{a}_t^{g,\gamma}(D) = \mathbb{1}_B \widehat{b}_t^{g,\gamma}(D) = 0$. □

With the help of Lemma 2.4.2 and Lemma 2.4.3 we can now prove the main theorem in this section.

Theorem 2.4.1. *Fix $\gamma > 0$, $t \in \mathcal{T}$. Assume that*

$$\widehat{P}^{ask}(\varphi, D) = \widehat{a}_t^{g,\gamma}(\varphi, D), \quad \widehat{P}^{bid}(\varphi, D) = \widehat{b}_t^{g,\gamma}(\varphi, D), \tag{2.50}$$

and consequently,

$$S_t(\varphi, D) = \mathbb{1}_{\varphi \geq 0} \widehat{a}_t^{g,\gamma}(\varphi, D) - \mathbb{1}_{\varphi < 0} \widehat{b}_t^{g,\gamma}(-\varphi, D),$$

for any $D \in \mathcal{D}$, and $\varphi \in L^\infty(\mathcal{F}_t)$. Then, NGD holds for $\mathcal{H}(t)$ at level $\gamma > 0$, if and only if there is no arbitrage opportunity at time t.

Proof. ($\Longrightarrow$) Assume that NGD holds for $\mathcal{H}(t)$ at level $\gamma > 0$, and assume that there is an arbitrage opportunity at time t. Then, according to Proposition 2.4.2, there exists $\widehat{D} \in \widehat{\mathcal{D}}(t)$ and $H \in \mathcal{H}(t)$ such that $\sum_{s=t}^T (\widehat{D}_s + H_s) \geq 0$ and $\mathbb{P}(\sum_{s=t}^T (\widehat{D}_s + H_s) > 0) > 0$.

Due to (2.50), there exists $D \in \mathcal{D}$, and $\varphi \in L^\infty(\mathcal{F}_t)$ such that

$$\sum_{s=t+1}^{T} (H_s - \varphi D_s) - (\mathbb{1}_{\varphi \le 0} \widehat{a}_t^{g,\gamma}(-\varphi, D) - \mathbb{1}_{\varphi > 0} \widehat{b}_t^{g,\gamma}(\varphi, D)) \ge 0,$$

$$\mathbb{P}(\sum_{s=t+1}^{T} (H_s - \varphi D_s) - (\mathbb{1}_{\varphi \le 0} \widehat{a}_t^{g,\gamma}(-\varphi, D) - \mathbb{1}_{\varphi > 0} \widehat{b}_t^{g,\gamma}(\varphi, D)) > 0) > 0.$$

In view of Corollary 2.4.1, we get that

$$\sum_{s=t+1}^{T} (H_s - \varphi D_s) + \widehat{b}_t^{g,\gamma}(\varphi D) \ge 0,$$

$$\mathbb{P}(\sum_{s=t+1}^{T} (H_s - \varphi D_s) + \widehat{b}_t^{g,\gamma}(\varphi D) > 0) > 0. \tag{2.51}$$

Since NGD holds for $\mathcal{H}(t)$ at level γ, then by Lemma 2.4.2,

$$\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} (\varphi D_s - H_s)\Big|\mathcal{F}_t\Big] \ge \widehat{a}_t^{g,\gamma}(\varphi D) \ge \widehat{b}_t^{g,\gamma}(\varphi D) \ge \sum_{s=t+1}^{T} (\varphi D_s - H_s),$$

and (2.51) implies that

$$\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} (\varphi D_s - H_s)\Big|\mathcal{F}_t\Big] > \sum_{s=t+1}^{T} (\varphi D_s - H_s),$$

with strictly positive probability. Consequently, by Theorem 2.1.1, we get that

$$\begin{aligned}\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} (\varphi D_s - H_s)\Big] &= \mathcal{E}_{g_\gamma}\Big[\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} (\varphi D_s - H_s)\Big|\mathcal{F}_t\Big]\Big] \\ &\ge \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T} (\varphi D_s - H_s)\Big],\end{aligned}$$

with strict inequality holding true on some set A such that $\mathbb{P}(A) > 0$, which leads to contradiction. Hence, there exists no arbitrage opportunity.

($\Longleftarrow$) To prove that absence of arbitrage at time t implies that NGD holds for $\mathcal{H}(t)$ at level $\gamma > 0$, we will show that if NGD does not hold for $\mathcal{H}(t)$, then there exists an arbitrage opportunity at time t. In fact, if NGD does not hold true, then there exists some $H' \in \mathcal{H}(t)$ and $A \in \mathcal{F}_t$ such that $\mathbb{P}(A) > 0$ and $\rho_t^{g_\gamma}(H') < 0$ on A. Without loss of generality, we assume that $\mathbb{1}_{A^c}\rho_t^{g_\gamma}(H) \ge 0$ for any $H \in \mathcal{H}(t)$.

Let us consider $D = (0, \dots, 0)$, then by Proposition 2.4.1, the acceptability ask price of D is

$$\widehat{a}_t^{g,\gamma}(D) = \operatorname*{ess\,inf}_{H \in \mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[- \sum_{s=t+1}^{T} H_s \Big| \mathcal{F}_t \Big] \leq \rho_t^{g_\gamma}(H'),$$

and therefore

$$\widehat{a}_t^{g,\gamma}(D)(\omega) \leq \rho_t^{g_\gamma}(H')(\omega) < 0,$$

for any $\omega \in A$. Also note that $D \in \mathcal{H}(t)$, then Lemma 2.4.3 implies that $\mathbb{1}_{A^c}\widehat{a}_t^{g,\gamma}(D) = 0$. Thus, the pair $(\widehat{D}, \phi)$, where $\widehat{D} = (0, \dots, 0, -\widehat{a}_t^{g,\gamma}(D), 0, \dots, 0)$ and $\phi = 0$, satisfies that

$$V_T(\phi) + \sum_{s=t}^{T} \widehat{D}_s = -\widehat{a}_t^{g,\gamma}(D) \geq 0,$$

$$\mathbb{P}\Big(V_T(\phi) + \sum_{s=t}^{T} \widehat{D}_s > 0 \Big) = \mathbb{P}(A) > 0.$$

According to Proposition 2.4.2, there exists an arbitrage opportunity at time t, and the proof is complete. □

Remark 2.4.2. *In [BCIR13], the authors prove, in case of* $\widehat{\mathcal{D}} = \{0\}$ *and pricing according to dynamic coherent risk measure, that absence of arbitrage is equivalent to NGD at the some level* $\gamma > 0$*, and the derived acceptability ask and bid prices are no-arbitrage prices. In our set-up, those two results are implications of Theorem 2.4.1.*

In fact, the notion of NGD not only plays an essential role in no-arbitrage pricing of derivative market, but also is crucial to the study of properties of the proposed acceptability ask and bid prices. We will show that, under NGD condition, acceptability ask and bid prices satisfy nice properties similar to those in Theorem 2.3.1. To start, we give the next two propositions without proof as they are direct applications of Lemma 2.4.2 and 2.4.3.

Proposition 2.4.3. *Assume that NGD holds for $\mathcal{H}(t)$ at level $\gamma > 0$ and some fixed $t \in \mathcal{T}$. Then, for any $D \in \mathcal{D}$, we have $\widehat{a}_t^{g,\gamma}(D) \geq \widehat{b}_t^{g,\gamma}(D)$.*

Recall Theorem 2.3.1 P2, the pricing operators $a_t^{g,\gamma}$ and $b_t^{g,\gamma}$ satisfy the property of non-negative spread. Proposition 2.4.3 shows that under the assumption of NGD, the acceptability ask and bid prices $\widehat{a}_t^{g,\gamma}$ and $\widehat{b}_t^{g,\gamma}$ also admit a similar result. Such proposition is in accord with observations from the real market. We also want to stress that, as discussed in [BCIR13], if NGD does not hold, then the ask-bid spread is equal to $-\infty$.

The following result shows that under NGD, by implementing the acceptability pricing method, the set-up cost of cash flows which can be super-hedged at zero cost is indeed 0.

Proposition 2.4.4. *Assume that NGD holds for $\mathcal{H}(t)$ at level $\gamma > 0$, at time $t \in \mathcal{T}$. Then, $\widehat{a}_t^{g,\gamma}(D) = \widehat{b}_t^{g,\gamma}(D) = 0$, for any $D \in \mathcal{H}(t)$.*

Next we will show that pricing the underlying securities $D \in \mathcal{M}$ by the bid and ask acceptability prices defined by (2.44) and (2.43) yields the market prices $P^{\text{ask/bid}}$ of these securities.

Proposition 2.4.5. *Fix $t \in \mathcal{T}$. Assume that NGD holds for $\mathcal{H}(t)$ at level $\gamma > 0$, at time $t \in \mathcal{T}$. Then,*

$$\widehat{a}_t^{g,\gamma}(\varphi, \widetilde{D}) = P_t^{ask}(\varphi, \widetilde{D})$$
$$\widehat{b}_t^{g,\gamma}(\varphi, \widetilde{D}) = P_t^{bid}(\varphi, \widetilde{D}),$$

for any $\varphi \in L_+^\infty(\mathcal{F}_t)$, and $\widetilde{D} \in \mathcal{M}$.

Proof. Since $\widetilde{D} \in \mathcal{M}$, then $\overline{H} := (0, \ldots, 0, \varphi\widetilde{D}_{t+1} - P_t^{\text{ask}}(\varphi, \widetilde{D}), \varphi\widetilde{D}_{t+2}, \ldots, \varphi\widetilde{D}_T) \in$

$\mathcal{H}(t)$. Therefore, by Proposition 2.4.4, we get $\widehat{a}_t^{g,\gamma}(\overline{H}) = 0$, which is equivalent to

$$\operatorname*{ess\,inf}_{H\in\mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(\varphi\widetilde{D}_s - H_s) - P_t^{\text{ask}}(\varphi,\widetilde{D})\Big|\mathcal{F}_t\Big] = 0.$$

Note that $P_t^{\text{ask}}(\varphi,\widetilde{D})$ is $\mathcal{F}_t$-measurable, and it does not depend on the argument $H \in \mathcal{H}(t)$ over which the ess inf is taken. Hence, we immediately get that

$$P_t^{\text{ask}}(\varphi,\widetilde{D}) = \operatorname*{ess\,inf}_{H\in\mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(\varphi\widetilde{D}_s - H_s)\Big|\mathcal{F}_t\Big] = \widehat{a}_t^{g,\gamma}(\varphi,\widetilde{D}).$$

The proof for the bid price is analogous. □

In [CM10] and [BCIR13], the authors consider ask-bid prices produced by coherent acceptability indices. However, in some literature there are arguments that coherent acceptability indices fail to take liquidity risk into account. In our set-up, acceptability indices are assumed to be quasi-concave and we will show that corresponding ask-bid prices reflect liquidity risk as in the following proposition.

Proposition 2.4.6. *The acceptability ask and bid prices satisfy*

$$\widehat{a}_t^{g,\gamma}(\lambda D^1 + (1-\lambda)D^2) \leq \lambda\widehat{a}_t^{g,\gamma}(D^1) + (1-\lambda)\widehat{a}_t^{g,\gamma}(D^2), \tag{2.52}$$

$$\widehat{b}_t^{g,\gamma}(\lambda D^1 + (1-\lambda)D^2) \geq \lambda\widehat{b}_t^{g,\gamma}(D^1) + (1-\lambda)\widehat{b}_t^{g,\gamma}(D^2), \tag{2.53}$$

for $D^1, D^2 \in \mathcal{D}$, $\lambda \in \mathcal{F}_t$, $0 \leq \lambda \leq 1$, *at level* $\gamma > 0$, *at time* $t \in \mathcal{T}$.

Proof. In view of Proposition 2.4.1, and by convexity of g-expectation, for any $H^1, H^2 \in \mathcal{H}(t)$, $\lambda \in \mathcal{F}_t$,

$$\begin{aligned}\lambda\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s^1 - H_s^1)\Big|\mathcal{F}_t\Big] &+ (1-\lambda)\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s^2 - H_s^2)\Big|\mathcal{F}_t\Big] \\ &\geq \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(\lambda D_s^1 + (1-\lambda)D_s^2 - H_s^3)\Big|\mathcal{F}_t\Big],\end{aligned}$$

where $H^3 = \lambda H^1 + (1-\lambda)H^2$. Due to convexity of $\mathcal{H}(t)$ (see Lemma 2.4.1), we have that $H^3 \in \mathcal{H}(t)$. Consequently, using Proposition 2.4.1, we continue

$$\begin{aligned}
&\lambda\widehat{a}_t^{g,\gamma}(D^1) + (1-\lambda)\widehat{a}_t^{g,\gamma}(D^2) \\
&\qquad = \operatorname*{ess\,inf}_{H^1,H^2\in\mathcal{H}(t)} \Big(\lambda\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s^1 - H_s^1)\Big|\mathcal{F}_t\Big] + (1-\lambda)\mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(D_s^2 - H_s^2)\Big|\mathcal{F}_t\Big]\Big) \\
&\qquad \geq \operatorname*{ess\,inf}_{H^1,H^2\in\mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(\lambda D_s^1 + (1-\lambda)D_s^2 - H_s^3)\Big|\mathcal{F}_t\Big] \\
&\qquad \geq \operatorname*{ess\,inf}_{H\in\mathcal{H}(t)} \mathcal{E}_{g_\gamma}\Big[\sum_{s=t+1}^{T}(\lambda D_s^1 + (1-\lambda)D_s^2 - H_s)\Big|\mathcal{F}_t\Big] \\
&\qquad = \widehat{a}_t^{g,\gamma}(\lambda D^1 + (1-\lambda)D^2).
\end{aligned}$$

The proof of (2.53) is similar. This conclude the proof.

□

As an immediate consequence of Proposition 2.4.4 and Proposition 2.4.6 we deduce the following result about market impact on acceptability ask and bid prices. Namely we show that the acceptability bid and ask prices may not be homogenous in number of shares traded - larger number of shares one trades, a higher price per share it will cost.

Corollary 2.4.2. *Assume that NGD holds for $\mathcal{H}(t)$ at level $\gamma > 0$, $t \in \mathcal{T}$. Then, the acceptability ask and bid prices satisfy the following inequalities*

$$\begin{aligned}
&\widehat{a}_t^{g,\gamma}(\lambda\varphi, D) \leq \lambda\widehat{a}_t^{g,\gamma}(\varphi, D),\ \widehat{b}_t^{g,\gamma}(\lambda\varphi, D) \geq \lambda\widehat{b}_t^{g,\gamma}(\varphi, D),\ \lambda,\varphi \in L_+^\infty(\mathcal{F}_t),\ 0 \leq \lambda \leq 1; \\
&\widehat{a}_t^{g,\gamma}(\lambda\varphi, D) \geq \lambda\widehat{a}_t^{g,\gamma}(\varphi, D),\ \widehat{b}_t^{g,\gamma}(\lambda\varphi, D) \leq \lambda\widehat{b}_t^{g,\gamma}(\varphi, D),\ \lambda,\varphi \in L_+^\infty(\mathcal{F}_t),\ \lambda \geq 1,
\end{aligned}$$

for $\gamma > 0$, $t \in \mathcal{T}$.

Similar to the discussion in Section 2.3.4, we study the relationship between the acceptability ask and bid prices in the case when these prices are generated by different families of drivers and and different acceptability levels. We start with a result analogous to Proposition 2.3.3.

Proposition 2.4.7. *The acceptability ask and bid prices at time* $t \in \mathcal{T}$ *of a cash flow* $D \in \mathcal{D}$*, satisfy the following inequalities* $\widehat{a}_t^{g,\gamma_1}(D) \leq \widehat{a}_t^{g,\gamma_2}(D)$ *and* $\widehat{b}_t^{g,\gamma_1}(D) \geq \widehat{b}_t^{g,\gamma_2}(D)$*, for* $\gamma_2 \geq \gamma_1 > 0$.

Proof. Since $g_{\gamma_1} \leq g_{\gamma_2}$, according to Theorem 2.1.2, we get that

$$\mathcal{E}_{g_{\gamma_1}}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big] \leq \mathcal{E}_{g_{\gamma_2}}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big],$$

for any $H \in \mathcal{H}(t)$. Hence, using the representation (2.45), we obtain

$$\begin{aligned}
\widehat{a}_t^{g,\gamma_1}(D) &= \operatorname*{ess\,inf}_{H\in\mathcal{H}(t)} \mathcal{E}_{g_{\gamma_1}}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big] \\
&\leq \operatorname*{ess\,inf}_{H\in\mathcal{H}(t)} \mathcal{E}_{g_{\gamma_2}}\Big[\sum_{s=t+1}^{T}(D_s - H_s)\Big|\mathcal{F}_t\Big] \\
&\leq \widehat{a}_t^{g,\gamma_2}(D).
\end{aligned}$$

Analogously, one proofs the corresponding inequality for the bid prices. □

Corollary 2.4.3. *Assume that NGD holds for* $\mathcal{H}(t)$ *at time* $t \in \mathcal{T}$*, and at levels* $\gamma_1, \gamma_2 > 0$*. Then,* $\widehat{a}_t^{g,\gamma_2}(D) \geq \widehat{b}_t^{g,\gamma_1}(D)$*, for any* $D \in \mathcal{D}$.

Proof. If $\gamma_1 \geq \gamma_2$, then by Proposition 2.4.3 and 2.4.7, we have that $\widehat{a}_t^{g,\gamma_2}(D) \geq \widehat{b}_t^{g,\gamma_2}(D) \geq \widehat{b}_t^{g,\gamma_1}(D)$. Similarly, if $\gamma_1 \leq \gamma_2$, then $\widehat{a}_t^{g,\gamma_2}(D) \geq \widehat{a}_t^{g,\gamma_1}(D) \geq \widehat{b}_t^{g,\gamma_1}(D)$. This completes the proof. □

Finally we want to mention that we were not able to establish a general result on comparison of acceptability bid and ask prices, similar to Proposition 2.3.2. Generally speaking we do not know if $\widehat{a}^{g^1,\gamma_1} \geq \widehat{b}^{g^2,\gamma_2}$, for some arbitrary family of drivers

g^1, g^2, and levels $\gamma_1, \gamma_2 > 0$. In the nutshell, this is due to the lack of an appropriate form of time consistency property for $\widehat{a}^{g,\gamma}$, and $\widehat{b}^{g,\gamma}$, similar to Property P5, in Theorem 2.3.1. We leave the answer to this question for further investigations. Nevertheless, we do have a result that shows that once two couterparties, who may use different acceptability levels and different drivers, find that their prices are such that $\widehat{a}_t^{g^1,\gamma_1}(D) \leq \widehat{b}_t^{g^2,\gamma_2}(D)$ for all $D \in \mathcal{D}$, then the bid and ask prices coincide and hence the trade will go through.

Proposition 2.4.8. *Fix $t \in \mathcal{T}$. Let g^1 and g^2 be two families of drivers. Assume that $\widehat{a}_t^{g^1,\gamma_1}(D) \leq \widehat{b}_t^{g^2,\gamma_2}(D)$, and for any $D \in \mathcal{D}$. Then,*

$$\widehat{a}_t^{g^1,\gamma_1}(D) = \widehat{b}_t^{g^1,\gamma_1}(D) = \widehat{a}_t^{g^2,\gamma_2}(D) = \widehat{b}_t^{g^2,\gamma_2}(D),$$

for any $D \in \mathcal{D}$.

Proof. Since $\widehat{a}_t^{g^1,\gamma_1}(D) \leq \widehat{b}_t^{g^2,\gamma_2}(D)$ for any $D \in \mathcal{D}$, then we have that

$$\widehat{a}_t^{g^1,\gamma_1}(-D) \leq \widehat{b}_t^{g^2,\gamma_2}(-D), \qquad D \in \mathcal{D}.$$

According to Proposition 2.4.1, it is clear that

$$\begin{aligned} \widehat{a}_t^{g^1,\gamma_1}(D) &= -\widehat{b}_t^{g^1,\gamma_1}(-D), \\ \widehat{a}_t^{g^2,\gamma_2}(D) &= -\widehat{b}_t^{g^2,\gamma_2}(-D). \end{aligned}$$

Therefore, $-\widehat{b}_t^{g^1,\gamma_1}(D) = \widehat{a}_t^{g^1,\gamma_1}(-D) \leq \widehat{b}_t^{g^2,\gamma_2}(-D) = -\widehat{a}_t^{g^2,\gamma_2}(D)$, for any $D \in \mathcal{D}$. Hence, $\widehat{a}_t^{g^2,\gamma_2}(D) \leq \widehat{b}_t^{g^1,\gamma_1}(D)$, and due to Proposition 2.4.3, we get that

$$\widehat{b}_t^{g^2,\gamma_2}(D) \leq \widehat{a}_t^{g^2,\gamma_2}(D) \leq \widehat{b}_t^{g^1,\gamma_1}(D) \leq \widehat{a}_t^{g^1,\gamma_1}(D), \qquad D \in \mathcal{D}.$$

Thus, using our initial assumptions, we have that

$$\widehat{b}_t^{g^2,\gamma_2}(D) = \widehat{a}_t^{g^2,\gamma_2}(D) = \widehat{b}_t^{g^1,\gamma_1}(D) = \widehat{a}_t^{g^1,\gamma_1}(D), \qquad D \in \mathcal{D}.$$

This concludes the proof. □

CHAPTER 3

RECURSIVE CONSTRUCTION OF CONFIDENCE REGIONS

3.1 Preliminaries

Let $(\Omega, \mathcal{F})$ be a measurable space, and $\mathbf{\Theta} \subset \mathbb{R}^d$ be a non-empty set, which will play the role of the parameter space throughout.[6] On the space $(\Omega, \mathcal{F})$ we consider a discrete time, real valued random process $Z = \{Z_n,\ n \geq 0\}$.[7] We postulate that this process is observed, and we denote by $\mathbb{F} = (\mathcal{F}_n, n \geq 0)$ its natural filtration. The (true) law of Z is unknown, and assumed to belong to a parameterized family of probability distributions on $(\Omega, \mathcal{F})$, say $\{\mathbb{P}_\theta, \theta \in \mathbf{\Theta}\}$. It will be convenient to consider $(\Omega, \mathcal{F})$ to be the canonical space for Z, and to consider Z to be the canonical process (see Appendix A for details). Consequently, the law of Z under $\mathbb{P}_\theta$ is the same as $\mathbb{P}_\theta$. The (true) law of Z will be denoted by $\mathbb{P}_{\theta^*}$; accordingly, $\theta^* \in \mathbf{\Theta}$ is the (unknown) true parameter.

The set of probabilistic models that we are concerned with is $\{(\Omega, \mathcal{F}, \mathbb{F}, Z, \mathbb{P}_\theta),\ \theta \in \mathbf{\Theta}\}$. The model uncertainty addressed in this work occurs if $\mathbf{\Theta} \neq \{\theta^*\}$, which we assume to be the case. Our objective is to provide a recursive construction of confidence regions for θ^*, based on accurate observations of realizations of process Z through time, and satisfying desirable asymptotic properties.

In what follows, all equalities and inequalities between random variables will be understood in $\mathbb{P}_{\theta^*}$ almost surely sense. We denote by $\mathbb{E}_{\theta^*}$ the expectation operator corresponding to probability $\mathbb{P}_{\theta^*}$.

---

[6]In general, the parameter space may be infinite dimensional, consisting for example of dynamic factors, such as deterministic functions of time or hidden Markov chains. In this study, for simplicity, we chose the parameter space to be a subset of $\mathbb{R}^d$.

[7]The study presented in this thesis extends to the case when process Z takes values in $\mathbb{R}^d$, for $d > 1$. We focus here the case of $d = 1$ for simplicity of presentation.

We impose the the following structural standing assumption.

Assumption M:

(i) Process Z is a time homogenous Markov process under any $\mathbb{P}_\theta,\ \theta \in \mathbf{\Theta}$.

(ii) Process Z is an ergodic Markov process under $\mathbb{P}_{\theta^*}$.[8]

(iii) The transition kernel of process Z under any $\mathbb{P}_\theta,\ \theta \in \mathbf{\Theta}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, that is, for any Borel subset of $\mathbb{R}$

$$\mathbb{P}_\theta(Z_1 \in A \mid Z_0 = x) = \int_A p_\theta(x,y)dy,$$

for some positive and measurable function p_θ.[9]

For any $\theta \in \mathbf{\Theta}$ and $n \geq 1$, we define $\pi_n(\theta) := \log p_\theta(Z_{n-1}, Z_n)$.

Remark 3.1.1. *In view of the Remark A.0.2, the process Z is a stationary process under $\mathbb{P}_{\theta^*}$. Consequently, under $\mathbb{P}_{\theta^*}$, for each $\theta \in \mathbf{\Theta}$ and for each $n \geq 0$, the law of $\pi_n(\theta)$ is the same as the law of $\pi_1(\theta)$.*

We will need to impose several technical assumptions in what follows. We begin with the assumption

R0. For any $\theta \in \mathbf{\Theta}$, $\pi_1(\theta)$ is integrable under $\mathbb{P}_{\theta^*}$.

Then, we have the following result.

Proposition 3.1.1. *Assume that M and R0 hold. Then,*

[8]See Appendix A for the definition of ergodicity that we postulate here.

[9]This postulate is made solely in order to streamline the presentation. In general, our methodology works for Markov processes for which the transition kernel is not absolutely continuous with respect to the Lebesgue.

(i) For any $\theta \in \Theta$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \pi_i(\theta) = \mathbb{E}_{\theta^*}[\pi_1(\theta)].$$

(ii) Moreover, for any $\theta \in \Theta$,

$$\mathbb{E}_{\theta^*}[\pi_1(\theta^*)] \geq \mathbb{E}_{\theta^*}[\pi_1(\theta)].$$

Proof. Fix $\theta \in \Theta$, since Z is ergodic under $\mathbb{P}_{\theta^*}$ and $\mathbb{E}_{\theta^*}[\pi_1(\theta)] < \infty$, then according to Proposition A.0.1 we have that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \pi_i(\theta) = \mathbb{E}_{\theta^*}[\pi_1(\theta)]$$

which proves (i).

Now we prove that (ii) holds. In fact, denote by f_{Z_1} the density function of Z_1 under $\mathbb{P}_{\theta^*}$, we have that

$$\begin{aligned}
&\mathbb{E}_{\theta^*}[\pi_1(\theta)] - \mathbb{E}_{\theta^*}[\pi_1(\theta^*)] \\
=&\mathbb{E}_{\theta^*}\Big[\log \frac{p_\theta(Z_1, Z_2)}{p_{\theta^*}(Z_1, Z_2)}\Big] = \int_{\mathbb{R}} \mathbb{E}_{\theta^*}\Big[\log \frac{p_\theta(Z_1, Z_2)}{p_{\theta^*}(Z_1, Z_2)}\Big| Z_1 = z_1\Big] f_{Z_1}(z_1) dz_1 \\
\leq& \int_{\mathbb{R}} \Big[\log \mathbb{E}_{\theta^*}\Big[\frac{p_\theta(Z_1, Z_2)}{p_{\theta^*}(Z_1, Z_2)}\Big| Z_1 = z_1\Big] f_{Z_1}(z_1) dz_1 \\
=& \int_{\mathbb{R}} \log \int_{\mathbb{R}} \frac{p_\theta(z_1, z_2)}{p_{\theta^*}(z_1, z_2)} p_{\theta^*}(z_1, z_2) dz_2 f_{Z_1}(z_1) dz_1 \\
=& \int_{\mathbb{R}} \log \int_{\mathbb{R}} p_\theta(z_1, z_2) dz_2 f_{Z_1}(z_1) dz_1 = 0,
\end{aligned}$$

where the inequality holds due to Jensen's inequality. □

In the statement of the technical assumptions R1-R8 below we use the notations

$$\psi_n(\theta) = \nabla \pi_n(\theta), \quad \Psi_n(\theta) = \mathsf{H}\pi_n(\theta), \quad \beta_n(\theta) = \mathbb{E}_{\theta^*}[\psi_n(\theta)|\mathcal{F}_{n-1}], \tag{3.1}$$

where ∇ denotes the gradient vector and H denotes the Hessian matrix with respect to θ, respectively.

R1. For each $x, y \in \mathbb{R}$ the function $p_{\cdot}(x,y) : \mathbf{\Theta} \to \mathbb{R}_+$ is three times differentiable, and

$$\nabla \int_{\mathbb{R}} p_\theta(x,y)dy = \int_{\mathbb{R}} \nabla p_\theta(x,y)dy, \quad \mathsf{H} \int_{\mathbb{R}} p_\theta(x,y)dy = \int_{\mathbb{R}} \mathsf{H} p_\theta(x,y)dy. \tag{3.2}$$

R2. For any $\theta \in \mathbf{\Theta}$, $\psi_1(\theta)$ and $\Psi_1(\theta)$ are integrable under $\mathbb{P}_{\theta^*}$. The function $\mathbb{E}_{\theta^*}[\pi_1(\,\cdot\,)]$ is twice differentiable in θ, and

$$\nabla \mathbb{E}_{\theta^*}[\pi_1(\theta)] = \mathbb{E}_{\theta^*}[\psi_1(\theta)], \quad \mathsf{H}\mathbb{E}_{\theta^*}[\pi_1(\theta)] = \mathbb{E}_{\theta^*}[\Psi_1(\theta)].$$

R3. There exists a unique $\theta \in \mathbf{\Theta}$ such that

$$\mathbb{E}_{\theta^*}[\psi_1(\theta)] = 0.$$

R4. There exists a constant $c > 0$ such that, for any $n \geq 1$ and $\theta \in \mathbf{\Theta}$,

$$\mathbb{E}_{\theta^*}[\|\psi_n(\theta)\|^2 \mid \mathcal{F}_{n-1}] \leq c(1 + \|\theta - \theta^*\|^2). \tag{3.3}$$

R5. There exist some positive constants $K_i, i = 1, 2, 3$, such that for any $\theta, \theta_1, \theta_2 \in \mathbf{\Theta}$, and $n \geq 1$,[10]

$$(\theta - \theta^*)^T b_n(\theta) \leq -K_1 \|\theta - \theta^*\|^2, \tag{3.4}$$

$$\|\beta_n(\theta_1) - \beta_n(\theta_2)\| \leq K_2 \|\theta_1 - \theta_2\|, \tag{3.5}$$

$$\mathbb{E}_{\theta^*}[\|\Psi_n(\theta_1) - \Psi_n(\theta_2)\| \mid \mathcal{F}_{n-1}] \leq K_3 \|\theta_1 - \theta_2\|. \tag{3.6}$$

R6. There exists a positive constant K_4, such that for any $\theta \in \mathbf{\Theta}$, and $n \geq 1$,

$$\mathbb{E}_{\theta^*}[\|\mathsf{H}\psi_n(\theta)\| | \mathcal{F}_{n-1}] \leq K_4. \tag{3.7}$$

R7. For any $n \geq 1$,

$$\sup_{\theta \in \mathbf{\Theta}} \mathbb{E}_{\theta^*} \|\psi_n(\theta) - \beta_n(\theta)\|^2 < \infty. \tag{3.8}$$

---

[10]Superscript T will denote the transpose.

R8. For each $\theta \in \mathbf{\Theta}$ the Fisher information matrix

$$I(\theta) := \mathbb{E}_\theta[\psi_1(\theta)\psi_1^T(\theta)]$$

exists and is positive definite. Moreover, $I(\theta)$ is continuous with respect to θ.

R9.

$$\lim_{n\to\infty} \mathbb{E}_{\theta^*}\Big[\sup_{0\le i\le n}\Big|\frac{1}{\sqrt{n}}\psi_i(\theta^*)\Big|\Big] = 0. \tag{3.9}$$

Remark 3.1.2. *(i) Note that in view of the Remark 3.1.1 properties assumed in R2, R3, and R8 imply that analogous properties hold with time n in place of time 1.*
(ii) According to Proposition C.0.6, we have that if R4-R6 hold, then (3.3)-(3.7) are also satisfied for any $\mathcal{F}_{n-1}$-measurable random vector $\boldsymbol{\theta} \in \mathbf{\Theta}$.

As stated above, our aim is to provide a recursive construction of the confidence regions for θ^*. In the sequel, we will propose a method for achieving this goal that will be derived from a suitable recursive point estimator of θ^*. Note that due to Proposition 3.1.1 (ii) and Assumption R3, we have that θ^* is the unique solution of

$$\mathbb{E}_{\theta^*}[\psi_1(\theta)] = 0. \tag{3.10}$$

Therefore, point-estimating θ^* is equivalent to point-estimating the solution of the equation (3.10). Since θ^* is unknown, equation (3.10) is not really known to us. We will therefore apply an appropriate version of the so called *stochastic approximation* method, which is a recursive method used to point-estimate zeros of functions that can not be directly observed. This can be done in our set-up since, thanks to Proposition 3.1.1 (i), we are provided with a sequence of observed random variables $\frac{1}{n}\sum_{i=1}^n \psi_i(\theta)$ that $\mathbb{P}_{\theta^*}$ almost surely converges to $\mathbb{E}_{\theta^*}[\psi_1(\theta)]$ – a property, which will enable us to adopt the method of stochastic approximation. Accordingly, in the next two sections, we will introduce two recursive point estimators of θ^*, and we will derive properties of these estimators that are relevant for us.

3.2 $\sqrt{n}$-consistent base point estimator

In this section we consider a recursive point estimator $\tilde{\theta} = \{\tilde{\theta}_n,\, n \geq 1\}$ of θ^*, that will be defined in (3.11). Towards this end, we fix a positive constant η such that $\eta K_1 > \frac{1}{2}$, where K_1 was introduced in Assumption R6, and we define the process $\tilde{\theta}$ recursively as follows,

$$\tilde{\theta}_n = \tilde{\theta}_{n-1} + \frac{\eta}{n}\psi_n(\tilde{\theta}_{n-1}), \quad n \geq 1, \tag{3.11}$$

with the initial guess $\tilde{\theta}_0$ being an element in $\boldsymbol{\Theta}$, where ψ_n was defined in (3.1).

Given the definition of ψ_n, we see that $\tilde{\theta}_n$ is updated from $\tilde{\theta}_{n-1}$ based on new observation Z_n available at time n; of course, Z_{n-1} is used as well. We note that the recursion (3.11) is a version of the stochastic approximation method, which is meant to recursively approximate roots of the unknown equations, such as equation (3.10) (see e.g. [RM51], [KW52], [LS87], [KC78], [KY03]).

Remark 3.2.1. *It is implicitly assumed in the recursion* (3.11) *that* $\tilde{\theta}_n \in \boldsymbol{\Theta}$*. One typical and easy way of making sure that this happens is to choose* $\boldsymbol{\Theta}$ *as the "largest possible set" that* θ^* *is an element of. So typically, one takes* $\boldsymbol{\Theta} = \mathbb{R}^d$*. However, this is not always possible, in which case one needs to implement a version of constrained stochastic approximation method (cf. e.g. [KC78] or [BK02]). We are not considering constrained stochastic approximation in this chapter. This is planned for a future work.*

Remark 3.2.2. *As we will see later, the requirement* $\eta K_1 > \frac{1}{2}$ *guarantees that* $\tilde{\theta}$ *converges at a rate of* $\frac{1}{\sqrt{n}}$ *in probability. Other than that, the importance of choice of* η *is neglectable. Hence, a simple choice of* η *would be* $\lceil \frac{1}{2K_1} \rceil + 1$ *where* $\lceil \cdot \rceil$ *denotes the ceiling function.*

As mentioned above, we are interested in the study of asymptotic properties of confidence regions that we will construct recursively in Section 3.4. These asymptotic properties crucially depend on the asymptotic properties of our recursive (point) estimators. One of such required properties is asymptotic normality. In this regard we stress that although the theory of asymptotic normality for stochastic approximation estimators is quite a mature field (see e.g. [Sac58], [Fab68], [LR79]), the existing results do not apply to $\tilde{\theta}$ as they require $\psi_n(\tilde{\theta}_{n-1}) - \mathbb{E}_{\theta^*}[\psi_n(\tilde{\theta}_{n-1})]$ to be a martingale, the property, which is not satisfied in our set-up. Thus, we need to modify the base estimator $\tilde{\theta}$ to the effect of producing a recursive estimator that is asymptotically normal. In the next section we will construct such estimator, denoted there as $\hat{\theta}$, and we will study its asymptotic properties in the spirit of the method proposed by Fisher [Fis25]. Motivated by finding estimators that share the same asymptotic property as maximum likelihood estimators (MLEs), Fisher proposed in [Fis25] that if an estimator is $\sqrt{n}$-consistent (see below), then appropriate modification of the estimator has the same asymptotic normality as the MLE. This subject was further studied by LeCam in [LeC56] and [LeC60], where a more general class of observation than i.i.d. observations are considered.

Accordingly, we will show that $\tilde{\theta}$ is strongly consistent, and, moreover it maintains $\sqrt{n}$ convergence rate, i.e.

$$\mathbb{E}_{\theta^*}\|\tilde{\theta}_n - \theta^*\|^2 = O(n^{-1}). \tag{3.12}$$

An estimator that satisfies this equality is said to be $\sqrt{n}$-*consistent.*

We begin with the following proposition, which shows that the estimator $\tilde{\theta}$ is strongly consistent. For convenience, throughout, we will use the notation $\Delta_n := \tilde{\theta}_n - \theta^*$, $n \geq 1$.

Proposition 3.2.1. *Assume that* (3.3)*, and* (3.4) *are satisfied, then*

$$\lim_{n\to\infty} \tilde{\theta}_n = \theta^*, \quad \mathbb{P}_{\theta^*} - a.s.$$

Proof. Let us fix $n \geq 1$. Clearly, $\Delta_n = \Delta_{n-1} + \frac{\eta}{n}\psi_n(\theta^* + \Delta_{n-1})$, so that

$$\|\Delta_n\|^2 = \|\Delta_{n-1}\|^2 + \frac{2\eta}{n}\Delta_{n-1}^T\psi_n(\theta^* + \Delta_{n-1}) + \frac{\eta^2}{n^2}\|\psi_n(\theta^* + \Delta_{n-1})\|^2.$$

Taking conditional expectation on both sides leads to

$$\mathbb{E}_{\theta^*}[\|\Delta_n\|^2|\mathcal{F}_{n-1}] \leq \|\Delta_{n-1}\|^2 + \frac{2\eta}{n}\Delta_{n-1}^T\beta_n(\theta^* + \Delta_{n-1}) + \frac{c\eta^2}{n^2}(1 + \|\Delta_{n-1}\|^2) \quad (3.13)$$

$$\leq \|\Delta_{n-1}\|^2 + \frac{c\eta^2}{n^2}(1 + \|\Delta_{n-1}\|^2), \quad (3.14)$$

where the first inequality comes from (3.3) and the second is implied by (3.4). Let

$$\mathbf{Y}_m := \|\Delta_m\|^2 \prod_{k=m+1}^{\infty}(1 + \frac{c\eta^2}{k^2}) + \sum_{k=m+1}^{\infty}\frac{c\eta^2}{k^2}\prod_{j=k+1}^{\infty}(1 + \frac{c\eta^2}{j^2}), \quad m \geq 0.$$

Then, (3.14) yields that

$$\mathbb{E}_{\theta^*}[\mathbf{Y}_{m+1}|\mathcal{F}_m] \leq \mathbf{Y}_m, \quad m \geq 0,$$

and therefore process $\mathbf{Y}$ is a supermartingale. Noting that $\mathbf{Y}$ is a positive process, and invoking the supermartingale convergence theorem, we conclude that hence the sequence $\{\mathbf{Y}_m, m \geq 0\}$ converges $\mathbb{P}_{\theta^*}$ almost surely. This implies that the sequence $\{\|\Delta_m\|, m \geq 0\}$ converges, and we will show now that its limit is zero. According to (3.13), we have

$$\mathbb{E}_{\theta^*}\|\Delta_m\|^2 \leq \mathbb{E}_{\theta^*}\|\Delta_{m-1}\|^2 + \frac{2\eta}{m}\mathbb{E}_{\theta^*}\left[\Delta_{m-1}^T\beta_m(\theta^* + \Delta_{m-1})\right] + \frac{c\eta^2}{m^2}\mathbb{E}_{\theta^*}\left[1 + \|\Delta_{m-1}\|^2\right]$$

$$\leq \mathbb{E}_{\theta^*}\|\Delta_1\|^2 + \sum_{k=1}^{m}\frac{2\eta}{k}\mathbb{E}_{\theta^*}\left[\Delta_{k-1}^T\beta_k(\theta^* + \Delta_{k-1})\right] + \sum_{k=1}^{m}\frac{c\eta^2}{k^2}\mathbb{E}_{\theta^*}\left[1 + \|\Delta_{k-1}\|^2\right].$$

Hence, we get

$$\sum_{k=1}^{m}\frac{2\eta}{k}\mathbb{E}_{\theta^*}\left|\Delta_{k-1}^T\beta_k(\theta^* + \Delta_{k-1})\right| \leq \mathbb{E}_{\theta^*}\|\Delta_1\|^2 - \mathbb{E}_{\theta^*}\|\Delta_m\|^2 + \sum_{k=1}^{m}\frac{c\eta^2}{k^2}\mathbb{E}_{\theta^*}\left[1 + \|\Delta_{k-1}\|^2\right].$$

Since

$$\lim_{m\to\infty}\|\Delta_m\|^2 = \lim_{m\to\infty}\mathbf{Y}_m < \infty,$$

and

$$\lim_{m\to\infty}\sum_{k=1}^{m}\frac{c\eta^2}{k^2}\mathbb{E}_{\theta^*}\left[1+\|\Delta_{k-1}\|^2\right]\le\sum_{k=1}^{\infty}\frac{c\eta\mathbb{E}_{\theta^*}[1+\mathbf{Y}_1]}{k^2}<\infty,$$

then, the series

$$\sum_{k=1}^{m}\frac{1}{k}\mathbb{E}_{\theta^*}\left|\Delta_{k-1}^T\beta_k(\theta^*+\Delta_{k-1})\right|,\quad m\ge1,$$

converges $\mathbb{P}_{\theta^*}$ almost surely, and thus

$$\lim_{k\to\infty}\mathbb{E}_{\theta^*}\left|\Delta_{k-1}^T\beta_k(\theta^*+\Delta_{k-1})\right|=0.$$

This implies that there exists a subsequence $\Delta_{m_k-1}^T b_{m_k}(\theta^*+\Delta_{m_k-1})$ which converges $\mathbb{P}_{\theta^*}$ almost surely to zero, as $k\to\infty$. According to (3.4), we also have that

$$\|\Delta_{m_k-1}\|^2\le\frac{1}{K_1}\|\Delta_{m_k-1}^T\beta_{m_k}(\theta^*+\Delta_{m_k-1})\|.$$

Therefore, $\lim_{k\to\infty}\Delta_{m_k-1}=0$, $\mathbb{P}_{\theta^*}$ almost surely, and this concludes the proof. □

Proposition 3.2.2. *Assume that* (3.3)*,* (3.4)*,* (3.5) *and* (3.8) *hold. Then,*

$$\mathbb{E}_{\theta^*}\|\tilde{\theta}_n-\theta^*\|^2=O(n^{-1}).$$

Proof. Putting $V_n(\tilde{\theta}_{n-1}):=\psi_n(\tilde{\theta}_{n-1})-\beta_n(\tilde{\theta}_{n-1})$, from (3.11) we immediately have that

$$\Delta_n=\Delta_{n-1}+\frac{\eta}{n}\beta_n(\tilde{\theta}_{n-1})+\frac{\eta}{n}V_n(\tilde{\theta}_{n-1}),$$

that yields

$$\mathbb{E}_{\theta^*}\|\Delta_n\|^2=\mathbb{E}_{\theta^*}\|\Delta_{n-1}+\frac{\eta}{n}\beta_n(\tilde{\theta}_{n-1})\|^2+\frac{\eta^2}{n^2}\mathbb{E}_{\theta^*}\|V_n(\tilde{\theta}_{n-1})\|^2.$$

From here, applying consequently (3.8), (3.5), (3.4), and note that $\beta_n(\theta^*)=0$, we get

$$\begin{aligned}\mathbb{E}_{\theta^*}\|\Delta_n\|^2&=\mathbb{E}_{\theta^*}\left\|\Delta_{n-1}+\frac{\eta}{n}\beta_n(\tilde{\theta}_{n-1})\right\|^2+O(n^{-2})\\&\le\mathbb{E}_{\theta^*}\left[\|\Delta_{n-1}\|^2+\frac{\eta^2K_2^2}{n^2}\|\Delta_{n-1}\|^2+\frac{2\eta}{n}\Delta_{n-1}^T\beta_n(\tilde{\theta}_{n-1})\right]+O(n^{-2})\\&\le\left(1+\frac{\eta^2K_2^2}{n^2}-\frac{2\eta K_1}{n}\right)\mathbb{E}_{\theta^*}\|\Delta_{n-1}\|^2+D_1n^{-2}.\end{aligned}$$

Clearly, for any $\varepsilon > 0$, and for large enough n, we get

$$\mathbb{E}_{\theta^*}\|\Delta_n\|^2 \leq (1-(2K_1\eta-\varepsilon)n^{-1})\mathbb{E}_{\theta^*}\|\Delta_{n-1}\|^2 + D_1 n^{-2}. \tag{3.15}$$

For ease of writing, we put $p := 2K_1\eta - \varepsilon$ and $c_n := \mathbb{E}_{\theta^*}\|\Delta_n\|^2$. Take ϵ sufficiently small, so that $p > 1$, and then chose an integer $N > p$. Then, for $n > N$ we have by (3.15) that

$$\begin{aligned} c_n &\leq c_N \prod_{j=N+1}^{n}(1-\frac{p}{j}) + D_1 \sum_{j=N+1}^{n} \frac{1}{j^2} \prod_{k=j+1}^{n}(1-\frac{p}{k}) \\ &\leq c_N \prod_{j=N+1}^{n}(1-\frac{p}{j}) + D_1 \sum_{j=N+1}^{n} \frac{1}{j^2}. \end{aligned}$$

Using the fact that $\sum_{j=m}^{n} 1/j^2 \sim 1/n$ and $\prod_{j=m}^{n}(1-p/j) \sim 1/n^p$, for any fixed $m, p \geq 1$, we immediately get that $c_n \leq O(n^{-1})$. This concludes the proof. □

3.3 Quasi-asymptotically linear estimator

In this section we define a new estimator denoted as $\{\hat{\theta}_n, n \geq 1\}$ and given recursively by

$$\begin{aligned} \hat{\theta}_n &= -I^{-1}(\tilde{\theta}_n)I_n\tilde{\theta}_n + I^{-1}(\tilde{\theta}_n)\Gamma_n, \\ \Gamma_n &= \frac{n-1}{n}\Gamma_{n-1} + \frac{1}{n}(\mathrm{Id}+\eta I_n)\psi_n(\tilde{\theta}_{n-1}), \\ I_n &= \frac{n-1}{n}I_{n-1} + \frac{1}{n}\Psi_n(\tilde{\theta}_{n-1}), \quad n \geq 1, \\ \Gamma_0 &= 0, \quad I_0 = 0, \end{aligned} \tag{3.16}$$

where Id is the unit matrix. Since $\tilde{\theta}_n$, I_n, and Γ_n are updated from time $n-1$ based on the new observation Z_n available at time n, then the estimator $\hat{\theta}$ indeed is recursive. This estimator will be used in Section 6 for recursive construction of confidence regions for θ^*.

Remark 3.3.1. *In the argument below we will use the following representations of Γ_n and I_n,*

$$\Gamma_n = \sum_{j=1}^{n}(\mathrm{Id}+\eta I_j)\,\psi_j(\tilde{\theta}_{j-1}), \quad I_n = \frac{1}{n}\sum_{i=1}^{n}\Psi_i(\tilde{\theta}_{i-1}).$$

Next, we will show that $\hat{\theta}$ is weakly consistent and asymptotically normal. We will derive asymptotic normality of $\hat{\theta}$ from the property of quasi-asymptotic linearity, which is related to the property of asymptotic linearity (cf. [Shi84]), and which is defined as follows:

Definition 3.3.1. *An estimator $\{\bar{\theta}_n, n \geq 1\}$ of θ^* is called a quasi-asymptotically linear estimator if there exist a $\mathbb{P}_{\theta^*}$-convergent, adapted matrix valued process G, and adapted vector valued processes ϑ and ε, such that*

$$\bar{\theta}_n - \vartheta_n = \frac{G_n}{n}\sum_{i=1}^{n}\psi_i(\theta^*) + \varepsilon_n,\ n \geq 1, \quad \vartheta_n \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} \theta^*, \quad \sqrt{n}\varepsilon_n \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} 0.$$

Our definition of quasi-asymptotically linear estimator is motivated by the classic concept of asymptotically linear estimator (see e.g. [Sha10]): $\check{\theta}$ is called (locally) asymptotically linear if there exists a matrix process $\{\check{G}_n, n \geq 1\}$ such that

$$\check{\theta}_n - \theta^* = \check{G}_n \sum_{i=1}^{n}\psi_i(\theta^*) + \varepsilon_n,$$

where $\check{G}_n^{-1/2}\varepsilon_n \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} 0$. Asymptotic linearity is frequently used in the proof of asymptotic normality of estimators. However, in general, asymptotic linearity can not be reconciled with the full recursiveness of the estimator. The latter property is the key property involved in construction of recursive confidence regions. Moreover, the property of asymptotic linearity requires that the matrices $\check{G}_n$ are invertible, which is a very stringent requirement, not easily fulfilled. These are the reasons why we propose the concept of quasi-asymptotic linearity since, it can be reconciled with recursiveness and does not require that matrices G_n are invertible. As it will be shown below, the fully recursive estimator $\hat{\theta}$ is quasi-asymptotically linear.

In what follows, we will make use of the following representation for $\hat{\theta}$

$$\hat{\theta}_n = -I^{-1}(\tilde{\theta}_n)I_n\tilde{\theta}_n + \frac{1}{n}I^{-1}(\tilde{\theta}_n)\sum_{j=1}^{n}(\mathrm{Id} + \eta I_j)\,\psi_j(\tilde{\theta}_{j-1}). \tag{3.17}$$

Theorem 3.3.1. *Assume that R1–R8 hold, then the estimator $\hat{\theta}$ is $\mathbb{P}_{\theta^*}$–weakly consistent.*[11]
Moreover, $\hat{\theta}$ is quasi-asymptotically linear estimator for θ^*.

Proof. First, we show the generalized asymptotic linearity of $\hat{\theta}$. Due to Taylor's expansion, we have that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\psi_i(\theta^*) - \frac{1}{n}\sum_{i=1}^{n}\psi_i(\tilde{\theta}_{i-1}) &= -\frac{1}{n}\sum_{i=1}^{n}\Psi_i(\tilde{\theta}_{i-1})\Delta_{i-1} + \frac{1}{n}\sum_{i=1}^{n}\Delta_{i-1}^T \mathsf{H}\psi_i(\zeta_{i-1})\Delta_{i-1} \\ &=: A_n + B_n, \end{aligned} \tag{3.18}$$

where $\zeta_{i-1}, 1 \leq i \leq n$, is in a neighborhood of θ^* such that $\|\zeta_{i-1} - \theta^*\| \leq \|\tilde{\theta}_{i-1} - \theta^*\|$. Note that

$$\begin{aligned} A_n =& -\frac{1}{n}\sum_{i=1}^{n}\Psi_i(\tilde{\theta}_{i-1})\Big(\Delta_n - \sum_{j=i}^{n}\frac{\eta}{j}\psi_j(\tilde{\theta}_{j-1})\Big) \\ =& - I_n\Delta_n + \frac{\eta}{n}\sum_{i=1}^{n} I_i\psi_i(\tilde{\theta}_{i-1}), \end{aligned}$$

and by (3.18), we get

$$I_n\Delta_n = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{Id} + \eta I_i)\psi_i(\tilde{\theta}_{i-1}) - \frac{1}{n}\sum_{i=1}^{n}\psi_i(\theta^*) + B_n.$$

Therefore, using the representation (3.17), we immediately have

$$\hat{\theta}_n + I^{-1}(\tilde{\theta}_n)I_n\theta^* = \frac{I^{-1}(\tilde{\theta}_n)}{n}\sum_{i=1}^{n}\psi_i(\theta^*) - I^{-1}(\tilde{\theta}_n)B_n. \tag{3.19}$$

Next we will show that

$$\mathbb{P}_{\theta^*}\text{-}\lim_{n\to\infty} I_n = -I(\theta^*). \tag{3.20}$$

First, by (3.6), we deduce that

$$\mathbb{E}_{\theta^*}\Big[\frac{1}{n}\sum_{i=1}^{n}\|\Psi_i(\tilde{\theta}_{i-1}) - \Psi_i(\theta^*)\|\Big] \leq \frac{K_3}{n}\sum_{i=1}^{n}\mathbb{E}_{\theta^*}\|\Delta_{i-1}\|.$$

---

[11] That is, $\hat{\theta}_n \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} \theta^*$.

Due to Proposition 3.2.2, $\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}_{\theta^*}\|\Delta_{i-1}\| \le \frac{1}{n}\sum_{j=1}^{n} j^{-1/2} = O(n^{-1/2})$. Hence,

$$\frac{1}{n}\sum_{i=1}^{n}\|\Psi_i(\tilde{\theta}_{i-1}) - \Psi_i(\theta^*)\| \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} 0. \tag{3.21}$$

Therefore,

$$\mathbb{P}_{\theta^*}\text{-}\lim_{n\to\infty} I_n = \mathbb{P}_{\theta^*} - \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Psi_i(\tilde{\theta}_{i-1}) = \mathbb{P}_{\theta^*} - \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Psi_i(\theta^*). \tag{3.22}$$

Next, observe that in view of Proposition A.0.1 we get

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Psi_i(\theta^*) = \mathbb{E}_{\theta^*}[\Psi_1(\theta^*)] = \mathbb{E}_{\theta^*}[\mathsf{H}\pi_1(\theta^*)] = \mathbb{E}_{\theta^*}[\mathsf{H}\log p_{\theta^*}(Z_0, Z_1)].$$

Invoking the usual chain rule we obtain that

$$\begin{aligned}\mathsf{H}\log p_{\theta^*}(Z_0, Z_1) =& \frac{\mathsf{H}p_{\theta^*}(Z_0, Z_1)}{p_{\theta^*}(Z_0, Z_1)} - \frac{\nabla p_{\theta^*}(Z_0, Z_1)\nabla p_{\theta^*}(Z_0, Z_1)^T}{p^2_{\theta^*}(Z_0, Z_1)}\\ =& \frac{\mathsf{H}p_{\theta^*}(Z_0, Z_1)}{p_{\theta^*}(Z_0, Z_1)} - \psi_1(\theta^*)\psi_1^T(\theta^*),\end{aligned}$$

so that

$$\mathbb{E}_{\theta^*}[\mathsf{H}\log p_{\theta^*}(Z_0, Z_1)] = \mathbb{E}_{\theta^*}[\frac{\mathsf{H}p_{\theta^*}(Z_0, Z_1)}{p_{\theta^*}(Z_0, Z_1)}] - I(\theta^*).$$

We will now show that $\mathbb{E}_{\theta^*}\left[\frac{\mathsf{H}p_{\theta^*}(Z_0,Z_1)}{p_{\theta^*}(Z_0,Z_1)}\right] = 0$. In fact, denote by f_{Z_0} the density function of Z_0 under $\mathbb{P}_{\theta^*}$ and in view of (3.2), we have

$$\begin{aligned}\mathbb{E}_{\theta^*}\left[\frac{\mathsf{H}p_{\theta^*}(Z_0, Z_1)}{p_{\theta^*}(Z_0, Z_1)}\right] =& \mathbb{E}_{\theta^*}\left[\mathbb{E}_{\theta^*}\left[\frac{\mathsf{H}p_{\theta^*}(Z_0, Z_1)}{p_{\theta^*}(Z_0, Z_1)} \mid Z_0\right]\right]\\ =& \int_{\mathbb{R}}\mathbb{E}_{\theta^*}\left[\frac{\mathsf{H}p_{\theta^*}(Z_0, Z_1)}{p_{\theta^*}(Z_0, Z_1)} \mid Z_0 = z_0\right] f_{Z_0}(z_0)dz_0\\ =& \int_{\mathbb{R}}\int_{\mathbb{R}}\frac{\mathsf{H}p_{\theta^*}(z_0, z_1)}{p_{\theta^*}(z_0, z_1)}p_{\theta^*}(z_0, z_1)dz_1 f_{Z_0}(z_0)dz_0\\ =& \int_{\mathbb{R}}\int_{\mathbb{R}}\mathsf{H}p_{\theta^*}(z_0, z_1)dz_1 f_{Z_0}(z_0)dz_0\\ =& \int_{\mathbb{R}}\mathsf{H}\int_{\mathbb{R}}p_{\theta^*}(z_0, z_1)dz_1 f_{Z_0}(z_0)dz_0\\ =& \int_{\mathbb{R}}(\mathsf{H}1) f_{Z_0}(z_0)dz_0 = 0.\end{aligned}$$

Recalling (3.22) we conclude that (3.20) is satisfied.

By Assumption R8 and strong consistency of $\tilde{\theta}$ we obtain that

$$\lim_{n\to\infty} I^{-1}(\tilde{\theta}_n) = I^{-1}(\theta^*) \quad \mathbb{P}_{\theta^*} - a.s., \tag{3.23}$$

which, combined with (3.20) implies that

$$-I^{-1}(\tilde{\theta}_n)I_n\theta^* \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} \theta^*. \tag{3.24}$$

Next, we will show that

$$\sqrt{n}B_n \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} 0. \tag{3.25}$$

Indeed, by (3.7), $\sqrt{n}\mathbb{E}_{\theta^*}\|B_n\| \le \frac{K_4}{\sqrt{n}}\sum_{i=1}^n \mathbb{E}_{\theta^*}\|\Delta_{i-1}\|^2$, and consequently, in view of Proposition 3.2.2,

$$\lim_{n\to\infty} \sqrt{n}\mathbb{E}_{\theta^*}\|B_n\| \le \lim_{n\to\infty} \frac{K_4}{\sqrt{n}}\log n = 0,$$

which implies (3.25).

Now, taking $\vartheta_n = -I^{-1}(\tilde{\theta}_n)I_n\theta^*$, $G_n = I^{-1}(\tilde{\theta}_n)$ and $\varepsilon_n = I^{-1}(\tilde{\theta}_n)B_n$, we deduce quasi-asymptotic linearity of $\hat{\theta}$ from (3.19), (3.23), (3.24) and (3.25).

Finally, we will show the weak consistency of $\hat{\theta}$. By ergodicity of Z, in view of Proposition A.0.1, and using the fact that θ^* is a (unique) solution of (3.10), we have that

$$\frac{1}{n}\sum_{i=1}^n \psi_i(\theta^*) = \mathbb{E}_{\theta^*}[\psi_1(\theta^*)] = 0, \quad \mathbb{P}_{\theta^*} - \text{a.s.}$$

Thus, $\lim_{n\to\infty} \frac{I^{-1}(\tilde{\theta}_n)}{n}\sum_{i=1}^n \psi_i(\theta^*) = 0$ $\mathbb{P}_{\theta^*}$ almost surely. This, combined with (3.19), (3.24) and (3.25) implies that $\hat{\theta}_n \xrightarrow{\mathbb{P}_{\theta^*}} \theta^*$, as $n\to\infty$. The proof is complete. □

The next result, which will be used in analysis of asymptotic properties of the recursive confidence region for θ^* in Section 6, is an application of Theorem 3.3.1.

Proposition 3.3.1. *Assume that R1–R9 are satisfied. Then, there exists an adapted process* ϑ *such that*

$$\vartheta_n \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} \theta^*, \tag{3.26}$$

and

$$\sqrt{n}(\hat{\theta}_n - \vartheta_n) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, I^{-1}(\theta^*)). \tag{3.27}$$

Proof. Let $\vartheta_n = -I^{-1}(\tilde{\theta}_n)I_n\theta^*$, $G_n = I^{-1}(\tilde{\theta}_n)$ and $I^{-1}(\tilde{\theta}_n)B_n = \varepsilon_n$. Then, property (3.26) follows from (3.24).

In order to prove (3.27), we note that according to Theorem 3.3.1 we have

$$\hat{\theta}_n - \vartheta_n = \frac{G_n}{n}\sum_{i=1}^{n}\psi_i(\theta^*) + \varepsilon_n, \quad \sqrt{n}\varepsilon_n \xrightarrow[n\to\infty]{\mathbb{P}_{\theta^*}} 0.$$

Next, Proposition B.0.5 implies that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_i(\theta^*) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, I(\theta^*)).$$

Consequently, since by (3.23) $G_n \xrightarrow{\mathbb{P}_{\theta^*}} I^{-1}(\theta^*)$, using Slutsky's theorem we get

$$\frac{G_n}{\sqrt{n}}\sum_{i=1}^{n}\psi_i(\theta^*) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, I^{-1}(\theta^*)).$$

The proof is complete. □

We end this section with the following technical result, which will be used in our construction of confidence region in Section 6. Towards this end, for any $\theta \in \boldsymbol{\Theta}$ and $n \geq 1$, we define[12]

$$\begin{aligned} U_n(\theta) &:= n(\hat{\theta}_n - \theta)^T I(\tilde{\theta}_n)(\hat{\theta}_n - \theta) \\ &= n\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma_n^{ij}(\hat{\theta}_n^i - \theta^i)(\hat{\theta}_n^j - \theta^i), \end{aligned} \tag{3.28}$$

where $(\sigma_n^{ij})_{i,j=1,\ldots,d} = I(\tilde{\theta}_n)$, and, as usual, we denote by χ_d^2 a random variable that has the chi-squared distribution with d degrees of freedom.

[12]We use superscripts here to denote components of vectors and matrices.

Corollary 3.3.1. *With* $\vartheta_n = -I^{-1}(\tilde{\theta}_n)I_n\theta^*$*, we have that*

$$U_n(\vartheta_n) \xrightarrow[n\to\infty]{d} \chi^2_d.$$

Proof. From Assumption R8, strong consistency of $\tilde{\theta}$ and Proposition 3.3.1, and employing the Slutsky's theorem again, we get that

$$\sqrt{nI(\tilde{\theta}_n)}(\hat{\theta}_n - \vartheta_n) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \mathrm{Id}).$$

Therefore,

$$U_n(\vartheta_n) = n(\hat{\theta}_n - \vartheta_n)^T I(\tilde{\theta}_n)(\hat{\theta}_n - \vartheta_n) \xrightarrow{d} \varsigma^T\varsigma,$$

where $\varsigma \sim \mathcal{N}(0, \mathrm{Id})$. The proof is thus complete since $\varsigma^T\varsigma \stackrel{d}{=} \chi^2_d$.

□

3.4 Recursive construction of confidence regions

This section is devoted to the construction of the recursive confidence region based on quasi-asymptotically linear estimator $\hat{\theta}$ developed in Section 3.3. We start with introducing the definition of the approximate confidence region.

Definition 3.4.1. *Let* $\{\mathcal{V}_n, n \geq 1\}$ *be such that* $\mathcal{V}_n : \mathbb{R}^{n+1} \to 2^{\boldsymbol{\Theta}}$ *and* $\mathcal{V}_n(z)$ *is a connected set*[13] *for any* $z \in \mathbb{R}^{n+1}$*,* $n \geq 1$*. The set* $\{\mathcal{V}_n(Z_0^n), n \geq 1\}$*, with* $Z_0^n := (Z_0, \ldots, Z_n)$*,* $n \geq 1$*, is called a sequence of approximate confidence regions for* θ^**, at significant level* $\alpha \in (0,1)$*, if there exists a weakly consistent estimator* ϑ *of* θ^**, such that*

$$\lim_{n\to\infty} \mathbb{P}_{\theta^*}(\vartheta_n \in \mathcal{V}_n(Z_0^n)) = 1 - \alpha.$$

[13] A connected set is a set that cannot be represented as the union of two or more disjoint nonempty open subsets.

Such sequence of approximate confidence regions can be constructed, as the next result shows, by using the asymptotic results obtained in Section 3.3. Recall the notation $U_n(\theta) = n(\hat{\theta}_n - \theta)^T I(\tilde{\theta}_n)(\hat{\theta}_n - \theta)$, for $\theta \in \boldsymbol{\Theta}$, $n \geq 1$.

Proposition 3.4.1. *Fix a confidence level α, and let $\kappa \in \mathbb{R}$ be such that $\mathbb{P}_{\theta^*}(\chi^2_d < \kappa) = 1 - \alpha$. Then, the set $\{\mathcal{T}_n, n \geq 1\}$ such that*

$$\mathcal{T}_n := \{\theta \in \boldsymbol{\Theta} : U_n(\theta) < \kappa\}$$

is a sequence of approximate confidence regions for θ^*.

Proof. As in Section 3.3, we take $\vartheta_n = -I^{-1}(\hat{\theta}_n)I_n\theta^*$, which in view of Proposition 3.3.1 is a weakly consistent estimator of θ^*. Note that $U_n(\,\cdot\,)$ is a continuous function, and thus $\mathcal{T}_n$ is a connected set, for any $n \geq 1$. By Corollary 3.3.1, $U_n(\vartheta_n) \xrightarrow{d} \chi^2_d$, and since $\mathbb{P}_{\theta^*}(\vartheta_n \in \mathcal{T}_n) = \mathbb{P}_{\theta^*}(U_n(\vartheta_n) < \kappa)$, we immediately have that $\lim_{n\to\infty} \mathbb{P}_{\theta^*}(\vartheta_n \in \mathcal{T}_n) = 1 - \alpha$. This concludes the proof. □

Next, we will show that the approximate confidence region $\mathcal{T}_n$ can be computed in a recursive way, by taking into account its geometric structure. By the definition, the set $\mathcal{T}_n$ is the interior of a d-dimensional ellipsoid, and hence $\mathcal{T}_n$ is uniquely determined by its extreme $2d$ points. Thus, it is enough to establish a recursive formula for computing the extreme points. Let us denote by

$$(\theta^1_{n,k}, \ldots, \theta^d_{n,k}), \quad k = 1, \ldots, 2d,$$

the coordinates of these extreme points; that is $\theta^i_{n,k}$, denotes the ith coordinate of the kth extreme point of ellipsoid $\mathcal{T}_n$.

First, note that the matrix $I(\tilde{\theta}_n)$ is positive definite, and hence it admits the

Cholesky decomposition:

$$I(\tilde{\theta}_n) = L_n L_n^T = \begin{bmatrix} l_n^{11} & 0 & \cdots & 0 \\ l_n^{21} & l_n^{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ l_n^{d1} & l_n^{d2} & \cdots & l_n^{dd} \end{bmatrix} \begin{bmatrix} l_n^{11} & l_n^{21} & \cdots & l_n^{d1} \\ 0 & l_n^{22} & \cdots & l_n^{d2} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & l_n^{dd} \end{bmatrix},$$

where l_n^{ij} $i, j = 1, \ldots, d$, are given by

$$l_n^{ii} = \sqrt{\sigma_n^{ii} - \sum_{k=1}^{i-1} (l_n^{ik})^2},$$

$$l_n^{ij} = \frac{1}{l_n^{ii}} \Big(\sigma_n^{ij} - \sum_{k=1}^{j-1} l_n^{ik} l_n^{jk} \Big).$$

Thus, we have that $U_n(\theta) = n(u_{n,1}^2(\theta) + u_{n,2}^2(\theta) + \cdots + u_{n,d}^2(\theta))$, where

$$u_{n,i}(\theta) = \sum_{j=i}^{d} l_n^{ji} (\hat{\theta}_n^j - \theta^j), \quad i = 1, \ldots, d,$$

and thus $\mathcal{T}_n = \{\theta : \sum_{j=1}^d (u_{n,j}(\theta))^2 < \frac{\kappa}{n}\}$.

By making the coordinate transformation $\theta \mapsto \varrho$ given by $\varrho = L_n^T(\hat{\theta}_n - \theta)$, the set $\mathcal{T}_n$ in the new system of coordinates can be written as $\mathcal{T}_n = \{\varrho : \sum_{i=1}^d (\varrho^i)^2 < \frac{\kappa}{n}\}$. Hence, $\mathcal{T}_n$, in the new system of coordinates, is determined by the following $2d$ extreme points of the ellipsoid:

$$(\varrho_1^1, \ldots, \varrho_1^d) = (\sqrt{\frac{\kappa}{n}}, 0, \ldots, 0),$$

$$(\varrho_2^1, \ldots, \varrho_2^d) = (-\sqrt{\frac{\kappa}{n}}, 0, \ldots, 0),$$

$$\cdots$$

$$(\varrho_{2d-1}^1, \ldots, \varrho_{2d-1}^d) = (0, \ldots, 0, \sqrt{\frac{\kappa}{n}}),$$

$$(\varrho_{2d}^1, \ldots, \varrho_{2d}^d) = (0, \ldots, 0, -\sqrt{\frac{\kappa}{n}}).$$

Then, in the original system of coordinates, the extreme points (written as vectors) are given by

$$\begin{aligned}(\theta^1_{n,2j-1},\dots,\theta^d_{n,2j-1})^T &= \hat{\theta}_n - \sqrt{\frac{\kappa}{n}}(L_n^T)^{-1}e_j,\\ (\theta^1_{n,2j},\dots,\theta^d_{n,2j})^T &= \hat{\theta}_n + \sqrt{\frac{\kappa}{n}}(L_n^T)^{-1}e_j,\end{aligned} \qquad j = 1,\dots,d, \tag{3.29}$$

where $\{e_j\}$, $j = 1,\dots,d$, is the standard basis in $\mathbb{R}^d$.

Finally, taking into account the recursive constructions (3.11), (3.16), and the representation (3.29), we have the following recursive scheme for computing the approximate confidence region.

Recursive construction of the confidence region

Initial Step: $\quad \Gamma_0 = 0, \quad I_0 = 0, \quad \tilde{\theta}_0 \in \boldsymbol{\Theta}$.

n^{th} Step:

Input: $\tilde{\theta}_{n-1}, I_{n-1}, \Gamma_{n-1}, Z_{n-1}, Z_n$.

Output:
$$\begin{aligned}
\tilde{\theta}_n &= \tilde{\theta}_{n-1} + \frac{\eta}{n}\psi_n(\tilde{\theta}_{n-1}),\\
I_n &= \frac{n-1}{n}I_{n-1} + \frac{1}{n}\Psi_n(\tilde{\theta}_{n-1}),\\
\Gamma_n &= \frac{n-1}{n}\Gamma_{n-1} + \frac{1}{n}(\mathrm{Id} + \eta I_n)\psi_n(\tilde{\theta}_{n-1}),\\
(\theta^1_{n,2j},\dots,\theta^d_{n,2j})^T &= -I^{-1}(\tilde{\theta}_n)I_n\tilde{\theta}_n + I^{-1}(\tilde{\theta}_n)\Gamma_n + \sqrt{\frac{\kappa}{n}}(I_n^{-1/2})^T e_j,\\
(\theta^1_{n,2j-1},\dots,\theta^d_{n,2j-1})^T &= -I^{-1}(\tilde{\theta}_n)I_n\tilde{\theta}_n + I^{-1}(\tilde{\theta}_n)\Gamma_n - \sqrt{\frac{\kappa}{n}}(I_n^{-1/2})^T e_j,\\
&\qquad j = 1,\dots,d.
\end{aligned}$$

From here, we also conclude that there exists a function τ, independent of n, such that

$$\mathcal{T}_n = \tau(\mathcal{T}_{n-1}, Z_n). \tag{3.30}$$

The above recursive relationship goes to heart of application of recursive confidence regions in the robust adaptive control theory originated in [BCC+16b], since it makes it possible to take the full advantage of the dynamic programming principle in the context of such control problems.

We conclude this section by proving that the confidence region converges to the singleton θ^*. Equivalently, it is enough to prove that the extreme points converge to the true parameter θ^*.

Proposition 3.4.2. *For any $k \in \{1, \dots, 2d\}$, we have that*

$$\mathbb{P}_{\theta^*}\text{-}\lim_{n\to\infty} \theta_{n,k} = \theta^*.$$

Proof. By Assumption R8 and Theorem 3.2.1 (strong consistency of $\tilde{\theta}$), we have that $L_n \xrightarrow[n\to\infty]{\text{a.s.}} I^{1/2}(\theta^*)$, and consequently, we also have that

$$\sqrt{\frac{\kappa}{n}} e_j^T L_n^{-1} \xrightarrow[n\to\infty]{\text{a.s}} 0. \tag{3.31}$$

Of course, the last convergence holds true in the weak sense too. Passing to the limit in (3.29), in $\mathbb{P}_{\theta^*}$ probability sense, and using (3.31) and weak consistency of $\hat{\theta}$ (Theorem 3.3.1), we finish the proof.

□

3.5 Examples

In this section we will present three illustrative examples of the recursive construction of confidence regions developed above. We start with our main example, Example 3.5.1, of a Markov chain with Gaussian transitional densities where both the conditional mean and conditional standard deviation are the parameters of interest. Example 3.5.2 is dedicated to the case of i.i.d. Gaussian observations, which is a particular case of the first example.

Generally speaking, the simple case of i.i.d. observations for which the MLE exists and asymptotic normality holds true, one can recursively represent the sequence of confidence intervals constructed in the usual (off-line) way, and the theory developed in this chapter is not really needed. The idea is illustrated in Example 3.5.3 by considering again the same experiment as in Example 3.5.2. In fact, as mentioned above, this idea served as the starting point for the general methodology presented in the thesis.

Example 3.5.1. *Let us consider a Markov process $\{Z_n\}$ with a Gaussian transition density function*

$$p_\theta(x,y) = \frac{1}{\sqrt{1-\rho^2}\sqrt{2\pi}\sigma} e^{-\frac{(y-\rho x-(1-\rho)\mu)^2}{2\sigma^2(1-\rho^2)}}, \quad n \geq 1,$$

and such that $Z_0 \sim \mathcal{N}(\mu, \sigma^2)$.

We assume that the correlation parameter $\rho \in (-1,1)$ is known, and the unknown parameter is $\theta = (\mu, \sigma) \in \Theta$, where $\Theta = [a_1, a_2] \times [b_1, b_2]$, and $a_1 \leq a_2$, $b_1 \leq b_2$ are some fixed real numbers with $b_1 > 0$.

In the Appendix C we show that the process Z satisfies the Assumption M, and the conditions R0-A9.

Thus, all the results derived in the previous sections hold true. Moreover, for a given confidence level α, we have the following explicit formulas for the nth step of the recurrent construction of the confidence regions:

$$\tilde{\mu}_n = \tilde{\mu}_{n-1} + \frac{\eta(Z_n - \rho Z_{n-1} - (1-\rho)\tilde{\mu}_{n-1})}{n\tilde{\sigma}^2_{n-1}(1+\rho)},$$

$$\tilde{\sigma}^2_n = \tilde{\sigma}^2_{n-1} - \frac{\eta}{n\tilde{\sigma}_{n-1}} + \frac{\eta(Z_n - \rho Z_{n-1} - (1-\rho)\tilde{\mu}_{n-1})^2}{n(1-\rho^2)\tilde{\sigma}^3_{n-1}}),$$

$$I_n = \frac{n-1}{n} I_{n-1} + \frac{1}{n} \begin{bmatrix} -\frac{1-\rho}{(1+\rho)\tilde{\sigma}^2_{n-1}} & -\frac{2(Z_n-\rho Z_{n-1}-(1-\rho)\tilde{\mu}_{n-1})}{(1+\rho)\tilde{\sigma}^3_{n-1}} \\ -\frac{2(Z_n-\rho Z_{n-1}-(1-\rho)\tilde{\mu}_{n-1})}{(1+\rho)\tilde{\sigma}^3_{n-1}} & \frac{1}{\tilde{\sigma}^2_{n-1}} - \frac{3(Z_n-\rho Z_{n-1}-(1-\rho)\tilde{\mu}_{n-1})^2}{(1-\rho^2)\tilde{\sigma}^4_{n-1}} \end{bmatrix},$$

$$\Gamma_n = \frac{n-1}{n}\Gamma_{n-1} + \frac{1}{n}(Id + \eta I_n) \begin{bmatrix} \tilde{\mu}_{n-1} \\ \tilde{\sigma}^2_{n-1} \end{bmatrix},$$

and, for $j \in \{1, 2, 3, 4\}$,

$$\begin{bmatrix} \mu_{n,j} \\ \\ \sigma^2_{n,j} \end{bmatrix} = -\begin{bmatrix} \frac{(1+\rho)\tilde{\sigma}_n^2}{1-\rho} & 0 \\ \\ 0 & \frac{\tilde{\sigma}_n^2}{2} \end{bmatrix} I_n \begin{bmatrix} \tilde{\mu}_n \\ \\ \tilde{\sigma}_n^2 \end{bmatrix} + \begin{bmatrix} \frac{(1+\rho)\tilde{\sigma}_n^2}{1-\rho} & 0 \\ \\ 0 & \frac{\tilde{\sigma}_n^2}{2} \end{bmatrix} \Gamma_n + \varpi_j \frac{\kappa}{n} \begin{bmatrix} \sqrt{\frac{1+\rho}{1-\rho}}\tilde{\sigma}_n & 0 \\ \\ 0 & \frac{\tilde{\sigma}_n}{\sqrt{2}} \end{bmatrix} u_j,$$

where $\varpi_1 = \varpi_3 = -1$, $\varpi_2 = \varpi_4 = 1$, $u_1 = u_2 = e_1$, $u_3 = u_4 = e_2$, η *is a constant such that* $\eta > \frac{b_2^3}{4b_1}$, $\eta > \frac{(1+\rho)b_2^3}{2(1-\rho)b_1}$, *and* $\mathbb{P}_{\theta^*}(\chi_2^2 < \kappa) = 1 - \alpha$.

Example 3.5.2. *Let* Z_n, $n \geq 0$, *be a sequence of i.i.d. Gaussian random variables with an unknown mean* μ *and unknown standard deviation* σ. *Clearly, this important case is a particular case of Example 3.5.1, with* $\rho = 0$, *and the same recursive formulas for confidence regions by taking* $\rho = 0$ *in the above formulas.*

Example 3.5.3. *We take the same set-up as in the previous example - i.i.d Gaussian random variables with unknown mean and standard deviation. We will use the fact that in this case, the MLE estimators for* μ *and* σ^2 *are computed explicitly and given by*

$$\hat{\mu}_n = \frac{1}{n+1}\sum_{i=0}^{n} Z_i, \quad \hat{\sigma}_n^2 = \frac{1}{n+1}\sum_{i=0}^{n}(Z_i - \hat{\mu}_n)^2, \quad n \geq 1,$$

It is well known that $(\hat{\mu}, \hat{\sigma}^2)$ *are asymptotically normal, namely*

$$\sqrt{n}(\hat{\mu}_n - \mu^*, \hat{\sigma}_n^2 - (\sigma^*)^2) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, I^{-1}),$$

where

$$I = \begin{bmatrix} (\sigma^*)^2 & 0 \\ \\ 0 & 2(\sigma^*)^4 \end{bmatrix}.$$

First, note that $(\hat{\mu}_n, \hat{\sigma}_n^2)$ *satisfies the following recursion:*

$$\begin{aligned} \hat{\mu}_n &= \frac{n}{n+1}\hat{\mu}_{n-1} + \frac{1}{n+1}Z_n, \\ \hat{\sigma}_n^2 &= \frac{n}{n+1}\hat{\sigma}_{n-1}^2 + \frac{n}{(n+1)^2}(\hat{\mu}_n - Z_n)^2, \quad n \geq 1. \end{aligned} \tag{3.32}$$

Second, due to asymptotic normality, we also have that, $U_n \xrightarrow[n\to\infty]{d} \chi_2^2$, *where* $U_n := \frac{n}{\hat{\sigma}_n^2}(\hat{\mu}_n - \mu^*)^2 + \frac{n}{2\hat{\sigma}_n^4}(\hat{\sigma}_n^2 - (\sigma^*)^2)^2$. *Now, for a given confidence level* α, *we let* $\kappa \in \mathbb{R}$ *be such that* $\mathbb{P}_{\theta^*}(\chi_2^2 < \kappa) = 1 - \alpha$, *and then, the confidence region for* (μ, σ^2) *is given by*

$$\mathcal{T}_n := \left\{ (\mu, \sigma^2) \in \mathbb{R}^2 : \frac{n}{\hat{\sigma}_n^2}(\hat{\mu}_n - \mu)^2 + \frac{n}{2\hat{\sigma}_n^4}(\hat{\sigma}_n^2 - \sigma^2)^2 < \kappa \right\}.$$

Similar to the previous cases, we note that $\mathcal{T}_n$ *is the interior of an ellipse (in* $\mathbb{R}^2$*), that is uniquely determined by its extreme points*

$$(\mu_{n,1}, \sigma_{n,1}^2) = \left(\hat{\mu}_n + \sqrt{\frac{\kappa}{n}}\hat{\sigma}_n, \hat{\sigma}_n^2 \right), \qquad (\mu_{n,2}, \sigma_{n,2}^2) = \left(\hat{\mu}_n - \sqrt{\frac{\kappa}{n}}\hat{\sigma}_n, \hat{\sigma}_n^2 \right),$$
$$(\mu_{n,3}, \sigma_{n,3}^2) = \left(\hat{\mu}_n, \left(1 + \sqrt{\frac{2\kappa}{n}} \right) \hat{\sigma}_n^2 \right), \quad (\mu_{n,4}, \sigma_{n,4}^2) = \left(\hat{\mu}_n, \left(1 - \sqrt{\frac{2\kappa}{n}} \right) \hat{\sigma}_n^2 \right).$$

Therefore, taking into account (3.32), *we have a recursive formula for computing these extreme points, and thus the desired recursive construction of the confidence regions* $\mathcal{T}_n$.

CHAPTER 4
FUTURE WORK

In Chapter 2, we have introduced an axiomatic approach for modeling bid and ask prices in general financial markets. Our framework is constructed in a discrete-time setting. Therefore, a major future work will be to develop analogues of our dynamic conic-finance theory in the continuous-time set-up. Such extension will be divided into four major research tasks:

1. The first step is to develop continuous-time versions of dynamic risk measures and dynamic acceptability indices for dividend-paying securities, and to establish the connection between such dynamic assessment indices and backward stochastic differential equations.

2. The next research problem is to define a market model under the continuous-time assumption, and to introduce the relevant financial definitions such as value process, self-financing trading strategy, and arbitrage.

3. As in the current study, special attention will be devoted to dividend paying securities. The main difficulty will be to prove that dynamic conic finance via backward stochastic differential equations is a no-arbitrage pricing framework.

4. Finally, it is interesting to study the limit behavior of the theory developed in this work as the time increment goes to 0.

In Chapter 3, we initiated the theory of recursive confidence regions. In part, this theory hinges on the theory of recursive identification for stochastic dynamical systems, such as a Markov chain, which is the main model studied here. Although the results in the existing literature on statistical inference for Markov processes are quite general, not much work has been done on the recursive identification methods for Markov processes. Our results provide a useful contribution in this regard, but, they are subject to assumption of ergodicity imposed on our Markov chain. We leave the study of more general cases to the future work.

It is implicitly assumed that our base estimator at each time step belongs to the parameter space. One way to ensure this is to choose the parameter space as the largest possible set that the true parameter lies in. However, this is not always possible so that one needs to consider constrained recursive point estimators. We leave for the future work the study of recursive confidence regions generated via constrained recursive point estimation algorithms.

APPENDIX A

ERGODIC THEORY FOR MARKOV CHAINS

In this section, we will briefly discuss the theory of ergodicity for (time homogeneous) Markov processes in discrete time. Note that for fixed transition kernel Q and initial distribution μ, all the corresponding Markov processes have the same law. With this in mind, we will present results regarding ergodicity of Markov processes associated to the canonical construction from Q and μ. We start with recalling the notion for ergodicity of general dynamical systems.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $T : \Omega \to \Omega$ be a measure preserving map, i.e. a map such that $\mathbb{P}(T^{-1}(A)) = \mathbb{P}(A)$ for every $A \in \mathcal{F}$. Then, the corresponding dynamical system is defined as the quadruple $(\Omega, \mathcal{F}, \mathbb{P}, T)$. Define $\mathcal{G} := \{A \in \mathcal{F} \, : \, T^{-1}(A) = A\}$, and note that $\Omega, \emptyset \in \mathcal{G}$. Then, we have the following

Definition A.0.1. *A dynamical system* $(\Omega, \mathcal{F}, \mathbb{P}, T)$ *is said to be ergodic if for any* $A \in \mathcal{G}$ *we have* $\mathbb{P}(A) = 0$ *or* $\mathbb{P}(A) = 1$.

One important result in the theory of dynamical system is the celebrated Birkhoff's Ergodic Theorem (See e.g. [Bir31], [vN32b], [vN32a]).

Theorem A.0.1 (Birkhoff's Ergodic Theorem)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P}, T)$ *be an ergodic dynamical system. If* $f \in L^1(\Omega, \mathcal{F}, \mathbb{P})$*. Then,*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n \omega) = \mathbb{E}_{\mathbb{P}}[f] \quad \mathbb{P} - a.s.$$

We now proceed by introducing the canonical construction of time homogeneous Markov chains. Let $(\mathcal{X}, \mathfrak{X})$ be a measurable space. Also, let $Q : \mathcal{X} \times \mathfrak{X} \to [0,1]$ be a transition kernal and π be a probability measure on $(\mathcal{X}, \mathfrak{X})$ such that $\pi(A) = \int_{\mathcal{X}} Q(x, A)\pi(dx)$, for any $A \in \mathfrak{X}$. Such measure π is called an invariant probability measure of Q. For every $n \geq 0$, we define a probability measure $\mathbb{P}^{Q,n}_{\pi}$ on $(\mathcal{X}^{n+1}, \mathfrak{X}^{n+1})$, where $\mathfrak{X}^{n+1}$ is the product σ-algebra on $\mathcal{X}^{n+1}$, by

$$\mathbb{P}^{Q,n}_{\pi}(A_0 \times \ldots \times A_n) = \int_{A_0} \cdots \int_{A_n} Q(x_{n-1}, dx_n) \cdots Q(x_0, dx_1)\pi(dx_0),$$

for any $A_0, \dots, A_n \in \mathfrak{X}$. The sequence of measures $\{\mathbb{P}_\pi^{Q,n}\}_{n>0}$ is consistent. That is,

$$\mathbb{P}_\pi^{Q,n}(A_0 \times A_1 \ldots \times A_n) = \mathbb{P}_\pi^{Q,n+m}(A_0 \times A_1 \times \ldots \times A_n \times \mathcal{X}^m),$$

holds true for any integer $m > 0$, and $A_0, \dots, A_n \in \mathfrak{X}$. Therefore, by Kolmogorov's extension theorem, such family of measures extends to a unique measure $\mathbb{P}_\pi^Q$ on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}})$, such that

$$\mathbb{P}_\pi^Q(A_0 \times A_1 \ldots \times A_n \times \mathcal{X}^\infty) = \mathbb{P}_\pi^{Q,n}(A_0 \times A_1 \ldots \times A_n), \quad A_0, \dots, A_n \in \mathfrak{X}. \tag{A.1}$$

With a slight abuse of notation, we denote by T the (one step) shift map on $\mathcal{X}^{\mathbb{N}}$

$$(T(\omega))_k = \omega_{k+1}, \quad \omega \in \mathcal{X}^{\mathbb{N}}.$$

Due to the construction of $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$ and the fact that π is an invariant measure, then it can be verified that T is measure preserving, and therefore $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q, T)$ is a dynamical system.

Next, define a process X on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$ by

$$X(\omega) = \omega, \quad \omega \in \mathcal{X}^{\mathbb{N}},$$

so that, in particular, $X_n(\omega) = \omega(n)$ for any integer $n \geq 0$. A process defined in this way is called a *canonical process* on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$.

We now state and prove the following result,

Lemma A.0.1. *A canonical process X on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$ is a time homogenous Markov chain with transition kernel Q, and thus it is called the canonical Markov chain on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$. Moreover, the initial distribution of X coincides with π, so that*

$$\mathbb{P}_\pi^Q(X_0 \in A) = \pi(A).$$

Consequently, process X is a stationary process, that is, for any $n \geq 1$, the law of $(X_j, X_{j+1}, \dots, X_{j+n})$ under $\mathbb{P}_\pi^Q$ is independent of j, $j \geq 0$.

Proof. For any $n > 1$, denote by $dx_{0:n} := dx_n \times \cdots \times dx_0$. According to (A.1) and the definition of $\mathbb{P}_\pi^{Q,n}$, we obtain that

$$\mathbb{P}_\pi^Q(dx_{0:n}) = \mathbb{P}_\pi^{Q,n}(dx_{0:n}) = Q(x_{n-1}, dx_n) \cdots Q(x_0, dx_1)\pi(dx_0).$$

Next, for any $A_0, \dots, A_n \in \mathfrak{X}$, we get that

$$\begin{aligned}\mathbb{P}_\pi^Q(A_n \times \cdots \times A_0) &= \mathbb{E}_\pi^Q[\mathbb{1}_{A_n \times \cdots \times A_0}] \\ &= \int_{A_{n-1}\times\cdots\times A_0} \mathbb{E}_\pi^Q[\mathbb{1}_{A_n} \mid X_{n-1} = x_{n-1}, \dots, X_0 = x_0]\mathbb{P}_\pi^Q(dx_{0:n-1}).\end{aligned} \tag{A.2}$$

On the other hand, we also have that

$$\begin{aligned}\mathbb{P}_\pi^Q(A_n \times \cdots A_0) = \mathbb{P}_\pi^{Q,n}(A_n \times \cdots A_0) &= \int_{A_0} \cdots \int_{A_n} Q(x_{n-1}, dx_n) \cdots Q(x_0, dx_1)\pi(dx_0) \\ &= \int_{A_{n-1}\times\cdots\times A_0} \int_{A_n} Q(x_{n-1}, dx_n)\mathbb{P}_\pi^Q(dx_{0:n-1}).\end{aligned} \tag{A.3}$$

(A.2) and (A.3) yield that

$$\begin{aligned}\mathbb{P}_\pi^Q(X_n \in A_n \mid X_{n-1} = x_{n-1}, \dots, X_0 = x_0) &= \mathbb{E}_\pi^Q[\mathbb{1}_{A_n} \mid X_{n-1} = x_{n-1}, \dots, X_0 = x_0] \\ &= \int_{A_n} Q(x_{n-1}, dx_n) \\ &= \mathbb{P}_\pi^Q(X_n \in A_n \mid X_{n-1} = x_{n-1}).\end{aligned}$$

Therefore, we conclude that X is a Markov chain.

Now we prove the initial distribution of X is π. By definition of X we have

$$\mathbb{P}_\pi^Q(X_0 \in A) = \mathbb{P}_\pi^Q(\omega(0) \in A) = \mathbb{P}_\pi^Q(A \times \mathcal{X}^\infty).$$

Then, according to (A.1), it is true that

$$\mathbb{P}_\pi^Q(A \times \mathcal{X}^\infty) = \mathbb{P}_\pi^{Q,0}(A) = \pi(A).$$

Therefore,

$$\mathbb{P}_\pi^Q(X_0 \in A) = \pi(A),$$

and π is the initial distribution of X.

We finish the proof by showing the stationarity of X. That is to prove for any fixed $n \geq 1$, the probability $\mathbb{P}_\pi^Q(X_j \in A_0, \ldots, X_{n+j} \in A_n)$ is independent of $j \geq 0$. Since π is invariant measure Q, then it is clear that $\mathbb{P}_\pi^Q(X_j \in A_0) = \pi(A_0)$. Next, we have

$$\mathbb{P}_\pi^Q(X_j \in A_0, \ldots, X_{n+j} \in A_n) = \mathbb{P}_\pi^Q(\mathcal{X}^j \times A_0 \times \cdots \times A_n \times \mathcal{X}^\infty),$$

where the right hand side is equal to $\mathbb{P}_\pi^{Q,n+j}(\mathcal{X}^j \times A_0 \times \cdots \times A_n)$ by (A.1). Finally, according to the definition of $\mathbb{P}_\pi^{Q,n+j}$, we have

$$\begin{aligned}
\mathbb{P}_\pi^{Q,n+j}(\mathcal{X}^j \times A_0 \times \ldots \times A_n) &= \int_{\mathcal{X}} \int_{A_0} \cdots \int_{A_n} Q(x_{n+j-1}, dx_{n+j}) \\
&\qquad \cdots Q(x_{n-1}, dx_n) \cdots Q(x_0, dx_1) \pi(dx_0), \\
&= \int_{A_0} \cdots \int_{A_n} Q(x_{n+j-1}, dx_{n+j}) \cdots Q(x_j, dx_{j+1}) \pi(dx_j) \\
&= \mathbb{P}_\pi^{Q,n}(A_0 \times \ldots \times A_n) = \mathbb{P}_\pi^Q(A_0 \times \ldots \times A_n \times \mathcal{X}^\infty) \\
&= \mathbb{P}_\pi^Q(X_0 \in A_0, \ldots, X_n \in A_n).
\end{aligned}$$

We now conclude that X is a stationary process. □

Remark A.0.1. *If a transition kernel Q admits an invariant measure π, then it is customary to say that π is an invariant measure for any Markov chain corresponding to Q. In particular, π is the invariant measure for the canonical Markov chain X on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$.*

We proceed by defining the notion of ergodicity for a canonical Markov chain X.

Definition A.0.2. *The canonical Markov chain X on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$ is said to be ergodic if $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q, T)$ is an ergodic dynamical system.*

Remark A.0.2. *Note that since an ergodic Markov chain X is, in particular, a canonical Markov chain on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_\pi^Q)$, then it is a stationary process.*

Through the rest of this section X denote the canonical Markov chain defined on $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_{\pi}^{Q})$. The following technical result is one of the key technical results used in this thesis. In its formulation we denote by $\mathbb{E}_{\pi}^{Q}$ the expectation under measure $\mathbb{P}_{\pi}^{Q}$.

Proposition A.0.1. *Let X be ergodic. Then for any g such that $\mathbb{E}_{\pi}^{Q}[g(X_0, \ldots, X_n)] < \infty$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} g(X_i, \ldots, X_{i+n}) = \mathbb{E}_{\pi}^{Q}[g(X_0, \ldots, X_n)] \quad \mathbb{P}_{\pi}^{Q} - a.s.$$

Proof. By definition, we have that $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_{\pi}^{Q}, T)$ is an ergodic dynamical system.

For fixed $n > 0$, take $f : \mathcal{X}^{\mathbb{N}} \to \mathbb{R}$ defined as $f(\omega) := g(\omega(0), \ldots, \omega(n))$ for any $\omega \in \mathcal{X}^{\mathbb{N}}$. Note that

$$\omega(j) = X_j(\omega), \quad j \geq 0,$$

and

$$T^i(\omega)(j) = X_{i+j}(\omega), \quad i, j \geq 0.$$

Then, according to Birkhoff's ergodic theorem, we get that for almost every $\omega \in \mathcal{X}^{\mathbb{N}}$:

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} g(X_i(\omega), \ldots, X_{i+n}(\omega)) &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} f(T^i(\omega)) \\ &= \mathbb{E}_{\pi}^{Q}[f] = \mathbb{E}_{\pi}^{Q}[g(X_0, \ldots, X_n)]. \end{aligned}$$

□

We finish this section with providing a brief discussion regarding sufficient conditions for the Markov chain X to be ergodic. Towards this end, first note that, in general, a transition kernel Q possesses more than one invariant measures, and we quote the following structural result regarding the set of invariant measures of Q,

Proposition A.0.2. *[Var01] Let $Q : \mathcal{X} \times \mathfrak{X} \to [0, 1]$ be a (one step) transition kernel. Then, the set Π_Q of all invariant probability measures of Q is convex. Also, given a* *measure* $\pi \in \Pi$*, the system* $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_{\pi}^{Q}, T)$ *is ergodic if and only if* π *is an extremal point of* Π*. Furthermore, any two ergodic measures are either identical or mutually singular.*

Proposition A.0.2 implies

Corollary A.0.1. *If a transition kernel* Q *has a unique invariant probability measure* π*, then the system* $(\mathcal{X}^{\mathbb{N}}, \mathfrak{X}^{\mathbb{N}}, \mathbb{P}_{\pi}^{Q}, T)$ *is ergodic.*

One powerful tool for checking the uniqueness of invariant probability measure is the notion of positive Harris chain. There are several equivalent definitions of positive Harris Markov chain, and we will use the one from [HLL00].

Definition A.0.3. *The Markov chain* X *with transition kernel* Q *is called a positive Harris chain if*

(a) there exists a σ*-finite measure* μ *on* $\mathfrak{X}$ *such that for any* $x_0 \in \mathcal{X}$*, and* $B \in \mathfrak{X}$ *with* $\mu(B) > 0$

$$\mathbb{P}(X_n \in B \text{ for some } n < \infty | X_0 = x_0) = 1,$$

(b) there exists an invariant probability measure for Q*.*

Proposition A.0.3. *If* X *is a positive Harris chain, then* X *is ergodic.*

Proof. It is well known (cf. e.g. [MT93]) that a positive Harris chain admits a unique invariant measure. Thus, the result follows form Corollary (A.0.1).

□

APPENDIX B

CLT FOR MULTIVARIATE MARTINGALES

In this section, for a matrix A with real valued entries we denote by $|A|$ the sum of the absolute values of its entries.

In [CP05] Proposition 3.1, the authors gave the following version of the central limit theorem for discrete time multivariate martingales.

Proposition B.0.4. *On a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *let* $D = \{D_{n,j}, 0 \leq j \leq k_n, n \geq 1\}$ *be a triangular array of* d*-dimensional real random vectors, such that, for each* n*, the finite sequence* $\{D_{n,j}, \mathbb{1} \leq j \leq k_n\}$ *is a martingale difference process with respect to some filtration* $\{\mathcal{F}_{n,j}, j \geq 0\}$*. Set*

$$D_n^* = \sup_{1 \leq j \leq k_n} |D_{n,j}|, \quad U_n = \sum_{j=1}^{k_n} D_{n,j} D_{n,j}^T.$$

Also denote by $\mathscr{U}$ *the* σ*-algebra generated by* $\bigcup_j \mathscr{H}_j$ *where* $\mathscr{H}_j := \liminf_n \mathcal{F}_{n,j}$*. Suppose that* D_n^* *converges in* L^1 *to zero and that* U_n *converges in probability to a* $\mathscr{U}$ *measurable* d*-dimensional, positive semi-definite matrix* U*. Then, the random vector* $\sum_{j=1}^{k_n} D_{n,j}$ *converges* $\mathscr{U}$*-stably to the Gaussian kernel* $\mathcal{N}(0, U)$*.*

Remark B.0.3. *$\mathscr{U}$-stable convergence implies convergence in distribution; it is enough to take the entire Ω in the definition of $\mathscr{U}$-stable convergence. See for example [AE78] or [HL15].*

We will apply the above proposition to the process $\{\psi_n(\theta^*), n \geq 0\}$ such that Assumption M, R8 and R9 are satisfied. To this end, let us define the triangular array $\{D_{n,j}, 1 \leq j \leq n, n \geq 1\}$ as

$$D_{n,j} = \frac{1}{\sqrt{n}} \psi_j(\theta^*),$$

and let us take $\mathcal{F}_{n,j} = \mathcal{F}_j$.

First, note that $\mathbb{E}_{\theta^*}[\psi_j(\theta^*) | \mathcal{F}_{j-1}] = 0$, so that for any $n \geq 1$, $\{D_{n,j}, 1 \leq j \leq n\}$ is a martingale difference process with respect to $\{\mathcal{F}_j, 0 \leq j \leq n\}$. Next, R9 implies that $D_n^* := \sup_{1\le j\le n} \frac{1}{\sqrt{n}}|\psi_j(\theta^*)|$ converges in L^1 to 0. Finally, stationarity, R8 and ergodicity guarantee that

$$U_n := \frac{1}{n}\sum_{j=1}^{n} \psi_j(\theta^*)\psi_j^T(\theta^*) \to \mathbb{E}_{\theta^*}[\psi_1(\theta^*)\psi_1^T(\theta^*)] \quad \mathbb{P}_{\theta^*} - a.s.$$

The limit $I(\theta^*) = \mathbb{E}_{\theta^*}[\psi_1(\theta^*)\psi_1^T(\theta^*)]$ is positive semi-definite, and it is deterministic, so that it is measurable with respect to any σ-algebra. Therefore, applying Proposition B.0.4 and Remark B.0.3 we obtain

Proposition B.0.5. *Assume that Assumption M, R8, and R9 are satisfied. Then,*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} \psi_j(\theta^*) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, I(\theta^*)).$$

APPENDIX C

TECHNICAL SUPPLEMENT

Assumptions R4–R6 are stated for any deterministic vector $\theta \in \mathbf{\Theta}$. In this section, we show that if (3.3)-(3.7) hold for $\theta \in \mathbf{\Theta}$, then for any random vectors $\boldsymbol{\theta}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ that are $\mathcal{F}_{n-1}$ measurable and take values in $\mathbf{\Theta}$, analogous inequalities are true.

Proposition C.0.6. *Assume that R4-R6 are satisfied. Then, for any fixed $n \geq 1$ and for any random vectors $\boldsymbol{\theta}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ that are $\mathcal{F}_{n-1}$ measurable and take values in $\mathbf{\Theta}$, we have*

$$\mathbb{E}_{\theta^*}[\|\psi_n(\boldsymbol{\theta})\|^2 | \mathcal{F}_{n-1}] \leq c(1 + \|\boldsymbol{\theta} - \theta^*\|^2), \tag{C.1}$$

$$(\boldsymbol{\theta} - \theta^*)^T b_n(\boldsymbol{\theta}) \leq -K_1 \|\boldsymbol{\theta} - \theta^*\|^2, \tag{C.2}$$

$$\|b_n(\boldsymbol{\theta})\| \leq K_2 \|\boldsymbol{\theta} - \theta^*\|, \tag{C.3}$$

$$\mathbb{E}_{\theta^*}[\|\Psi_n(\boldsymbol{\theta}_1) - \Psi_n(\boldsymbol{\theta}_2)\| | \mathcal{F}_{n-1}] \leq K_3 \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|, \tag{C.4}$$

$$\mathbb{E}_{\theta^*}[\|\mathsf{H}\psi_n(\boldsymbol{\theta})\| | \mathcal{F}_{n-1}] \leq K_4. \tag{C.5}$$

Proof. We will only show that (C.2) is true. The validity of the remaining inequalities can be proved similarly. Also, without loss of generality, we assume that $d = 1$.

From (3.4), we have for any $\theta \in \mathbf{\Theta}$, $(\theta - \theta^*)\mathbb{E}_{\theta^*}[\psi_n(\theta) \mid \mathcal{F}_{n-1}] \leq K_1|\theta - \theta^*|$. If $\boldsymbol{\theta}$ is a simple random variable, i.e. there exists a partition $\{A_m, 1 \leq m \leq M\}$ of Ω, where M is a fixed integer, such that $A_m \in \mathcal{F}_{n-1}$, $1 \leq m \leq M$, and $\boldsymbol{\theta} = \sum_{m=1}^{M} c_m \mathbb{1}_{A_m}$, where $c_m \in \boldsymbol{\Theta}$. Then, we have that

$$\begin{aligned}
(\boldsymbol{\theta} - \theta^*)\, b_n(\boldsymbol{\theta}) &= (\sum_{m=1}^{M} c_m \mathbb{1}_{A_m} - \theta^*)\mathbb{E}_{\theta^*}[\psi_n(\boldsymbol{\theta})|\mathcal{F}_{n-1}] \\
&= \sum_{m=1}^{M} \mathbb{1}_{A_m}(c_m - \theta^*)\mathbb{E}_{\theta^*}[\mathbb{1}_{A_m}\psi_n(\boldsymbol{\theta})|\mathcal{F}_{n-1}] \\
&= \sum_{m=1}^{M} \mathbb{1}_{A_m}(c_m - \theta^*)\mathbb{E}_{\theta^*}[\mathbb{1}_{A_m}\psi_n(c_m)|\mathcal{F}_{n-1}] \\
&= \sum_{m=1}^{M} \mathbb{1}_{A_m}(c_m - \theta^*)\mathbb{E}_{\theta^*}[\psi_n(c_m)|\mathcal{F}_{n-1}] \\
&\le -\sum_{m=1}^{M} \mathbb{1}_{A_m} K_1 |c_m - \theta^*|^2 = -K_1|\boldsymbol{\theta} - \theta^*|^2.
\end{aligned}$$

From here, using the usual limiting argument we conclude that (C.2) holds true for any $\mathcal{F}_{n-1}$ measurable random variable $\boldsymbol{\theta}$. □

In the rest of this section we will verify that the Assumption M and the properties R0–R9 are satisfies in Example 3.5.1.

It is clear that the Markov chain $\{Z_n, n \ge 0\}$, as defined in Example 3.5.1, satisfies (i) and (iii) in Assumption M. Next we will show that Z is a positive Harris chain (see Definition A.0.3). For any Borel set $B \in \mathcal{B}(\mathbb{R})$ with strictly positive Lebesgue measure, and any $z_0 \in \mathbb{R}$, we have that

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}_{\theta^*}(Z_n \notin B, \ldots, Z_1 \notin B \mid Z_0 = z_0) \\
&\quad = \lim_{n\to\infty} \mathbb{P}_{\theta^*}(Z_n \notin B \mid Z_{n-1} \notin B) \cdots \mathbb{P}_{\theta^*}(Z_2 \notin B \mid Z_1 \notin B)\mathbb{P}_{\theta^*}(Z_1 \notin B \mid Z_0 = z_0) \\
&\quad = \lim_{n\to\infty} \mathbb{P}_{\theta^*}(Z_2 \notin B \mid Z_1 \notin B)^{n-1}\mathbb{P}_{\theta^*}(Z_1 \notin B \mid Z_0 = z_0) = 0,
\end{aligned}$$

and thus Z satisfies Definition A.0.3.(a). Also, since the density (with respect to the Lebesgue measure) of Z_1 is

$$f_{Z_1,\theta^*}(z_1) = \int_{\mathbb{R}} p_{\theta^*}(z_0, z_1) f_{Z_0,\theta^*}(z_0) dz_0 = \frac{1}{\sqrt{2\pi}\sigma^*} e^{-\frac{(z_1-\mu^*)^2}{2(\sigma^*)^2}},$$

then $Z_1 \sim \mathcal{N}(\mu^*, (\sigma^*)^2)$, and consequently, we get that $Z_n \sim \mathcal{N}(\mu^*, (\sigma^*)^2)$ for any $n \geq 0$. This implies that $\mathcal{N}(\mu^*, (\sigma^*)^2)$ is an invariant distribution for Z. Thus, Z is a positive Harris chain, and respectively, by Proposition A.0.3, Z is an ergodic process.

As far as propreties R0–R9, we fist note that

$$\begin{aligned}
\psi_n(\theta) &= \nabla \log p_\theta(Z_{n-1}, Z_n) \\
&= \Big(\frac{Z_n - \rho Z_{n-1} - (1-\rho)\mu}{\sigma^2(1+\rho)}, -\frac{1}{\sigma} + \frac{(Z_n - \rho Z_{n-1} - (1-\rho)\mu)^2}{(1-\rho^2)\sigma^3}\Big)^T, \\
b_n(\theta) &= \mathbb{E}_{\theta^*}[\psi_n(\theta) | \mathcal{F}_{n-1}] \\
&= \Big(-\frac{(1-\rho)(\mu-\mu^*)}{\sigma^2(1+\rho)}, \frac{\sigma^{*,2} - \sigma^2}{\sigma^3} + \frac{(1-\rho)(\mu-\mu^*)^2}{(1+\rho)\sigma^3}\Big)^T, \\
\Psi_n(\theta) &= \begin{bmatrix} -\frac{1-\rho}{(1+\rho)\sigma^2} & -\frac{2(Z_n - \rho Z_{n-1} - (1-\rho)\mu)}{(1+\rho)\sigma^3} \\ -\frac{2(Z_n - \rho Z_{n-1} - (1-\rho)\mu)}{(1+\rho)\sigma^3} & \frac{1}{\sigma^2} - \frac{3(Z_n - \rho Z_{n-1} - (1-\rho)\mu)^2}{(1-\rho^2)\sigma^4} \end{bmatrix}.
\end{aligned}$$

We denote by $Y_n := Z_n - \rho Z_{n-1} - (1-\rho)\mu$, and we immediately deduce that that

$$\begin{aligned}
\mathbb{E}_{\theta^*}[Y_n \mid \mathcal{F}_{n-1}] &= (1-\rho)(\mu^* - \mu), \\
\mathbb{E}_{\theta^*}[Y_n^2 \mid \mathcal{F}_{n-1}] &= (1-\rho)^2(\mu - \mu^*)^2 + (\sigma^*)^2(1-\rho^2), \\
\mathbb{E}_{\theta^*}[Y_n^4 \mid \mathcal{F}_{n-1}] &= (1-\rho)^4(\mu^* - \mu)^4 + 6(1+\rho)(1-\rho)^3(\mu^* - \mu)^2(\sigma^*)^2 \\
&\quad + 3(\sigma^*)^4(1-\rho^2)^2.
\end{aligned} \tag{C.6}$$

From here, and using the fact that $\boldsymbol{\Theta}$ is bounded, it is straightforward, but tedious,[14] to show that R4, R5, R6, and R7 are satisfied. Also, it is clear note that R0 is true, and using (C.6) by direct computations we get that R1 and R2 are satisfied.

Since

$$\mathbb{E}_{\theta^*}[\psi_1(\theta)] = \left(\frac{(1-\rho)(\mu^* - \mu)}{\sigma^2(1+\rho)}, \frac{(\sigma^*)^2 - \sigma^2}{\sigma^3} + \frac{(1-\rho)(\mu - \mu^*)^2}{(1+\rho)\sigma^3}\right),$$

then R3 is clearly satisfied.

[14]The interested reader can contact the authors for details.

Again by direct evaluations, we have that

$$I(\theta) = \mathbb{E}_\theta[\psi_1(\theta)\psi_1(\theta)^T] = \begin{bmatrix} \frac{1-\rho}{(1+\rho)\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix},$$

which is positive definite matrix, and thus R8 is satisfied.

Finally, we will verify R9. By Jensen's inequality and Cauchy-Schwartz inequality, we have that

$$\begin{aligned} \exp\left(\mathbb{E}_{\theta^*} \sup_{0\le i\le n} |\psi_i(\theta^*)|\right) &\le \mathbb{E}_{\theta^*} \exp\left(\sup_{0\le i\le n} |\psi_i(\theta^*)|\right) = \mathbb{E}_{\theta^*}\left[\sup_{0\le i\le n} \exp|\psi_i(\theta^*)|\right] \\ &\le \sum_{i=1}^n \mathbb{E}_{\theta^*} \exp|\psi_i(\theta^*)| \\ &\le \sum_{i=1}^n \mathbb{E}_{\theta^*} \exp\left(\frac{|Y_i|}{\sigma^2(1+\rho)} + \frac{1}{\sigma} + \frac{Y_n^2}{(1-\rho)^2\sigma^3}\right) \\ &\le \sum_{i=1}^n \left(\mathbb{E}_{\theta^*} \exp(\frac{2|Y_i|}{\sigma^2(1+\rho)})\right)^{\frac{1}{2}} \left(\mathbb{E}_{\theta^*} \exp(\frac{2}{\sigma} + \frac{2Y_i^2}{(1-\rho)^2\sigma^3})\right)^{\frac{1}{2}}. \end{aligned}$$

Note that for $Y_i, i = 0, \ldots, n$ is normally distributed, and therefore, there exist two constants C_1 and C_2, that depend on θ^* such that

$$\mathbb{E}_{\theta^*} \exp\left(\frac{2|Y_i|}{\sigma^2(1+\rho)}\right) = C_1, \quad \mathbb{E}_{\theta^*} \exp\left(\frac{2}{\sigma} + \frac{2Y_i^2}{(1-\rho)^2\sigma^3}\right) = C_2.$$

Hence, we have that

$$\mathbb{E}_{\theta^*} \sup_{0\le i\le n} |\psi_i(\theta^*)| \le \log n + \frac{1}{2}\log C_1 C_2,$$

and, thus R9 is satisfied:

$$\lim_{n\to\infty} \mathbb{E}_{\theta^*}\left[\sup_{0\le i\le n}\left|\frac{1}{\sqrt{n}}\psi_i(\theta^*)\right|\right] \le \lim_{n\to\infty}\left(\frac{\log n}{\sqrt{n}} + \frac{\log C_1C_2}{2\sqrt{n}}\right) = 0.$$

BIBLIOGRAPHY

[ADEH99] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[AE78] D. J. Aldous and G. K. Eagleson. On Mixing and Stability of Limit Theorems. *The Annals of Probability*, 6(2):325–331, April 1978.

[AP11] B. Acciaio and I. Penner. Dynamic risk measures. *In G. Di Nunno and B. Øksendal (Eds.), Advanced Mathematical Methods for Finance, Springer*, pages 1–34, 2011.

[AS08] C. Acerbi and G. Scandolo. Liquidity risk theory and coherent measures of risk. *Quant. Finance*, 8(7):681–692, 2008.

[BBN14] S. Biagini and J. Bion-Nadal. Dynamic quasi-concave performance measures. *Journal of Mathematical Economics*, 55:143–153, December 2014.

[BCC15] T.R. Bielecki, I. Cialenco, and T. Chen. Dynamic conic finance via Backward Stochastic Difference Equations. *SIAM J. Finan. Math.*, 6(1):1068–1122, 2015.

[BCC16a] T.R. Bielecki, I. Cialenco, and T. Chen. Recursive construction of confidence regions. *submitted for publication*, 2016.

[BCC+16b] T.R. Bielecki, I. Cialenco, T. Chen, A. Cousin, and M. Jeanblanc. Adaptive Robust Hedging Under Model Uncertainty. In preparation, 2016.

[BCDK15] T.R. Bielecki, I. Cialenco, S. Drapeau, and M. Karliczek. Dynamic assessment indices. *Forthcoming in Stochastics: An International Journal of Probability and Stochastic Processes*, 2015.

[BCH+00] P. Briand, F. Coquet, Y. Hu, J. Memin, and S. Peng. A converse comparison theorem for BSDEs and related properties of g-expectation. *Electronic Communications in Probability*, 5:101–117, 2000.

[BCIR13] T.R. Bielecki, I. Cialenco, I. Iyigunler, and R. Rodriguez. Dynamic Conic Finance: Pricing and hedging via dynamic coherent acceptability indices with transaction costs. *International Journal of Theoretical and Applied Finance*, 16(01):1350002, 2013.

[BCP14] T. R. Bielecki, I. Cialenco, and M. Pitera. A unified approach to time consistency of dynamic risk measures and dynamic performance measures in discrete time. *Preprint*, 2014.

[BCP15] T.R. Bielecki, I. Cialenco, and M. Pitera. Time consistency of dynamic risk measures and dynamic performance measures in discrete time: LM-measure perspective. *Preprint*, 2015.

[BCR15] T.R. Bielecki, I. Cialenco, and R. Rodriguez. No-arbitrage pricing theory for dividend-paying securities in discrete-time markets with transaction costs. *Mathematical Finance*, 25(4):673–701, 2015.

[BCZ14] T.R. Bielecki, I. Cialenco, and Z. Zhang. Dynamic coherent acceptability indices and their applications to finance. *Mathematical Finance*, 24(3):411–441, 2014.

[BEK07] P. Barrieu and N. El Karoui. Pricing, hedging and optimally designing derivatives via minimization of risk measures. In R. Carmona, editor, *Indifference Pricing*. Princeton University Press, 2007.

[Bir31] G. Birkhoff. Proof of the Ergodic Theorem. *Proc. Natl. Acad. Sci. USA*, 17(12):656–660, 1931.

[BK02] Robert Buche and Harold J. Kushner. Rate of convergence for constrained stochastic approximation algorithms. *SIAM J. Control Optim.*, 40(4):1011–1041, 2001/02.

[BN09] J. Bion-Nadal. Bid-ask dynamic pricing in financial markets with transaction costs and liquidity risk. *Journal of Mathematical Economics*, 45(11):738–750, December 2009.

[BNS94] O.E. Barndorff-Nielsen and M. Sørensen. A review of some aspects of asymptotic likelihood theory for stochastic processes. *International Statistical Review*, 62(1):133–165, 1994.

[CE10] S.N. Cohen and R.J. Elliott. A general theory of finite state backward stochastic difference equations. *Stochastic Processes and their Applications*, 120(4):442–466, April 2010.

[CE11] S.N. Cohen and R.J. Elliott. Backward stochastic difference equations and nearly-time-consistent nonlinear expectations. *SIAM J. Control Optim., 49(1), 125139.*, 2011.

[CHMP02] F. Coquet, Y. Hu, J. Mémin, and S. Peng. Filtration-consistent nonlinear expectations and related g-expectations. *Probab. Theory Related Fields*, 123(1):1–27, 2002.

[CM09] A. Cherny and D.B. Madan. New measures for performance evaluation. *The Review of Financial Studies*, 22(7):2571–2606, 2009.

[CM10] A. Cherny and D.B. Madan. Markets as a counterparty: An introduction to conic finance. *International Journal of Theoretical and Applied Finance (IJTAF)*, 13(08):1149–1177, 2010.

[CP05] I. Crimaldi and L. Pratelli. Convergence Result for Multivariate Martingales. *Stoch. Proc. and their Applic.*, 115:571–577, 2005.

[CS13] P. Cheridito and M. Stadje. BSDeltaEs and BSDEs with non-Lipschitz drivers: Comparison, convergence and robustness. *Bernoulli*, 19:1047–1085, 2013.

[CSR00] J. Cochrane and J. Saa-Requejo. Beyond arbitrage: Good deal asset price bounds in incomplete markets. *Journal of Policital Economy*, 108:79–119, 2000.

[DK13] S. Drapeau and M. Kupper. Risk preferences and their robust representation. *Mathematics of Operations Research*, 38(1):28–62, February 2013.

[EHR89] J.-E. Englund, U. Holst, and D. Ruppert. Recursive estimators for stationary, strong mixing processes - a representation theorem and asymptotic distributions. *Stochastic Processes Appl.*, 31:203–222, 1989.

[EKQ97] Peng S. El Karoui, N. and M.C. Quenez. Backward stochastic differential equation in finance. *Mathematical Finance*, 7/1:1–71, 1997.

[ESC15] R.J. Elliott, T.K. Siu, and S. N. Cohen. Backward stochastic difference equations for dynamic convex risk measures on a binomial tree. *Journal of Applied probability, 2015, 52(3)*, 2015.

[Fab68] V. Fabian. On Asymptotic Normality in stochastic approximation. *Ann. Math. Statist.*, 39(4):1327–1332, 1968.

[Fab78] V. Fabian. On asymptotically efficient recursive estimation. *Ann. Statist.*, 6:854–867, 1978.

[Fis25] R. A. Fisher. Theory of statistical estimation. *Proc. Cambridge Philos. Soc.*, 22:700–725, 1925.

[FKV09] D. Filipovic, M. Kupper, and N. Vogelpoth. Separation and duality in locally L^0-convex modules. *Journal of Functional Analysis*, 256:3996 – 4029, 2009.

[FS04] H. Föllmer and A. Schied. *Stochastic finance. An introduction in discrete time*, volume 27 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2004.

[Hey97] C. Heyde. *Quasi-Likelihood And Its Application: A General Approach to Optimal Parameter Estimation*. Springer Series in Statistics. Springer-Verlag New York, 1997.

[HL15] E. Häusler and H. Luschgy. *Stable Convergence and Stable Limit Theorems*. Springer, 2015.

[HLL00] O. Hernández-Lerma and J. Lasserre. Further criteria for positive Harris recurrence of Markov chains. *Proceedings of the American Mathematical Society*, 129(5):1521–1524, 2000.

[KC78] H. Kushner and D. Clark. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, volume 26 of *Applied Mathematical Sciences*. Springer-Verlag New York, 1978.

[KN76] R.Z. Khas'minskii and M.B. Nevelson. *Stochastic Approximation and Recursive Estimation*, volume 47 of *Translations of Mathematical Monographs*. American Mathematical Society, 1976.

[KV09] M. Kupper and N. Vogelpoth. Complete L^0-Normed Modules and Automatic Continuity of Monotone Convex Functions. *VIF Working Papers Series*, 2009.

[KW52] J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, 23(3):462–466, 1952.

[KY03] H. Kushner and G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*, volume 35 of *Stochastic Modelling and Applied Probability*. Springer-Verlag New York, 2003.

[LeC56] L. LeCam. On the asymptotic theory of estimation and testing hypothesis. *Proccedings in the Third Berkely Symposium on Mathematical Statist. and Probab.*, 1:129–156, 1956.

[LeC60] L. LeCam. Locally asymptotic normal families of distributions. *Univ. California Publ. in Statist.*, 3(2):37–98, 1960.

[LR79] T. L. Lai and H. Robbins. Adaptive design and stochastic approximation. *Ann. Statist.*, 7(6):1196–1221, 1979.

[LS87] L. Ljung and Söderström. *Theory and Practice of Recursive Identification*, volume 4 of *The MIT Press Series in Signal Processing, Optimization, and Control.* MIT Press, 1987.

[MT93] S. Meyn and R. Tweedie. *Markov Chains and Stochastic Stability.* Cambridge Mathematical Library. Cambridge University Press, 1993.

[Pen97] S. Peng. Backward SDE and related g-expectations. *Backward stochastic differential equations, Pitman Research Notes in Mathematics Series*, 364:141–159, 1997.

[PP90] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 12:161–185, 1990.

[PR99] B.L.S. Prakasa Rao. *Semimartingales and their Statistical Inference*, volume 83 of *Chapman & Hall/CRC Monographs on Statistics & Applied Probability.* CRC Press, 1999.

[RG06] E. Rosazza Gianin. Risk measures via g-expectations. *Insurance: Mathematics and Economics*, 39(1):19–34, August 2006.

[RGS13] E. Rosazza Gianin and E. Sgarra. Acceptability indexes via g-expectations: an application to liquidity risk. *Mathematical Finance (2013) 7:457-475*, 2013.

[RM51] H. Robbins and S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22(3):400–407, 1951.

[Sac58] J. Sacks. Asymptotic distribution of stochastic approximation procedures. *Ann. Math. Statist.*, 29(2):373–405, 1958.

[Sha98] T. Sharia. On the recursive parameter estimation for the general discrete time statistical model. *Stochastic Processes Appl*, 73(2):151–172, 1998.

[Sha10] T. Sharia. Recursive parameter estimation: asymptotic expansion. *Annals of the Institute of Statistical Mathematics*, 62(2):343–362, 2010.

[Shi84] A. Shiryayev. *Probability*, volume 95 of *Graduate Texts in Mathematics.* Springer New York, 1984.

[Sta09] M. Stadje. Dynamic risk measures and BSDEs: from discrete to continuous time. *PhD Thesis*, 2009.

[Var01] S. R. S. Varadhan. *Probability Theory*, volume 7 of *Courant Lecture Notes.* American Mathematical Society, 2001.

[vN32a] J. von Neumann. Physical applications of the ergodic hypothesis. *Proc. Natl. Acad. Sci. USA*, 18(3):263–266, 1932.

[vN32b] J. von Neumann. Proof of the quasi-ergodic hypothesis. *Proc. Natl. Acad. Sci. USA*, 18(1):70–82, 1932.