Selected problems on discrete time stochastic control for dynamic risk and performance measures.

PhD Dissertation

Marcin Pitera

Advisor:
prof. dr hab. Łukasz Stettner

Jagiellonian University
Faculty of Mathematics and Informatics
Institute of Mathematics

August, 2014

Slightly modified version. (Date: November 6, 2015.)

Contents

Chapter 1

Introduction

The terms *risk* and *performance* are now a part of everyday language. Despite the fact, that the meaning of those words is clear on intuitive level, no good unifying definition has been proposed so far. One can even say that those words are closer to primitive notions, than undefined terms.

Let us mention just a few words, such as preference, hazard, fortune, safety, opportunity or uncertainty, which strongly relates to both *risk* and *performance*, allowing a wide spectrum of non-equivalent interpretations.

From the mathematical point of view, we are more interested in quantifying *risk* and *performance*, i.e. defining the proper methods of measurement. Through time, a lot of beautiful concepts were introduced to answer this question. Let us alone mention people such as Daniel Bernoulli, John von Neumann or Oskar Morgenstern who made a major contribution to this subject. As the description of evolution of mathematical methods quantifying *risk* and *performance*, or more generally speaking *preference* and *utility*, is not the main topic of this thesis, we refer interested reader to [70, 102, 77] and references therein for a a good overview on this subject and historic survey.

In this work we will focus on quantifying *the utility* (this term will cover both *risk* and *performance*) of a financial position and follow the modern (normative) approach, suggested e.g. by Artzner, Delbaen, Eber, and Heath [6]. The notation of *coherent risk measures* proposed in [6] attracted attention both on theoretical and practical level resulting in the development of many objects such as *convex risk measures* [76] or *acceptability indices* [45], which we will introduce in this thesis.

At the beginning, the theory of *risk measures* was developed in the static framework. On a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the (real valued) random variables coincided with the future (discounted) values of financial positions and risk measure was a map which assigned a number to any financial position, measuring it's risk. The normative approach from [6] required this map to satisfy certain sets of axioms such as *cash-additivity*, *convexity* or *monotonicity*. Those axioms have been also given a clear financial interpretation. The real values produced by a coherent risk measures could be also interpreted as minimal capital requirements sufficient for the financial position to be acceptable, i.e. having non-positive risk. The notation of *utility measures*, often considered as negatives of risk measures was very similar. The similar results also hold true for risk and utility measures quantifying the risk of cash-flows of some financial position, which are typically described by a stochastic process [36].

For convex risk measures (defined for random variables) satisfying certain regularity conditions, the duality theory of Fenchel-Legandre lead to so called *robust representation*, which allowed the risk measure ρ, to be expressed as

$$\rho(X) = -\inf_{\mathbb{Q}\in\mathcal{M}}[E_{\mathbb{Q}}[X] - \alpha(\mathbb{Q})],$$

for the set of all probability measures $\mathcal{M}$ and the penalty function $\alpha : \mathcal{M} \to [-\infty, \infty]$ [75]. In other words, the risk of a financial position could be considered as the penalised worst case expectation taken under probabilistic model $\mathcal{M}$, where the penalty function α describes how seriously we treat each scenario from $\mathcal{M}$.

The need to consider the evolution of preferences through time $\mathbb{T}$ lead to the definition of conditional and dynamic risk measures (and utility measures) [7]. The information available at time $t \in \mathbb{T}$ could be described by a σ-subfield of $\mathcal{F}$, denoted by $\mathcal{F}_t$, and the whole evolution of market is given by a filtered probability space, where the filtration is given by $\{\mathcal{F}_t\}_{t\in\mathbb{T}}$. Then, the conditional risk measure ρ_t is a mapping which assignes a $\mathcal{F}_t$-measurable random variable to any financial position, measuring it's conditional risk. The dynamic risk measure is simply a collection of conditional risk measures, i.e. the family of mappings $\{\rho_t\}_{t\in\mathbb{T}}$ [2]. Besides the conditional equivalents of properties from the static definition of risk measures, one has to introduce additional axioms, which will link preferences from different time points, i.e. explain the relation between $\rho_t(X)$ and $\rho_s(X)$ for any financial position X and time points $t, s \in \mathbb{T}$. Such axioms are related to so called *time-consistency conditions* [2], which will be one of the main topics of this thesis.

Dynamic utility measures very often act as objective functions in the stochastic control problems, while dynamic risk measures serve as constraints [109]. The need to consider their dynamic equivalents is also justified by the use of the dynamic programming approach [7]. Bellman principle of optimality, needed for the backward recursive reformulation of stochastic control problems, is tightly connected with time consistency condition, as will be shown later. Let us mention the fact, that in some types of optimisation problems, time-inconcistency (and thus the lack of the standard Bellman principle of optimality) could be present (cf. [26]) but we will not focus on those problems here.

In this thesis, we will focus on three types of utility maps, namely convex risk measures [75], acceptability indices [45] and limit growth indices [18], as they are often used in stochastic control problems. We have decided to show some representatives of each class, to explain why those classes are interesting.

Apart from theoretical study about time constancy, we have decided to show three explicit representative examples, which will show how to perform optimisation, using dynamic risk and performance measures. The first two examples will be connected to portfolio optimisation, while the last one will cover the subject of pricing an American options, by solving an optimal stopping problem.

This thesis is organized as follows. The Introduction will be followed by Chapter 2, where we shall introduce the basic framework, notations and definitions used throughout the whole thesis. Next, in Chapter 3 we will focus on time-consistency conditions, providing a new definition for this type of objects and linking it with the existing literature. Chapter 4 will be dedicated to the study of various specific families of dynamic risk and performance measures, namely dynamic convex and coherent risk measures, dynamic acceptability indices and dynamic limit growth indices. We will provide also many explicit examples with the emphasis on time consistency. Finally, in Chapter 5 we will deal with three stochastic control problems. The first problem will be connected with risk sensitive control for Markov decision processes on infinite time horizon. The second one will cover the classical risk-to-reward optimisation, using coherent risk measure as a constraint. Finally the last example will provide an explicit algorithm, which can numerically approximate conditional expectation and Snell envelopes, providing a recipe for pricing various types of American options.

Contribution

Three papers [18, 19, 99] could be seen as a main building blocks of this dissertation.

In particular, Chapter 3 is based on the first paper, *A unified approach to time consistency of dynamic risk measures and dynamic performance measures in discrete time*, written together with prof. T. R. Bielecki and I.Cialenco. Next, Section 4.3 is based on *Dynamic Limit Growth Indices in Discrete Time*, written together with prof. T. R. Bielecki and prof. I. Cialenco. Finally, Section 5.3 is based on the last paper, *The least squares method for option pricing revisited*, written together with prof. M. Klimek.

Our feeling is that our contribution to the subject of dynamic risk and performance measurement as well as dynamic stochastic control theory is mainly through the depth of construction, which unifies those theories and provide the space for new interesting objects as well as shed a new light on the old ones. While almost all of the proofs from Chapters 3 and 4 are rather elementary, i.e. not requiring any advanced mathematical tools and methods, we think they provide a nice framework for future development. Nevertheless, let us mention some of the results, which we consider noteworthy.

The new definition (Definition 3.1.3) of time-consistency, also introduced in [19] is promising, in the sense that it allows to handle both cash-additive and scale-invariant maps using the same tools. The difference between those two maps caused a lot of problems for many people, as the benchmark definition of time-consistency provided only limited answers.

The Dynamic equivalent of Risk Sensitive Criterion defined in [18] is an interesting object, which requires further attention. The proof that Risk Sensitive Criterion is in fact an acceptability index surprisingly was not present in the literature. We think, that the proof of supermartingale property for this map, i.e. Thoerem 4.3.11, 5), is interesting and non-trivial. It also show a nice application of a variation of conditional Borel-Cantelli lemma, exploiting the strength of power utility transformation.

While the proofs from Section 5.1 are present in the literature, we think that our reformulations can shed a new light on risk sensitive control problems, connecting them with risk measurement theory. In particular, our presentation of Bellman equation (5.10) and related results seem to be more intuitive than the usual formulation, which relates to Multiplicative Poisson Equations.

Apart from results mentioned above, we have decided to list some Propositions and Theorems, which we think contribute to the theory of dynamic risk measurement and for which the proofs are noteworthy, namely: Proposition 3.2.1, Proposition 3.2.13, Proposition 4.2.3, Proposition 4.2.4, Proposition 4.3.2, Theorem 4.3.11, Proposition 5.1.6, Theorem 5.3.3 and Theorem 5.3.6.

Acknowledgments / Thanks

In the beginning I would like to thank my advisor Łukasz Stettner for his support and patience. His profound expertise in the area of financial mathematics together with generosity in sharing has inspired my interest in this field. His talent to explain mathematical issues in simple terms and the ability to recognise quickly the essence of the problem has been very helpful during my work on this thesis, while the rigorous approach to mathematics helped to temper my impatience and imprecision. I am very grateful for many deep and inspiring discussions, from which I have learned a lot.

Additionally, I would like to thank Tomasz R. Bielecki and Igor Cialenco for our collaboration during my stay at Illinois Institute of Technology. Their guidance, support, enthusiasm, devotion to research as well as mathematical precision helped me a lot during the work on this thesis. I am also very grateful for bringing my attention to the subject of dynamic performance measurement and time consistency axioms. In fact, the papers about those subjects, written together by us, are the main building blocks of this thesis.

Furthermore, I sincerely thank Maciej Klimek for many interesting and fruitful discussions during my stay at Uppsala University. His expertise in the area of functional analysis and approximation techniques made me understand more the beauty and depth of mathematics. I would also like to thank him for our joint work, which resulted in the paper included in this thesis.

Finally, I also acknowledge the support by Project operated within the Foundation for Polish Science IPP Programme "Geometry and Topology in Physical Models" co-financed by the EU European Regional Development Fund, Operational Program Innovative Economy 2007-2013.

Chapter 2

Preliminaries

In this Chapter we will introduce basic concepts and ideas from dynamic risk measure theory and related fields. In particular, we will provide some general comments about certain aspects of the stochastic control theory and convex analysis, needed in this thesis. We will also introduce here the basic framework, i.e. definitions and notation, used throughout this thesis. Let us underline the fact that only selected aspects of dynamic risk measure theory are considered here. For a good survey about risk measures we refer to [77] and references therein.

Almost all results from this section are well known and used commonly in the field of dynamic risk measurement. While we will give definitions, which are more general than those usually seen in the literature, the generalisations are rather intuitive and straightforward. Nevertheless, we will provide proofs for statements, which have no direct counterparts in the existing literature.

In particular, Propositions 2.2.13 and 2.2.15 due to our knowledge did not occur in this form in the literature before, and could be seen as an original contribution. Despite this fact, we have decided to put them here, as they are very simple and the proofs are elementary.

This Chapter is organized as follows. In Section 2.1 we provide a set of some general underlying concepts that will be used throughout the thesis. In particular we introduce the space on which we will define all objects, give some comments about conditional expectations and recall definitions of conditional equivalents of essential infimum and supremum, for various types of objects.

Section 2.2 recalls basic properties of maps which try to quantify (measure) the risk or performance of objects describing terminal payoffs, portfolio cash-flows or dividend streams. Those objects are usually represented as random variables or adapted stochastic processes. We will also give some comment about how to extend such maps onto bigger space and give some insight about robust representations of risk and performance measures.

Next, in Section 2.3 we give a general overview about dynamic stochastic control problems which naturally arise, when we study maps introduced in the previous section.

Finally, in Section 2.4 we give some additional remarks concerning all objects introduced in the previous Sections in this Chapter. This Section will try to give some overview on the literature and present some new results and ideas, which were introduced recently. We put here all the remarks, which are not required in the main body of this Chapter. It was done to not confuse the reader with too many auxiliary facts. Still, they could help understand the objects, which were introduced before and shed some new light on them.

As we have mentioned above, the main purpose of this Chapter is to provide basic definitions and notation, which we will use later. Thus, we have decided to postpone all proofs from this

Chapter to Appendix, in order to make everything more transparent.

2.1 General framework

Throughout this thesis, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{T}}, \mathbb{P})$ will denote a discrete-time filtered probability space. We will consider both finite and infinite time horizons, i.e $\mathbb{T} = \{0, 1, \ldots, T\}$, for a fixed $T \in \mathbb{N}$, or $\mathbb{T} = \mathbb{N} \cup \{0\}$. We will also assume that $\mathcal{F}_0$ is trivial and $\mathcal{F} = \bigcup_{t\in\mathbb{T}} \mathcal{F}_t$. Furthermore, we will assume that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{T}}, \mathbb{P})$ is a *standard* filtered probability space (*Lebesgue-Rokhlin* probability space [123]), i.e. it is (a.s.) isomorphic to $([0,1]^{\mathbb{T}}, \mathcal{B}([0,1]^{\mathbb{T}}), \{\mathcal{F}'_t\}_{t\in\mathbb{T}}, \lambda^{\mathbb{T}})$, where $\mathcal{B}(\cdot)$ denotes the Borel σ-algebra of considered set[1], $\lambda^{\mathbb{T}}$ is a product of the Borel measures and $\{\mathcal{F}'_t\}_{t\in\mathbb{T}}$ is the filtration generated by the coordinate functions (see e.g. [103] for similar settings in the risk measure framework). While almost all of the results could be easily reformulated for the general case, we make this assumption to avoid many technical difficulties. In particular note that we could define conditional (regular) probabilities, through *Canonical system of measures* and $(\Omega, \mathcal{F}, \mathbb{P})$ always contain countably separated[2] subset of full measure (cf. [123, 127] and references therein for more detailed description of standard probability spaces). For clarity we will also assume that the (original) filtration is completed with sets of measure zero.[3]

For σ-algebra $\mathcal{G}$, such that $\mathcal{G} \subseteq \mathcal{F}$, we denote by $L^0(\Omega, \mathcal{G}, \mathbb{P})$ and $\bar{L}^0(\Omega, \mathcal{G}, \mathbb{P})$ the sets of all (a.s. identified) $\mathcal{G}$-measurable random variables with values in $(-\infty, \infty)$ and $[-\infty, \infty]$, respectively. Moreover, we define $\widehat{L}^p(\Omega, \mathcal{G}, \mathbb{P}) := \{X \in \bar{L}^0(\Omega, \mathcal{G}, \mathbb{P}) \mid (X \vee 0) \in L^p(\Omega, \mathcal{G}, \mathbb{P})\}$, for $p \in [0, \infty]$. We shall write $L^p := L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $L^p_t := L^p(\Omega, \mathcal{F}_t, \mathbb{P})$, for $p \in [0, \infty]$. Analogous definitions will apply to $\bar{L}^0$ and $\widehat{L}^p$. The financial interpretation of the elements of L^p_t (also $\widehat{L}^p_t$ and $\bar{L}^0_t$) will depend on the context. Usually they will illustrate terminal payoffs of some (zero cost) financial portfolio or cash-flows at time t. They might also correspond to the preference level. We will also use notation

$$\mathbb{V}^p := \{(V_t)_{t\in\mathbb{T}} \mid V_t \in L^p_t,\ t \in \mathbb{T}\}, \tag{2.1}$$
$$\mathbb{V}^p_{\ln} := \{(V_t)_{t\in\mathbb{T}} \mid \ln V_t \in L^p_t,\ V_t > 0,\ t \in \mathbb{T}\}, \tag{2.2}$$
$$\mathbb{V}^p_{\tau} := \{(V_t)_{t\in\mathbb{T}} \mid V_t \in L^p_t,\ V_t \geq 0,\ \text{and } V_t = V_{t\wedge\tau^V},\ t \in \mathbb{T}\}, \tag{2.3}$$

where $\tau^V := \inf\{t \in \mathbb{T} \mid V_t = 0\}$ and $p \in [0, \infty]$, to define various spaces of adapted stochastic processes. Similarly as in the previous case, we also define $\bar{\mathbb{V}}^0$, $\widehat{\mathbb{V}}^p$, etc. The financial intuition going behind (2.1), (2.2) and (2.3) will also depend on the context. Usually $\mathbb{V}^p_{\ln}$ and $\mathbb{V}^p_{\tau}$ will denote (cumulative) value processes of portfolios of financial securities, where the stopping time condition relates to bankrupcy event. Moreover, the elements of $\mathbb{V}^p$ can be viewed e.g. as cash-flows corresponding to some investment strategy or dividend stream of some financial portfolio.

Throughout this thesis, $\mathcal{X}$ will always denote the certain space of random variables or adapted stochastic processes, described in the previous paragraph. All equalities and inequalities will be understood in the almost sure sense. Let us now present some additional notation used for $\mathcal{X}$, which will allow us to keep the notation uniform:

---

[1] We will use notation $\mathcal{B}(\bar{\mathbb{R}}) := \{B \cup S :\ B \in \mathcal{B}(\mathbb{R}),\ S \in \{\emptyset, \{+\infty\}, \{-\infty\}, \{+\infty, -\infty\}\}$.

[2] i.e. there exists a sequence of sets $\{A_n\}_{n\in\mathbb{N}}$, such that $\{n :\ x \in A_n\} = \{n :\ y \in A_n\}$ implies $x = y$, for any $x, y \in \Omega$.

[3] For simplicity, in the static case (i.e. for $t = 0$), with slight abuse of notation, we will say that $\mathcal{F}_0$-adapted random variables are constants. Consequently, we will use $\mathbb{R}$ and $\bar{\mathbb{R}}$ to denote corresponding spaces of random variables, rather than defining the spaces of equivalent classes of random variables as in all other cases.

(1) If $L^\infty \subseteq \mathcal{X} \subseteq \bar{L}^0$, for $t \in \mathbb{T}$, we shall write

$$\begin{aligned}
\bar{\mathcal{X}} &= \bar{L}^0,\\
\widehat{\mathcal{X}} &= \{X \in \bar{\mathcal{X}} \mid (X \vee 0) \in \mathcal{X}\},\\
\mathcal{X}_t &= \mathcal{X} \cap \bar{L}^0_t,\\
1_{\{t\}} &= 1.
\end{aligned}$$

Moreover, for $X, X' \in \mathcal{X}$, we shall write $X \leq X'$, if $\mathbb{P}[X \leq X'] = 1$. In other words, we consider the usual almost sure order for random variables.

(2) If $\mathbb{V}^\infty \subseteq \mathcal{X} \subseteq \bar{\mathbb{V}}^0$, we shall write

$$\begin{aligned}
\bar{\mathcal{X}} &= \bar{\mathbb{V}}^0,\\
\widehat{\mathcal{X}} &= \{(V_t)_{t\in\mathbb{T}} \in \bar{\mathcal{X}} \mid (V_t \vee 0)_{t\in\mathbb{T}} \in \mathcal{X}\},\\
\mathcal{X}_t &= \{X \in \bar{L}^0_t | X = V_t,\ \text{for some } V \in \mathcal{X}\},\\
1_{\{t\}} &= (\underbrace{0, 0, \ldots, 0}_{t}, 1, 0, 0, \ldots).
\end{aligned}$$

Moreover, for $V, V' \in \mathcal{X}$, we shall write $V \leq V'$, if $\mathbb{P}[V_t \leq V'_t] = 1$, for all $t \in \mathbb{T}$.

We will often assume that the space $\mathcal{X}$ is embedded with certain topology. For example, if $p \in [1, \infty]$, we will assume that L^p is a standard Banach space, and the topology is induced by the norm $\|\cdot\|_p$. On the other hand, if L^0, we will use the topology of convergence in measure, i.e. topology generated by metric $d(X, Y) = E[|X - Y| \wedge 1]$ $(X, Y \in L^0)$. See Appendix A.3 for more detailed description.

Following e.g. [44, 38], while working in risk measure framework and using tools from convex analysis, we will adapt the convention

$$\infty - \infty = -\infty \quad \text{and} \quad 0 \cdot \pm\infty = 0. \tag{2.4}$$

Remark 2.1.1. Convention (2.4) has a financial interpretation. We will work with (utility) functions which illustrate preferences of portfolios, cash-flows, etc. For example, if the negative part of random variable is unpredictable (e.g. in a sense that the conditional expectation of the negative part is infinite), then such portfolios are of no interest for us, however big the positive part is. Consequently, our preference level for such random variables should be equal to $-\infty$. See also (2.6), where the generalized (robust) conditional expectation is defined explicitly.

Remark 2.1.2. In this thesis (to get rid of various minus signs) we will mostly work with mappings which admit (quasi)concavity, the reason why we introduce the convention (2.4). In particular, (quasi)convex maps will be treated as negatives of (quasi)concave maps. If we wanted to consider everything in convex framework, then the convention $\infty - \infty = \infty$ should be applied. See [143, 122] for more details about this convention in convex framework.

In particular, with (2.4) in mind, we will use standard (coordinate wise) additive and multiplicative operators both for random variables and adapted stochastic processes. Furthermore, we define multiplicative operator $(\cdot_t)$ by

$$\begin{aligned}
m \cdot_t X &= mX, && X \in \bar{L}^0,\ m \in \bar{L}^0_t,\\
m \cdot_t V &= (V_0, \ldots, V_{t-1}, mV_t, mV_{t+1}, \ldots), && V \in \bar{\mathbb{V}}^0,\ m \in \bar{L}^0_t.
\end{aligned} \tag{2.5}$$

In the case of random variables, operator $(\cdot_t)$ coincides with the standard multiplicative operator, but we introduce the notation to keep the definitions uniform. See Remark 2.4.1 for additional comments about operator $(\cdot_t)$.

2.1.1 Generalized conditional expectation

For $t \in \mathbb{T}$ and $X \in \bar{L}^0$, using (2.4), we define the (generalized) conditional expectation of X by

$$\begin{aligned} E[X|\mathcal{F}_t] &:= E[X^+|\mathcal{F}_t] - E[X^-|\mathcal{F}_t] \\ &= \lim_{n\to\infty} E[(X^+ \wedge n)|\mathcal{F}_t] - \lim_{n\to\infty} E[(X^- \wedge n)|\mathcal{F}_t] \end{aligned} \tag{2.6}$$

where $X^+ = (X \wedge 0)$ and $X^- = (-X \wedge 0)$. Let us now recall that the generalized conditional expectation is not a linear operator.

Proposition 2.1.3. *For any $X, Y \in \bar{L}^0$ and $s, t \in \mathbb{T}$, $s > t$ we get*

1) $E[\lambda X|\mathcal{F}_t] \leq \lambda E[X|\mathcal{F}_t]$ for $\lambda \in L^0_t$ and $E[\lambda X|\mathcal{F}_t] = \lambda E[X|\mathcal{F}_t]$ for $\lambda \in L^0_t$, $\lambda \geq 0$;

2) $E[X|\mathcal{F}_t] \leq E[E[X|\mathcal{F}_s]|\mathcal{F}_t]$ and $E[X|\mathcal{F}_t] = E[E[X|\mathcal{F}_s]|\mathcal{F}_t]$ for $X \geq 0$;

3) $E[X|\mathcal{F}_t] + E[Y|\mathcal{F}_t] \leq E[X + Y|\mathcal{F}_t]$ and $E[X|\mathcal{F}_t] + E[Y|\mathcal{F}_t] = E[X + Y|\mathcal{F}_t]$ if $X, Y \geq 0$;

The proof of this proposition is deferred to the Appendix A.1 (page 115). See Remark 2.1.4 and Remark 2.4.2 (page 21) for additional comment about (2.6).

Remark 2.1.4. All inequalities in Proposition 2.1.3 can be strict. Let $t = 0$ and $k, s, t \in \mathbb{T}$, be such that $k > s > 0$. Let $\xi \in L^0_k$, $\xi = \pm 1$, each with probability $\frac{1}{2}$. Let $Z \in L^0_s$ be such that $Z \geq 0$, Z is independent of ξ and $E[Z] = \infty$. With $\lambda = -1$, $X = \xi Z$ and $Y = -X$ we get strict inequalities in 1), 2) and 3).

2.1.2 Essential infimum and supremum

For a non-negative family $\{X_i\}_{i\in I}$ (the index set I might be uncountable), where $X_i \in L^0$, we will denote by $\operatorname{ess\,inf}_{i\in I} X_i$ a unique (up to a set of measure zero) random variable from $\bar{L}^0$, such that

- For all $i \in I$, $X_i \geq \operatorname{ess\,inf}_{i\in I} X_i$;
- If $Y \in \bar{L}^0$, is such that $X_i \geq Y$ for any $i \in I$, then $\operatorname{ess\,inf}_{i\in I} X_i \geq Y$.

We will call $\operatorname{ess\,inf}_{i\in I} X_i$ the essential infimum (also called essential lower bound) for a family $\{X_i\}_{i\in I}$. Analogously, we can define the essential supremum (essential upper bound) denoted by $\operatorname{ess\,sup}_{i\in I} X_i$. Next, we define (generalized) essential lower bound for a family of random variables $\{X_i\}_{i\in I}$, such that $X_i \in \bar{L}^0$, by

$$\operatorname*{ess\,inf}_{i\in I} X_i = \lim_{n\to\infty} \Big[\operatorname{ess\,inf}_{i\in I}(X_i^+ \wedge n)\Big] - \lim_{n\to\infty} \Big[\operatorname{ess\,sup}_{i\in I}(X_i^- \wedge n)\Big].$$

In particular, it is worth mentioning, that if $X_i \in \bar{L}^0_t$ for $i \in I$, then $\operatorname{ess\,inf}_{i\in I} X_i \in \bar{L}^0_t$. Futhermore if for all $i, j \in I$, there exists $k \in I$, such that $X_k \leq X_i \wedge X_j$, then there exists a sequence $i_n \in I$ $(n \in \mathbb{N})$, such that $\{X_{i_n}\}_{n\in\mathbb{N}}$ is non-increasing and $\operatorname{ess\,inf}_{i\in I} X_i = \inf_{n\in\mathbb{N}} X_{i_n} = \lim_{n\to\infty} X_{i_n}$ (see [97, Appendix A]). Analogous results are true for $\operatorname{ess\,sup}_{i\in I} X_i$.

We will also make use of a conditional equivalents of $\operatorname{ess\,inf}$ and $\operatorname{ess\,sup}$ defined for a single random variable. For $X \in L^\infty$ and $t \in \mathbb{T}$, we will denote by $\operatorname{ess\,inf}_t X$ a unique (up to a set of measure zero), random variable from $\bar{L}^0_t$, such that

- $X \geq \operatorname{ess\,inf}_t X$;
- If $Y \in \bar{L}^0_t$, is such that $X \geq Y$, then $\operatorname{ess\,inf}_t X \geq Y$.

We will call it the $\mathcal{F}_t$-conditional essential infimum of X (see [10] for the proof of existence and uniqueness). Respectively, we will call $\operatorname{ess\,sup}_t(X) := -\operatorname{ess\,inf}_t(-X)$ the $\mathcal{F}_t$-conditional essential supremum. Next, For any $t \in \mathbb{T}$ and $X \in \bar{L}^0$ we define the $\mathcal{F}_t$-conditional essential infimum by[4]

$$\operatorname{ess\,inf}_t X := \lim_{n\to\infty}\Big[\operatorname{ess\,inf}_t(X^+ \wedge n)\Big] - \lim_{n\to\infty}\Big[\operatorname{ess\,sup}_t(X^- \wedge n)\Big], \tag{2.7}$$

As in the previous case, we put $\operatorname{ess\,sup}_t(X) := -\operatorname{ess\,inf}_t(-X)$. For convenience, we present some fundamental properties of conditional essential infimum and supremum, for $\bar{L}^0$ setup, that will be used throughout the paper.

Proposition 2.1.5. *For any $X, Y \in \bar{L}^0$, $s \geq t$ $(s, t \in \mathbb{T})$, $A \in \mathcal{F}_t$ and $U \in \bar{L}^0_t$ we get*

1) $\operatorname{ess\,inf}_{\omega\in A} X = \operatorname{ess\,inf}_{\omega\in A}(\operatorname{ess\,inf}_t X)$*;*

2) If $\operatorname{ess\,inf}_{\omega\in B} X = \operatorname{ess\,inf}_{\omega\in B} U$ *for any $B \in \mathcal{F}_t$, then* $U = \operatorname{ess\,inf}_t X$*;*[5]

3) $X \geq \operatorname{ess\,inf}_t X$*;*

4) If $Z \in \bar{L}^0_t$, is such that $X \geq Z$, then $\operatorname{ess\,inf}_t X \geq Z$*.*

5) If $X \geq Y$, then $\operatorname{ess\,inf}_t X \geq \operatorname{ess\,inf}_t Y$*;*

6) $\mathbb{1}_A \operatorname{ess\,inf}_t X = \mathbb{1}_A \operatorname{ess\,inf}_t(\mathbb{1}_A X)$*;*

7) $\operatorname{ess\,inf}_s X \geq \operatorname{ess\,inf}_t X$ *;*

The similar results are true for $\{\operatorname{ess\,sup}_t\}_{t\in\mathbb{T}}$*.*

The proof of Proposition 2.1.5 for $X, Y \in L^\infty$ can be found in [10]. Basing on the fact that for any $n \in \mathbb{N}$ and $X, Y \in \bar{L}^0$ we get $X^+ \wedge n \in L^\infty$, $X^- \wedge n \in L^\infty$ and $X^+ \wedge X^- = 0$, the proof for $X, Y \in \bar{L}^0$ is straightforward.

Please note that the concept of $\operatorname{ess\,inf}$ for a family of random variables, slightly differs from the concept of conditional $\operatorname{ess\,inf}$ for a single random variable. We hope the notation will be transparent, and it will be clear from the context, which definition we have in mind.

[4] One could also extend $\operatorname{ess\,inf}_t$ from L^∞ to $\bar{L}^0$ e.g. noting that the function arctan is strictly monotone and bounded and setting $\operatorname{ess\,inf}_t X = \arctan^{-1}[\operatorname{ess\,inf}_t(\arctan X)]$. Such extension will coincide with (2.7).

[5] $\operatorname{ess\,inf}_{\omega\in A} X := \sup\{k \in \mathbb{R} : \ \mathbb{1}_A k \leq \mathbb{1}_A X\}$.

2.2 Utility measures

Let $\mathcal{X}$ denote the space of random variables or adapted stochastic processes, described in Section 2.1. In this subsection we will now recall various notation and definitions used in the literature, especially in the theory of risk measures, performance measures and utility theory. The definitions and properties are usually linked to the corresponding (static) definitions and properties from the classical convex analysis (cf. [122] and references therein).

Definition 2.2.1 (General properties of maps)**.** Let $t_0 \in \mathbb{T}$. For any $X, Y \in \mathcal{X}$, $A \in \mathcal{F}_t$, $s \in \mathbb{T}$, such that $s > t$, and $m, \lambda, \beta \in \mathcal{X}_t$,[6] such that $0 \leq \lambda \leq 1$, $\beta > 0$ and $\|\beta\|_\infty < \infty$,[7] we will say that the map $f\colon \mathcal{X} \to \bar{L}^0$ is

(SBA) *Subadditive* if $f(X+Y) \leq f(X) + f(Y)$;

(SPA) *Superadditive* if $f(X+Y) \geq f(X) + f(Y)$;

(AD) *Additive* if $f(X+Y) = f(X) + f(Y)$;

(N) *Normalized* if $f(0) = 0$;

(M) *Monotone* if (MI) or (MD) hold, where

- (MI) (*Monotone increasing*) $X \leq Y \Rightarrow f(X) \leq f(Y)$;
- (MD) (*Monotone decreasing*) $X \leq Y \Rightarrow f(X) \geq f(Y)$;

(P) *Proper* if (P1) or (P2) hold, where

- (P1) $f(X) \in \widehat{L}^0$ and there exists $X_0 \in \mathcal{X}$, such that $f(X_0) > -\infty$;
- (P2) $-f(X) \in \widehat{L}^0$ and there exists $X_0 \in \mathcal{X}$, such that $f(X_0) < \infty$;

(F) *Finite* if $f(X) \in L^0$;

(tA) *$\mathcal{F}_t$-adapted* if $f(X) \in \bar{L}^0_t$;[8]

(tL) *$\mathcal{F}_t$-local* if $\mathbb{1}_A f(X) = \mathbb{1}_A f(\mathbb{1}_A \cdot_t X)$;

(tCA) *$\mathcal{F}_t$-cash additive* if $f(X + m1_{\{t\}}) = f(X) + m$;

(tCCA) *$\mathcal{F}_t$-counter cash additive* if $f(X + m1_{\{t\}}) = f(X) - m$;[9]

(tQCC) *$\mathcal{F}_t$-quasi-concave* if $f(\lambda \cdot_t X + (1-\lambda) \cdot_t Y) \geq f(X) \wedge f(Y)$;

(tCC) *$\mathcal{F}_t$-concave* if $f(\lambda \cdot_t X + (1-\lambda) \cdot_t Y) \geq \lambda f(X) + (1-\lambda) f(Y)$;

(tQCV) *$\mathcal{F}_t$-quasi-convex* if $f(\lambda \cdot_t X + (1-\lambda) \cdot_t Y) \leq f(X) \vee f(Y)$;

---

[6] See page 6 for the definition of $\mathcal{X}_t$.

[7] Note that we only consider (essentially) bounded elements, so that $\beta \cdot_t X \in \mathcal{X}$, for any $X \in \mathcal{X}$. See Remark 2.4.5 for details.

[8] Note that for the static case ($t = 0$) we will write simply $f(X) \in \bar{\mathbb{R}}$.

[9] Note that for $\mathcal{X} \subseteq \bar{L}^0$, we get $1_{\{t\}} = 1$. On the other hand, for $\mathcal{X} \subseteq \bar{\mathbb{V}}^0$, the value $m1_{\{t\}}$ corresponds to a single cash flow m, received at time t, see page 6 for details.

(tCV) $\mathcal{F}_t$*-convex* if $f(\lambda \cdot_t X + (1-\lambda) \cdot_t Y) \leq \lambda f(X) + (1-\lambda) f(Y)$;

(tSI) $\mathcal{F}_t$*-scale invariant* if $f(\beta \cdot_t X) = f(X)$;

(tPH) $\mathcal{F}_t$*-positively homogeneous* if $f(\beta \cdot_t X) = \beta f(X)$;

(tLSC) $\mathcal{F}_t$*-lower semi-continous wrt.* η, if $\{Z \in \bar{L}^0_t \mid f(X) \leq Z\}$ is η-closed[10];

(tUSC) $\mathcal{F}_t$*-upper semi-continous wrt.* η, if $\{Z \in \bar{L}^0_t \mid f(X) \geq Z\}$ is η-closed.

Moreover, if $\mathcal{X} \subseteq \bar{\mathbb{V}}^0$, with the same notation as before, we will say that f satisfies[11]

(tIP) $\mathcal{F}_t$*-independent of the past* if $f(X) = f(X - 0 \cdot_t X)$;

(tTI) $\mathcal{F}_t$*-translation invariant* if $f(X + m1_{\{t\}}) = f(X + m1_{\{s\}})$.

If $t = 0$, then we will abandon the notation "$\mathcal{F}_t$-" and call the map f simply adapted (A), translation invariant (TI), cash additive (CA), counter cash additive (CCA), quasi-concave (QCC), concave (CC), quasi-convex (QCV), convex (CV), scale invariant (SI), positively homogeneous (PH), lower semicontinous (LSC) and lower semicontinous (USC).

Remark 2.2.2. Most of the properties introduced in Definition 2.2.1 have a natural financial interpretation. For example *quasi-concavity* (tQCC), *concavity* (tCC) or *superadditivity* (SPA) might correspond to the positive effect of portfolio diversification. Please cf. [78, 45] and references therein, for more details and financial interpretation of other properties introduced above.

Remark 2.2.3. Please note that if $\mathcal{X} \subseteq \bar{L}^0$ then for any $t \in \mathbb{T}$, the properties *independence of the past* (tIP) and *translation invariance* (tTI) are automatically satisfied.

See also Remark 2.4.3 (page 22) for more comment about *locality* (tL), and Remark 2.4.4 (page 22) for alternative definitions of properties from Definition 2.2.1 using the idea of *Acceptance sets*. In Remark 2.4.5 (page 22) one could find additional comment about L^0-modules, which allow to get rid of various technical assumptions from Definition 2.2.1.

Proposition 2.2.4 (Selected implications for L^∞)**.** *Let* $\mathcal{X} = L^\infty$ *and let* $f \colon \mathcal{X} \to \bar{L}^0$. *Then the map* f *satisfies the following implications:*

1) Positively homogeneus (tPH) + *Superadditive* (SPA) $\Rightarrow$ *Convex* (tCV);

2) Convex (tCV) + *Positively homogeneus* (tPH) $\Rightarrow$ *Superadditive* (SPA);

3) Adapted (tA)+*Monotone* (MI)+*Quasi-convex* (tQCV)+*Cash additive* (tCA)$\Rightarrow$*Convex* (tCV);

4) Monotone (MI) + *Cash additive* (tCA) $\Rightarrow$ *Local* (tL)

5) Concave (CC) $\Rightarrow$ *Local* (tL)

The proof of this proposition is deferred to the Appendix A.1 (page 116).

[10] i.e. closed with respect to topology η; if η will be clear from the context, we will simply write that f is (tLSC). For example, if $\mathcal{X} = L^p$, then we will usually use the topology induced by $\|\cdot\|_p$ norm (see [77, Appendix A.7], for details).

[11] Note that for $\mathcal{X} \subseteq \bar{L}^0$, those properties are automatically satisfied; this is the very reason why we have decided to distinguish those properties from the previous ones.

Definition 2.2.5 (Additional properties for random variables)**.** Let $\mathcal{X} \subseteq \bar{L}^0$. For any $X, Y \in \mathcal{X}$, we will say that the map $f\colon \mathcal{X} \to \bar{L}^0$ is

(FP) Admitting *Fatou property*, if $f(X) \geq \limsup_{n\to\infty} f(X_n)$ for $\mathcal{X}$-dominated sequence [12] $\{X_n\}_{n\in\mathbb{N}}$ such that $X_n \in \mathcal{X}$ and $X_n \xrightarrow{a.s.} X$.

(LP) Admitting *Lebesgue property*, if $f(X) = \lim_{n\to\infty} f(X_n)$ for $\mathcal{X}$-dominated sequence $\{X_n\}_{n\in\mathbb{N}}$ such that $X_n \in \mathcal{X}$ and $X_n \xrightarrow{a.s.} X$.

(LI) *Law-invariant* if $f(X) = f(Y)$, whenever $Law(X) = Law(Y)$.

Moreover we shall write that f admits (FP') if $(-f)$ admits (FP).

See Remarks 2.4.6 and 2.4.7 (page 22) for additional comment about *Fatou property*. Also, for a general survey about properties introduced in this Section, see e.g. [107, 57, 77].

2.2.1 Families of utility measures

In this subsection we will present some specific families of (dynamic) utility measures, which are present in the literature. In mathematical finance, there are two most important families of maps:

- **Risk measures** were studied e.g. in [77, 7, 78] and provide theoretical framework for maps, which try to explain how risky an asset (investment strategy, stream of cash flows, etc.) can be. In practise the risk is associated with the amount of an asset or set of assets (currency) which must be kept in reserve, so that the position will be acceptable by the regulator [110]. The standard example (used widely by practitioners) is Value at Risk [93].
- **Performance measures** were studied in [17, 45, 66] and corresponds to maps designed to measure how good a financial position could be. The value of Performance measure might e.g. denote the degree of arbitrage consistency in the market [45, 65, 29], compare the financial position with a benchmark index [9, 66] or present the ratio between reward and risk [39]. In this case, industry standard is Sharpe Ratio [9].

Unfortunately both Value at Risk and Sharpe Ratio lack a lot of properties which are desired from a practical point of view (e.g. in general Sharpe's Ratio is not monotone and Value at Risk is not subadditive). This justifies the need for additional measures, which will be more suitable (see [7, 45] for more detailed comment). Please see Introduction and Section 4 for a more detailed comment about (dynamic) risk and performance measures, examples, properties, etc.

Definition 2.2.6 (Families of maps – static case)**.** We will say that the map $f : \mathcal{X} \to \bar{L}^0$ is

UM *Utility measure*, if f is adapted (A), translation invariant (TI) and monotone increasing (MI);

RM *(Monetary) risk measure*, if f is adapted (A), translation invariant (TI), monotone decreasing (MD), normalized (N) and counter cash-additive (CCA);

- *Convex risk measure*, if f is additionally convex (CV);

[12] This means that there exists $Y \in \mathcal{X}$ such that for all $n \in \mathbb{N}$ we have $|X_n| \leq |Y|$.

- *Coherent risk measure*, if f is additionally positively homogeneous (PH) and superadditive (SPA);

PM *Performance measure*, if f is adapted (A), translation invariant (TI), monotone increasing (MI) and scale invariant (SI);

- *Acceptability index*, if f is additionally quasi-concave (QCC);

Moreover, if $\mathcal{X} \subseteq \bar{L}^0$, then we will say that **f** is

CE *Certainty equivalent*, if there exists $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$, u strictly increasing and continuous on $\bar{\mathbb{R}}$[13], such that for any $X \in \mathcal{X}$ and $t \in \mathbb{T}$:

$$f(X) = u^{-1}(E[u(X)]); \tag{2.8}$$

Additionally, we usually assume that maps from Definition 2.2.6 are also proper (P). See Remark 2.4.8 (page 23) for additional comment about this fact.

Remark 2.2.7. Sometimes, instead of dealing with a (monetary, convex or coherent) risk measure f, it is more convenient to consider the mapping $-f$ (cf. [44] and references therein). Note that such maps belong to the family of UMs (they are called *Monetary utility functions* or *Monetary utility measures* e.g. in [40, 36, 94, 79, 44]). Similar remark will apply for corresponding conditional and dynamic families of maps.

Definition 2.2.8 (Families of maps - conditional case)**.** We will say that the map $f : \mathcal{X} \to \bar{L}^0$ is

tUM $\mathcal{F}_t$-*cond. utility measure* if f is adapted (tA), translation invariant (tTI), independent of the past (tIP) and monotone increasing (MI);

tRM $\mathcal{F}_t$-*cond. (monetary) risk measure*, if f is adapted (tA), translation invariant (tTI), independent of the past (tIP), normalized (N), monotone decreasing (MD) and counter cash-additive (tCCA);

- $\mathcal{F}_t$-*cond. convex risk measure*, if f is additionally convex (tCV);
- $\mathcal{F}_t$-*cond. coherent risk measure*, if f is additionally positively homogeneous(tPH) and super additive (SPA);

tPM $\mathcal{F}_t$-*cond. performance measure*, if f is adapted (tA), translation invariant (tTI), independent of the past (tIP), monotone increasing (MI) and scale invariant (tSI);

- $\mathcal{F}_t$-*cond. acceptability index*, if f is additionally quasi-concave (tQCC);

Moreover, if $\mathcal{X} \subseteq \bar{L}^0$, then we will say that **f** is

tCE $\mathcal{F}_t$-*cond. certainty equivalent*, if there exists $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$, u strictly increasing and continuous on $\bar{\mathbb{R}}$, such that for any $X \in \mathcal{X}$ and $t \in \mathbb{T}$:

$$f_t(X) = u^{-1}(E[u(X)|\mathcal{F}_t]); \tag{2.9}$$

[13] i.e. strictly increasing and continuous of $\mathbb{R}$, with $u(\pm\infty) = \lim_{n\to\pm\infty} u(n)$.

Remark 2.2.9. Maps defined in Definition 2.2.6 are a special case of corresponding maps defined in Definition 2.2.8, with $t = 0$.

See Remark 2.4.9 (page 23) for the alternative definition of conditional map, when the static map is provided and it is law invariant (LI).

Having in mind Remark 2.4.3 (page 22), which explains why $\mathcal{F}_t$-locality plays crucial role in dynamic framework, we are now ready to present the definition of a dynamic map.

Definition 2.2.10 (Families of maps - dynamic case)**.** We will say that a family $\mathbf{f} = \{f_t\}_{t\in\mathbb{T}}$ of maps $f_t : \mathcal{X} \to \bar{L}^0$, is

1) *Adapted* (dA), *Local* (dL), etc., if for any $t \in \mathbb{T}$, f_t is adapted (tA), local (tL), etc. We will also use (dSBA), (dSPA), (dAD), (dN), (dM), (dMI), (dMD), (dP), (dP1), (dP2), (dF), (dIP), (dTI), (dCA), (dCCA), (dQCC), (dCC), (dQCV), (dCV), (dSI), (dPH), (dLI), (dFP), (dFP'), (dLP) to denote the corresponding property of $\mathbf{f}$;

2) *Dynamic map* if $\mathbf{f}$ is adapted (dA) and local (dL);

3) *Dynamic utility measure* (dUM), *Dynamic risk measure* (dRM), etc. if for any $t \in \mathbb{T}$, the map f_t is tUM, tRM, etc. We will also use symbols dPM and dCE.

Usually the filtration $\{\mathcal{F}_t\}_{t\in\mathbb{T}}$ corresponds to the evolution of time, so a dynamic utility describes the change of our preferences through time. See Remark 2.4.10 (page 23) for the spatial approach.

The main objects of study in this thesis will be $\mathcal{F}_t$-conditional utility measures and dynamic utility measures. Thus, from now on we will always use

- φ to denote utility measure (UM) or dynamic utility measure (dUM) (depending on the context).

- φ_t, to denote conditional utility measure (tUM).

We hope it will be clear from the context, which $\mathcal{X}$ we have in mind. Similarly, we will use f to denote a (general) map $f : \mathcal{X} \to \bar{L}^0$ and $\mathbf{f} = \{f_t\}_{t\in\mathbb{T}}$ to denote a (general) dynamic map.

2.2.2 How to extend an utility measure

In this subsection we will assume that $L^\infty \subseteq \mathcal{X} \subset L^0$. Let $\widetilde{\mathcal{X}}$ be such that $\mathcal{X} \subset \widetilde{\mathcal{X}}$ (for transparency, one might assume that $\mathcal{X} = L^\infty$ and $\widetilde{\mathcal{X}} = L^0$). For a given map $f : \mathcal{X} \to \bar{L}^0$, we want to create a map $f' : \widetilde{\mathcal{X}} \to \bar{L}^0$, which will inherit some or all of the properties of map f. In this subsection we will present some approaches to this problem. First of all let us say what we understand by the *extension* of tUM.

Definition 2.2.11 (Extension of utility measure)**.** Let $\varphi_t : \mathcal{X} \to \bar{L}^0$ be tUM and let $\widetilde{\mathcal{X}} \supset \mathcal{X}$. We will call a map $\widetilde{\varphi}_t \colon \widetilde{\mathcal{X}} \to \bar{L}^0$, *an* $\widetilde{\mathcal{X}}$*-extension of* φ_t, if $\widetilde{\varphi}_t$ is tUM and $\widetilde{\varphi}_t|_{\mathcal{X}} \equiv \varphi_t$. Similarly, given $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ dUM, we will say that $\widetilde{\varphi} = \{\widetilde{\varphi}_t\}_{t\in\mathbb{T}}$ is an $\widetilde{\mathcal{X}}$*-extension of* φ, if for any $t \in \mathbb{T}$, $\widetilde{\varphi}_t$ is $\widetilde{\mathcal{X}}$-extension of φ_t.

See Remark 2.4.11 (page 23) for a comment about extending a map using so called *robust representation* and Remark 2.4.12 (page 23) about the extensions which preserve additional properties.

It could be easily shown that for any tUM or dUM and $\mathcal{X} \subset \widetilde{\mathcal{X}}$ there exist tUM or dUM, which extends the map from $\mathcal{X}$ to $\widetilde{\mathcal{X}}$. The exemplary detailed procedure of construction is shown in Appendix A.4.1. One could also define the biggest and smallest maps, which extend tUM. To do this we will make use of the following sets

$$\mathcal{Y}_A^+(X) := \{Y \in \mathcal{X} \mid \ \mathbb{1}_A Y \geq \mathbb{1}_A X\},$$

$$\mathcal{Y}_A^-(X) := \{Y \in \mathcal{X} \mid \ \mathbb{1}_A Y \leq \mathbb{1}_A X\},$$

defined for $X \in \bar{L}^0$ and $A \in \mathcal{F}$.

Definition 2.2.12 (Upper and lower extensions of utility measure)**.** Let φ_t be tUM. We will denote by $\varphi_t^+ : \bar{L}^0 \to \bar{L}_t^0$, an *upper* $\bar{L}^0$*-extension of* φ_t, where[14]

$$\varphi_t^+(X) := \operatorname*{ess\,inf}_{A \in \mathcal{F}_t} \Big[\mathbb{1}_A \operatorname*{ess\,inf}_{Y \in \mathcal{Y}_A^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c}(\infty) \Big]. \tag{2.10}$$

Respectively, we will denote by $\varphi_t^- : \bar{L}^0 \to \bar{L}_t^0$, an *lower* $\bar{L}^0$*-extension of* φ_t, where

$$\varphi_t^-(X) := \operatorname*{ess\,sup}_{A \in \mathcal{F}_t} \Big[\mathbb{1}_A \operatorname*{ess\,sup}_{Y \in \mathcal{Y}_A^-(X)} \varphi_t(Y) + \mathbb{1}_{A^c}(-\infty) \Big]. \tag{2.11}$$

The next result shows that $\varphi^{\pm}$ are two 'extreme' extensions, and any other extension is sandwiched between them.

Proposition 2.2.13. *Let* φ_t *be tUM. Then* φ_t^- *and* φ_t^+ *are* $\bar{L}^0$*-extensions of* φ_t*. Moreover, let* $\widetilde{\varphi}_t$ *be an* $\bar{L}^0$*-extension of* φ_t*. Then, for any* $X \in \bar{L}^0$ *we get*

$$\varphi_t^-(X) \leq \widetilde{\varphi}_t(X) \leq \varphi_t^+(X). \tag{2.12}$$

The proof of this proposition is deferred to the Appendix A.1 (page 116).

Clearly, generally speaking the maps (2.10) and (2.11) are not equal, and thus the extensions of a dUM are not unique.

Corollary 2.2.14. *Let* $t \in \mathbb{T}$ *and* $\mathcal{B} \subseteq \bar{L}^0$ *be such that, for any* $A \in \mathcal{F}_t$, $\mathbb{1}_A\mathcal{B} \subseteq \mathcal{B}$ *and* $\mathbb{1}_A\mathcal{B} + \mathbb{1}_{A^c}\mathcal{B} \subseteq \mathcal{B}$*. For any* $\mathcal{F}_t$*-local and monotone*[15] *mapping* $f : \mathcal{B} \to \bar{L}_t^0$*, the maps* $f^{\pm}$ *defined analogously as in (2.10) and (2.11) will be extensions of* f *to* $\bar{L}^0$*, preserving locality and monotonicity.*

We omit the detailed proof here, as it is a simple generalization of Proposition 2.2.13. Next, let us present two ways of constructing maps on $\widehat{\mathcal{X}}$ and $\bar{\mathcal{X}}$ from maps on $\mathcal{X}$ (see page 6 for the definition of $\widehat{\mathcal{X}}$ and $\bar{\mathcal{X}}$).

1) For $f\colon \mathcal{X} \to \bar{L}^0$, we define a mapping $\widehat{f}\colon \widehat{\mathcal{X}} \to \bar{L}^0$ as

$$\widehat{f}(X) := \liminf_{n \to -\infty} f\Big(X \vee n\Big), \quad n \in \mathbb{Z}. \tag{2.13}$$

[14]We will use convention $\operatorname{ess\,sup} \emptyset = -\infty$ and $\operatorname{ess\,inf} \emptyset = \infty$.

[15]That is, $\mathcal{F}_t$-local and monotone on $\mathcal{B}$.

2) For $f\colon \mathcal{X} \to \bar{L}^0$, we also define a mapping $\bar{f}\colon \bar{\mathcal{X}} \to \bar{L}^0$ as

$$\bar{f}(X) := \liminf_{m\to\infty} \widehat{f}\Big(X \wedge m\Big), \quad m \in \mathbb{Z}. \tag{2.14}$$

Clearly, for monotone f, one can replace lim inf with lim in (2.13) and (2.14). Next propositions shows that the functions $\widehat{f}$ and $\bar{f}$ could inherit some of the properties of f, although generally speaking, $\widehat{f}$ and $\bar{f}$ are not $\widehat{\mathcal{X}}$-extensions and $\bar{\mathcal{X}}$-extenstions of f unless they satisfy some additional properties, like (FP) or (LP). (See Remark 2.2.16 for counterexample.)

Proposition 2.2.15. *Let $\varphi_t\colon \mathcal{X} \to \bar{L}^0_t$ be tUM. Then*

1) $\widehat{\varphi}_t$ and $\bar{\varphi}_t$ are tUM.

2) If φ_t is cash-additive (tCA) and $\varphi_t(0) \neq \infty$, then $\widehat{\varphi}_t$ and $\bar{\varphi}_t$ are cash-additive (tCA);

3) $\varphi_t(X) = \widehat{\varphi}_t(X) = \bar{\varphi}_t(X)$ for $X \in L^\infty$. Moreover, if φ_t satisfies Fatou property (FP), then $\varphi_t(X) = \widehat{\varphi}_t(X)$ for $X \in \mathcal{X}$ and if φ_t satisfies Lebesgue property (LP), then $\varphi_t(X) = \bar{\varphi}_t(X)$ for $X \in \mathcal{X}$.

The proof of this proposition is deferred to the Appendix A.1 (page 118).

Remark 2.2.16. In general $\widehat{f}$ and $\bar{f}$ might not be an extensions of f. On $\mathcal{X} = L^1$, It is sufficient to consider the example

$$f(X) = \begin{cases} 1, & \text{if } E[X] > 0,\\ 0, & \text{if } E[X] = 0.\\ -1, & \text{if } E[X] < 0.\end{cases}$$

It is easy to show, that this function is UM. For $X \sim N(0,1)$, we get $\widehat{f}(X) = 1$, $f(X) = 0$ and $\bar{f}(X) = -1$.

Remark 2.2.17. Please note that extensions (2.13) and (2.14) are a natural way of extending the map from L^∞ into any bigger space, especially if we want to preserve continuity, i.e. properties (FP) or (LP). This is the main reason, why usually we only deal with those type of extensions [56, 69] (see also Remark 2.4.12). Moreover, (2.13) and (2.14) are a natural extension of convention (2.4), especially if the map admits robust representation (2.16), which will be introduced in Subsection 2.2.3.

2.2.3 Robust representation

The duality is among the most important properties for UMs and dUMs. Many methods used to solve optimal stochastic control problems base on such representation, which allow us to move the original problem to the dual space and to obtain dual representations (cf. [79] and references therein). For example if $\mathcal{X}$ is a locally convex topological vector space (e.g. for $\mathcal{X} = L^p$, when $p \geq 1$), then Separation theorem (Hahn-Banach theorem) hold, which allow us to use many classical results from convex analysis, such as Fenchel-Moreau theorem (see Appendix A.3). The similar result is also true in quasi-convex framework (again see Appendix A.3). In this thesis we will mostly use the robust representation for convex RMs defined for random variables, the reason we only introduce Definition 2.2.18.

Following [75] (see also [78, Section 2.5], for the explanation of this notation) we will use notation $\mathcal{M}_1 := \mathcal{M}_1(\Omega, \mathcal{F})$ to denote the set of all probability measures on $(\Omega, \mathcal{F})$, $\mathcal{M}_1(\mathbb{P}) := \mathcal{M}_1(\Omega, \mathcal{F}, \mathbb{P})$ to denote the set of all probability measures on $(\Omega, \mathcal{F})$ absolutely continuous wrt. $\mathbb{P}$. For $q \in [1, \infty]$ and $t \in \mathbb{T}$, we will also write (see e.g. [95, 3, 75])

$$\mathcal{Q}_t^q := \begin{cases} \{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \in L^q,\ \mathbb{Q} = \mathbb{P} \text{ on } \mathcal{F}_t\} & \text{if } q \neq \infty, \\ \{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) \mid \mathbb{Q} \in ba(\mathcal{F}),\ \mathbb{Q} = \mathbb{P} \text{ on } \mathcal{F}_t\} & \text{if } q = \infty. \end{cases} \tag{2.15}$$

The number q will always denotes the conjugate index of p (i.e. $\frac{1}{p} + \frac{1}{q} = 1$) and $ba(\mathcal{F})$ will denote the set of all finitely additive signed measures on $\mathcal{F}$. For the static case we willl write $\mathcal{Q}^p := \mathcal{Q}_0^p$.

Definition 2.2.18 (Representable risk measures)**.** Let φ be a monetary RM (defined for random variables). We will call φ representable, if

$$\varphi(X) = -\inf_{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P})} [E_{\mathbb{Q}}[X] + \alpha^{\min}(\mathbb{Q})], \tag{2.16}$$

for some function $\alpha^{\min} \colon \mathcal{M}_1(\mathbb{P}) \to \mathbb{R} \cup \{\infty\}$. Similarly, we will say that $\{\varphi_t\}_{t \in \mathbb{T}}$, a monetary dRM is representable, if for all $t \in \mathbb{T}$,

$$\varphi_t(X) = -\operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P})} [E_{\mathbb{Q}}[X|\mathcal{F}_t] + \alpha_t^{min}(Q)], \tag{2.17}$$

for some function $\alpha_t^{\min} \colon \mathcal{M}_1(\mathbb{P}) \to \widehat{L}_t^0$.

Remark 2.2.19. Representation (2.16) has a financial interpretation. Any element of $\mathcal{M}_1(\mathbb{P})$ could be treated as a (risk) scenario (i.e. plausible probabilistic model) and the function $\alpha^{\min}$ is so called penalty function, which tell us how seriously we treat any particular scenario. Then φ could be seen as a penalised worst case expectation, taken over elements of $\mathcal{M}_1(\mathbb{P})$. This approach is also in line with the concept of stress testing (cf. [75] and references therein).

See Remark 2.4.13 (page 23) for more comment about function $\alpha^{\min}$, Remark 2.4.14 (page 23) for more general approach to robust representation using *minimal risk functions* and Remark 2.4.15 (page 24) about connection to dual spaces. Remark 2.4.16 (page 24) could provide some information about robust representations in conditional case and robust representation for maps defined for stochastic processes.

We will conclude this subsection with Theorems which link convex RMs and dRMs defined on L^p ($p \in [1, \infty]$) with robust representation. We refer to [95] and [3] for more details and proofs of Theorem 2.2.20 and 2.2.21, respectively.

Theorem 2.2.20. *Let φ be a monetary RM defined on L^p for $p \in [1, \infty]$. Moreover, let φ be proper* (P2)*. Then, the following are equivalent:*

1. *The mapping φ is convex* (CV) *and satisfies Fatou property* (FP')*.*

2. *The mapping φ admits robust representation*[16]

$$\varphi(X) = -\inf_{\mathbb{Q} \in \mathcal{Q}^q} [E_{\mathbb{Q}}[X] + \alpha^{min}(\mathbb{Q})], \tag{2.18}$$

where $\alpha^{min} : \mathcal{M}_1(\mathbb{P}) \to \mathbb{R}_+ \cup \{\infty\}$ is such that $\alpha^{min}(\mathbb{Q}) = \infty$, if $\mathbb{Q} \notin \mathcal{Q}^q$.

[16] Recall that q denote the conjugate index of p ($\frac{1}{p} + \frac{1}{q} = 1$)

Moreover, if φ is convex (CV), law-invariant (LI) and finite (F), then it satisfies Fatou property (FP').

The proof of Theorem 2.2.20 could be found e.g. in [95, Theorem 2.4 and 3.1]. See also [75, Section 3] for more comments about representation (2.18). See also Remark 2.4.17 (page 24) for more comment about Proposition 2.2.20. The similar result is also true in the dynamic setting.

Theorem 2.2.21. *Let φ be a monetary dRM defined on L^p, for $p \in [1,\infty]$. Moreover, let $-\varphi_t(X) \in \widehat{L}^1_t$ for any $X \in \mathcal{X}$ and $t \in \mathbb{T}$.*[17] *Then, the following are equivalent:*

1. *The mapping φ is convex (dCV) and satisfies Fatou property (dFP').*
2. *The mapping φ admits robust representation, i.e. for any $t \in \mathbb{T}$,*

$$\varphi_t(X) = \inf_{\mathbb{Q}\in\mathcal{Q}^q_t} [E_{\mathbb{Q}}[X|\mathcal{F}_t] + \alpha^{min}_t(\mathbb{Q})], \tag{2.19}$$

where $\alpha^{min}_t : \mathcal{M}_1(\mathbb{P}) \to (\bar{L}^0_t \vee 0)$[18] *is such that $\alpha^{min}(\mathbb{Q}) = \infty$, if $\mathbb{Q} \notin \mathcal{Q}^q_t$.*

Moreover, if φ is convex (dCV), law-invariant (dLI) and finite (dF), then it satisfies Fatou property (dFP').

2.3 Stochastic control with utility measures

In a very general (static) framework we will be interested in solving problems of the form

$$\sup_{X\in\mathcal{Z}} \varphi(X), \tag{2.20}$$

where φ is UM and $\mathcal{Z} \subseteq \mathcal{X}$. Of course we must impose some additional restrictions on φ and $\mathcal{Z}$, if want to solve (2.20) explicitly (e.g. to guarantee existence and uniqueness of the optimal solution). While we will generally assume that φ is quasi-concave (QCC) and upper semi-continuous (USC) (at least on the set $\mathcal{Z}$), the conditions imposed on the set $\mathcal{Z}$ will depend on the problem, which we will consider (cf. [117] and references therein, where sufficient conditions are presented in a more general context).

For example, in portfolio optimisation theory, we might require that for any $X \in \mathcal{Z}$, we get $X = F(Z)$, for some $Z \in \mathcal{H}$, where $F : \mathcal{H} \to \mathcal{X}$ is concave and $\mathcal{H}$ is a convex subset of some vector space. This would lead to problems of the form

$$\sup_{Z\in\mathcal{H}} \varphi(F(Z)), \tag{2.21}$$

which were studied e.g. in [109, 131] for negatives of (quasi)convex RMs.

The set $\mathcal{Z}$ could also be constructed using UMs, typically (convex) RMs. For example it could be given by $\mathcal{Z} = \{X \in \widehat{\mathcal{Z}} : \rho(X) \leq x\}$, where $\widehat{\mathcal{Z}}$ is the set of all admissible portfolios (e.g. self-financing, or corresponding to the choice of some appropriate weights in portfolio construction), ρ is RM, and $x \in \mathbb{R}$.

[17] Note that this condition might be considered as an extension of (P2) to the conditional case.
[18] $(\bar{L}^0_t \vee 0)$ is the set of nonnegative $\mathcal{F}_t$-measurable random variables with values in $\bar{\mathbb{R}}$.

Moreover, the optimal stopping problems could be presented in this form. For example let $\varphi : L^1 \to \mathbb{R}$, let C_0^T denote the set of all stopping times with values in $\mathbb{T} = \{0, 1, \ldots, T\}$ and let $V = (V_t)_{t=1}^T \in \mathbb{V}^1$. Then the class of problems

$$\sup_{\nu \in C_0^T} \varphi(V_\nu) \tag{2.22}$$

leads to optimal stopping problems studied e.g. in [77, Section 6.5] or [36, Section 5.1 and 5.2] for coherent dRMs. In particular if φ is a standard expectation then this leads to classical formulation of optimal stopping problem (cf. [118] and references therein).

Remark 2.3.1. The family of maps φ, for which the problem (2.20) was studied extensively, correspond to the class of *Certainty Equivalents* defined in (2.8). If $\mathcal{X} \subseteq \bar{L}^0$ and we have some given (possibly stochastic) utility function u (see e.g. [15, Section 2.1]), then we seek for φ, such that for any $X \in \mathcal{X}$, we get

$$u(\varphi(X)) = E[u(X)]. \tag{2.23}$$

The value $\varphi(X)$ corresponds to certainty equivalent of X. For transparency, in this thesis we will use this name only for $u : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ being a classic von Neumann-Morgenstern utility function. The corresponding family of maps was defined in (2.8) and could be obtain using formula

$$\varphi(X) = u^{-1}(E[u(X)]). \tag{2.24}$$

The families of the form (2.24) were extensively studied in the literature, especially in the actuarial risk theory, where they were reffered to as *Mean value principles* (see e.g. [85, Chapter 5, Section 4]). Note, that for such φ, the problem could be translated to so called *expected utility optimisation problem*, as maximisation of (2.24) is equivalent to maximisation of the function $E[u(\cdot)]$. Please also note that with $u(x) = x$, we get the standard expectation operator.

2.3.1 Bellman principle of optimality

In (2.20), the shape of $\mathcal{Z}$ is typically strongly connected with the filtration $\{\mathcal{F}_t\}_{t\in\mathbb{T}}$, usually through a (finite) family of adapted stochastic process $\{(X_t^i)_{t\in\mathbb{T}}\}_{i\in I}$, where $I = \{1, 2, \ldots, N\}$ for some $N \in \mathbb{N}$. They might describe e.g. the evolution of all market factors, securities, etc.

Moreover, any $X \in \mathcal{Z}$ could be seen as a function of $\{(X_t^i)_{t\in\mathbb{T}}\}_{i\in I}$ and a (predictable) control process $u = (u_t)_{t\in\mathbb{T}}$, which basically tell us what we decide to do at any time $t \in \mathbb{T}$ (see [12] for a more formal introduction and definition of u). In other words, at any time $t \in \mathbb{T}$ we could have an impact on X, changing the strategy (e.g. rebalancing a portfolio), where the choice will depend on the evolution of the whole market till time t as well as our previous control decisions (e.g. on $\{(X_k^i)\}_{k\leq t}$ and $(u_k)_{k\leq t}$, as they will determine our current wealth, shape of portfolio, etc.). Let X^u correspond to the choice of strategy u.

Basically, Bellman principle of optimality (also called dynamic programming principle) tell us that if $u^* = (u_t^*)_{t\in\mathbb{T}}$ is optimal control for the problem (2.20), then for any time $t \in \mathbb{T}$, the control $(u_k^*)_{k\geq t}$ should be optimal for the conditional version of the problem (2.20), considered at time t. We will refer to the sequence of conditional versions of the problem (2.20) for all $t \in \mathbb{T}$ as the dynamic programming equations for (2.20).

In a general framework, for a finite time horizon $\mathbb{T} = \{0, 1, \ldots, T\}$, where $T \in \mathbb{N}$, such sequence could be given recursively by

$$\begin{aligned} U_T(u_0, u_1, \ldots, u_T) &:= \Phi_T(X^u) \\ U_t(u_0, u_1, \ldots, u_t) &:= \operatorname*{ess\,sup}_{u_{t+1}} \Phi_t^{u_{t+1}}[U_{t+1}(u_1, \ldots, u_{t+1})], \qquad t \in \{T-1, \ldots, 0\}, \end{aligned} \tag{2.25}$$

where $\Phi_t^{u_{t+1}}$ corresponds to so called response function and reflects the impact of the choice of u_{t+1} at time t for a (preference) value function, $\Phi_T : \mathcal{X} \to \bar{L}_T^0$ shifts the final value of X^u into the space of preferences and $X^u \in \mathcal{Z}$ correspond to the element obtained using control $u = (u_1, \ldots, u_T)$. The maps $\Phi_t^{u_{t+1}}$ very often depend on some dUM, being a dynamized version of φ. Moreover we assume that $U_t \in \bar{L}_t^0$ for any $t \in \mathbb{T}$, and thus, we could define $U \in \bar{\mathbb{V}}^0$, as $U = (U_0, U_1, \ldots, U_T)$.

To make the problem (2.25) traceable, and reduce the number of parameters needed for dynamic programming equations we usually assume that Z is a Markov chain and use (time-consistent) dynamic version of φ (see e.g. [129]). A various specific approaches to this problem will be presented in Section 5. See also (2.27) for the exemplary application of dynamic programming principle in the context of optimal stopping.

2.3.2 Optimal stopping

For $t \in \mathbb{T}$ ($t \leq T$), let $\mathcal{C}_t^T$ denote the set of all stopping times with values in the set $\{t, t+1, \ldots, T\}$. In this subsections we will present some insight on the class of problems given by

$$\sup_{\nu \in \mathcal{C}_0^T} \mathrm{E} V_\nu, \tag{2.26}$$

for a fixed $V \in \mathbb{V}^1$. In other words we will consider problems of the form (2.22), for $\varphi(\cdot) = E[\cdot]$. Applying the dynamic programming principle to (2.26), we define $U \in \mathbb{V}^1$ recursively, by

$$\begin{aligned} U_T &= V_T, \\ U_t &= \max\left(V_t, \mathrm{E}[U_{t+1}|\mathcal{F}_t]\right), \qquad t \in \{T-1, \ldots, 0\}. \end{aligned} \tag{2.27}$$

Since V is assumed to be integrable, U is also integrable due to L^1-continuity of the conditional expectation operator. We will call U the *Snell envelope of* V [104]. The following theorem collects the standard properties of Snell envelopes:

Theorem 2.3.2. *Let $V \in \mathbb{V}^1$ and let U denote it's Snell envelope. Then:*

1. *U is the smallest supermartingale dominating V.*
2. *$U_t = \operatorname{ess\,sup}\{\mathrm{E}[V_\tau|\mathcal{F}_t] \ : \ \tau \in \mathcal{C}_t^T\}$.*
3. *Let $\tau_t = \min(s \geq t \mid U_s = V_s)$. Then $\tau_t \in \mathcal{C}_t^T$ and*
$$\begin{aligned} \tau_T &= T, \\ \tau_t &= t\,\mathbf{1}_{\{V_t \geq \mathrm{E}[U_{t+1}|\mathcal{F}_t]\}} + \tau_{t+1}\mathbf{1}_{\{V_t < \mathrm{E}[U_{t+1}|\mathcal{F}_t]\}}, \qquad t \in \{0, \ldots, T-1\}. \end{aligned}$$
4. *$U_t = \mathrm{E}[V_{\tau_t} \mid \mathcal{F}_t]$.*

5. $\mathrm{E}[U_{t+1}|\mathcal{F}_t] = \mathrm{E}[V_{\tau_{t+1}}|\mathcal{F}_t]$.

6. τ_0 *is optimal for* V*. In particular, for any optimal stopping time* σ*,*

$$U_0 = \mathrm{E}[V_{\tau_0}] = \mathrm{E}[V_\sigma].$$

7. τ_t *can also be defined recursively:*

$$\begin{aligned} \tau_T &= T, \\ \tau_t &= t\,\mathbf{1}_{\{V_t \geq \mathrm{E}[V_{\tau_{t+1}}|\mathcal{F}_t]\}} + \tau_{t+1}\mathbf{1}_{\{V_t < \mathrm{E}[V_{\tau_{t+1}}|\mathcal{F}_t]\}}, \qquad t \in \{0,\ldots,T-1\}. \end{aligned}$$

8. $\tau \in \mathcal{C}_0^T$ *is optimal for* V *if and only of*

   (a) $V_\tau = U_\tau$*,*

   (b) $U_{t\wedge\tau}$ *is a martingale (where* $t \wedge \tau = \min(t,\tau)$*).*

9. $U_0 = \max(Z_0, \mathrm{E}[V_{\tau_1}])$.

10. *The random variable* τ_0 *is the smallest optimal stopping time.*

11. *Let* $U_t = M_t - A_t$ *be the Doob decomposition of the Snell envelope into a martingale* M_t *and a non-decreasing predictable process* A_t *starting at* 0*. If* $K = \{t : \; A_{t+1} > 0\} \subset \{0,1,\ldots,T-1\}$*, then*

$$\varrho = \begin{cases} T & \text{if } K = \emptyset, \\ \min K & \text{if } K \neq \emptyset, \end{cases}$$

is the largest optimal stopping time.

The proof of Theorem 2.3.2 can be found e.g. in [104].

2.4 Additional remarks

Remark 2.4.1. It is worth noticing, that for stochastic processes, the space $\mathcal{X}$, embedded with the operator $(\cdot_t)$ does not define a proper L^0–module (see e.g. [67, 68]). It is enough to notice that the property $0 \cdot_t V = 0$ (for any $V \in \mathcal{X}$) does not hold. This is one of the basic properties of L^0–modules (their structure is similar to the structure of vector spaces). Nevertheless, the ideas presented in this thesis will be strictly connected with this theory. Moreover, most dynamic utility measures (see Subsection 2.2.1) considered in this thesis will admit so called *independence of the past* for which the operator $(\cdot_t)$ becomes an appropriate operator for an L^0–module, i.e. one might consider the corresponding operator, with $V_0, \ldots, V_{t-1}$ substituted with 0s in (2.5), for which $\mathcal{X}$ becomes an L^0–module (see [17] for details).

Remark 2.4.2. As stated in Remark 2.1.4, the operator $E[\cdot|\mathcal{F}_t]$ might be not linear in the general case. Nevertheless, such notation is strictly connected to the theory of (dynamic) risk measures. See [113, 114, 115] for the general concepts of *Nonlinear Expectations* (in particular *g-Expectations*) and [32] for the relation to *Choquet Integrals* (also called *Choquet Expectations*).

Remark 2.4.3. The properties similar or equivalent to *locality* (tL) were studied in many forms (see e.g. [57, Subsection 2.2]). Sometimes they were referred to as the *regularity* property. Moreover, some authors treat locality as a part of dynamic time-consistency property (see [103, Definition 1.1]). Moreover, nearly all of the properties considered in Definitions 2.2.1 are local in a sense that $\mathcal{F}_t$-locality allow us to treat them locally, even if we assume they hold just for $t = 0$. For example, let us assume that the map $f\colon \mathcal{X} \to \bar{L}^0$ is *scale invariant* (SI) and *local* (tL). Then, we get $f(\beta \cdot_t X) = f(X)$, for $\beta = \sum_{i=1}^{\infty} \mathbb{1}_{A_i}\beta_i$, where $A_i \in \mathcal{F}_t$, $\bigcup_{i=1}^{\infty} A_i = \Omega$, $A_i \cap A_j = \emptyset$ $(i \neq j)$, $\beta_i \in \mathbb{R}$, $\beta_i > 0$. Moreover, if we assume that f possess some form of continuity (e.g. Lebesgue property for $\mathcal{X} = L^1$, as in Definition 2.2.5, see also [103]), then we get $f(\beta \cdot_t X) = f(X)$ for $\beta \in \mathcal{X}_t$, $\beta > 0$ and consequently *conditional scale invariance* (tSI) holds.

Remark 2.4.4. In many cases for a fixed $t \in \mathbb{T}$ and $f : \mathcal{X} \to \bar{L}^0$, most of the properties introduced in Definition 2.2.1 could be alternatively defined using the corresponding properties for sets, using the idea of so called *($\mathcal{F}_t$-conditional) acceptance sets*, denoted by $A_t(m) = \{X \in \mathcal{X} \mid f(X) \geq m\}$ for $m \in \bar{L}^0_t$. See e.g. [77, Section 4.1] and [45] for details.

Remark 2.4.5. For $\mathcal{X} = L^p$ $(p \in [1, \infty])$, properties in Definition 2.2.1 are suited for the L^p-framework, which implies some drawbacks. Let us alone point out that we need additional restriction for scalar in the definition of (SI) (see [67] for detailed discussion and more drawbacks). To overcome this and many more problems, the theory of L^0-modules was developed. Intuitively speaking, instead of dealing with the space L^p_t, we can define the space

$$L^p_{\mathcal{F}_t} := L^0_t \cdot L^p = \{YX \mid Y \in L^0_t,\ X \in L^p\},$$

where the elements of the space L^0_t act as scalars (defining module over a ring [67]). Such space is a natural interpretation of the fact, that at time t, we can treat any $\mathcal{F}_t$-measurable random variable as a constant, so no additional restriction should be imposed on such random variables. Moreover, when we embed $L^p_{\mathcal{F}_t}$ with a certain kind of topology, it allow us to extend many theorems from classic (convex) functional analysis. This approach could also be applied, when we consider the space of adapted stochastic processes. See [67, 68], when this theory was initiated, for more formal and detailed description. Also, for a good review of existing literature, more detailed explanation of this idea, and it's application for L^p-type of maps, see [83, 88]. Nevertheless, it is worth mentioning that the theory of L^0-modules is not yet polished (see e.g. [145, 88] for potential problems).

Remark 2.4.6. The Fatou property (FP) has been studied in many various forms in existing literature (cf. [55, 13] and references therein). In risk measure framework, it is usually assumed that $(-f)$ satisfies Fatou property (FP), the reason we introduced notation (FP'). For maps defined on order complete Riesz spaces[19], which are monotone (MI), Fatou property (FP) coincide with so called *order upper semicontinuity* [13]. Moreover, for maps defined on L^p, where $p \in [0, \infty]$, upper semi-continuity (USC) (wrt. $\|\cdot\|_p$) implies Fatou property (FP) [13]. The Fatou property (FP) is also strictly connected to continuity from above, as shown e.g. in [13, 95]. Thus, Fatou property (FP) is crucial if we want the robust representation to hold (cf. [95]). The Lebesgue property (LP) has been studied e.g. in [112, 94].

Remark 2.4.7. On L^∞, we know that adaptivity (A), properness (P1), concavity (CC), monotonicity (MI), cash-additivity (CA) and law invariance (LI) imply the Fatou property (FP) (see [94] for

[19] Riesz space is a partially ordered vector space where the order structure is a lattice (i.e. for any $X, Y \in \mathcal{X}$, there exists Z, such that $X \vee Y \leq Z$

details). In particular, any Monetary risk measure (see Definition 2.2.6) on L^{∞}, which is law invariant(LI), satisfies Fatou property (FP'). Moreover, on L^p, for $p \in [1, \infty)$, any convex risk measure (see Definition 2.2.6), which is finite (F), satisfies Fatou property (FP') (cf. [95]).

Remark 2.4.8. We will usually assume that maps from Definition 2.2.6 also satisfy properness (P) (see [56] for discussion). While proper maps are the real object of study, it would be inconvenient to exclude maps without this property, as they arise in many natural situations (cf. [122]). Similar remark will apply for corresponding conditional and dynamic families of maps. Note that the properness (P) is also needed if we want to use certain classical tools from convex analysis like Fenchel-Moreau Theorem (cf. [95] and references therein). It is also worth mentioning, that in general we could not require the map to be finite as it would exclude many interesting cases. For example, if the space is atomless, then there is no real-valued coherent risk measure on L^0 (see [55, Theorem 5.1]).

Remark 2.4.9. As $\mathbb{R}$ equipped with Borel σ-algebra $\mathcal{B}$ is a Borel space, for any $t \in \mathbb{T}$ there always exist a unique probability kernel $\mathbb{P}^{(X,\mathcal{F}_t)} : \Omega \times \mathcal{B} \to [0,1]$ from $(\Omega, \mathcal{F}_t)$ to $(\mathbb{R}, \mathcal{B})$ such that for any $X \in \mathcal{X}$ we get

$$\mathbb{P}^{(X,\mathcal{F}_t)}(\cdot, B) = P[X \in B|\mathcal{F}_t](\cdot),$$

for $B \in \mathcal{B}$ [96]. In particular, for UMs which are law invariant (LI), we could try to define $\mathcal{F}_t$-conditional utility measure directly from the static one, i.e. for given f, we put

$$f_t(X)(\omega) = f(\mathbb{P}^{(X,\mathcal{F}_t)}(\omega, \cdot)).$$

Remark 2.4.10. In a general framework a family of σ-subfields of $\mathcal{F}$ could be associated with a different concept. For example, in the spatial framework it could describe a node in the financial system (as in systemic risk framework). Then a conditional utility measure will correspond to the value of local preferences, e.g. for a certain financial institution (see [74] for details).

Remark 2.4.11. Some tUM or dUM which are additionally quasi-concave (QCC) and satisfy Fatou property (FP) can be extended through so called robust representation (see Subsection 2.2.3 for details). Exemplary procedure for coherent dUMs could be found in [41]. See also [69].

Remark 2.4.12. It is known that every RM defined on L^{∞} which is additionally proper (P), law invariant (LI), convex (CV) and which negative admits Fatou property (FP) can be extended to RM defined on L^1 with the same properties. Moreover, such extension is unique [69], so L^1 could be treated as canonical space for this type of maps. To see when there exists extension of RM, which additionally preserve finiteness (F) and continuity, see [100].

Remark 2.4.13. Please note that the function $\alpha^{\min} : \mathcal{M}_1(\mathbb{P}) \to \mathbb{R} \cup \{\infty\}$ might be used to define a monetary RM. Given a representable Monetary RM one might recover $\alpha^{\min}$ noticing that[20]

$$\alpha^{min}(\mathbb{Q}) = -\inf_{X \in A_\rho} E_{\mathbb{Q}}[X],$$

for $A_\rho = \{X \in L^p \mid \varphi(X) \geq 0\}$. See [75, 78] for details.

Remark 2.4.14. The robust representation for a general case of UMs defined for random variables (i.e. when $\mathcal{X} = L^p$ for $p \geq 1$) is of the form

$$\varphi(X) = \inf_{\mathbb{Q} \in \mathcal{M}_1} R(\mathbb{Q}, E_{\mathbb{Q}}[X]), \tag{2.28}$$

[20]Note that $\inf_{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P})} \alpha^{\min}(Q) = 0$, as φ satisfies (N).

where R is so called *Minimal risk function* (see [63, 17, 82] for details). Robust representation hold usually when φ admits quasi-concavity (QCC) and upper semi-continuity (USC). This representation could be also extended to the case, when $\mathcal{X}$ is a locally convex topological L^0–module (see e.g. [17]). See also [37, 107, 63, 17] for information about robust representation in a more general framework (e.g. for stochastic processes).

Remark 2.4.15. Let $\mathcal{X} = L^p$, $p \in [1, \infty)$. Usually, for representable dRM defined on $\mathcal{X}$, instead of taking infimum on $\mathcal{M}_1(\mathbb{P})$ in (2.16) it is enough to consider it on the subset $Q^q \subseteq \mathcal{M}_1(\mathbb{P})$. The space Q^q corresponds to the dual space of $\mathcal{X}$ (considering the standard topology induced by $\|\cdot\|_p$ norm). See Theorems 2.2.20 and 2.2.21 for details.

Remark 2.4.16. There are several papers which characterise robust representations in the conditional case, both for random variables as well as adapted stochastic processes [3, 67, 63, 82, 17, 68]. We usually consider only probability measures which coincide with $\mathbb{P}$ on $\mathcal{F}_t$, see e.g. Theorem 2.2.21. Moreover, equation (2.17) seems to be a natural way to construct tRM from a (static) RM which admits (2.16), if only we get formula for conditional equivalent of $\alpha^{\min}$ (cf. [41, 69] and references therein). One could also consider conditional robust representation (2.28) in the general framework, as in Remark 2.4.14 (see e.g. [17, 68, 107]).

Remark 2.4.17. In Theorem 2.2.20 we could replace the Fatou property (FP') with lower semi-continuity (LSC) (wrt. $\|\cdot\|_p$), as those properties coincide in the static framework (see [95, Theorem 3.3]). Moreover, for $p = \infty$, the mapping φ in Theorem 2.2.20, point 2) is always finite, so any law invariant and convex RM is representable (cf. Remark 2.4.7).

Chapter 3

Time-consistency of dynamic utility measure

As we have pointed our in the introduction, a line of research that branched out from [8] was dedicated to extension of the theory of risk and performance measure to the dynamical, multi-period setup, where the flow of information is modeled by a filtration, say $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$, that is a component of the underlying probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. When the space $\mathcal{X}$ is considered, the risk measures are defined on the set of $\mathcal{F}$-measurable random variables that correspond to (terminal) cashflows, or, more generally, on the set of adapted stochastic processes that correspond to dividend streams or to cumulative cashflows. As shown in Section 2, most of the axioms from static case are transferred to the dynamic setup in a natural way, with addition of requirement that the measures are $\mathbb{F}$-adapted (tA) and, frequently, that they are independent of the past (tIP). From another point of view, an extension from one period to multi-period models can be realized through robust representation theorems, essentially by replacing expectations with conditional expectations (see Subsection 2.2.3). As defined in the previous Section, risk measures obtained by this procedure are referred to as conditional and dynamic risk measures (also cf. [124, 57, 24]).

As shown in one of the first papers that studied dynamic coherent risk measures, [121], if one is concerned about making noncontradictory decisions (from the risk/utility point of view) over the time, then an additional axiom, called time consistency, is needed. Over the past decade significant progress has been made towards expanding the theory of dynamic risk measures and their time consistency. For example, so called cocycle condition (for convex risk measures) was studied in [25, 73], recursive construction was exploited in [40], relation to acceptance and rejection sets was studied in [53], the concept of prudence was introduced in [116], connections to g-expectations were studied in [125], and the relation to Bellman's principle of optimalty was shown in [7]. For more details on we also refer the reader to [35, 38, 36, 50, 57, 92, 121, 81, 80, 124, 141, 54, 140, 82, 17], as well as to a comprehensive survey paper [2] and the references therein.

Let us briefly recall the concept of strong time consistency of dynamic monetary risk measures (dRMs), which is one of the most recognized forms of time consistency (see Appendix A.2 for other types). Assume that $\rho_t(X)$ is the value of a dynamic monetary risk measure at time $t \in \mathbb{T}$, that corresponds to the riskiness, at time t, of the cashflow X, with X being an $\mathcal{F}$-measurable random variable. The monetary risk measure is said to be strongly time consistent if for any $t < s$ ($t, s \in \mathbb{T}$), and any $\mathcal{F}$-measurable random variables $X, Y \in \mathcal{X}$ we have that

$$\rho_s(X) = \rho_s(Y) \quad \Rightarrow \quad \rho_t(X) = \rho_t(Y). \tag{3.1}$$

The financial interpretation of strong time consistency is clear – if X is as risky as Y at some future time s, then today, at time t, X is also as risky as Y. One of the main features of the strong time consistency is its connection to dynamic programming principle. It is not hard to show that in the L^∞ framework, a monetary risk measure is strongly time consistent if and only if

$$\rho_t = \rho_t(-\rho_s), \quad 0 \leq t < s. \tag{3.2}$$

All other forms of time consistency for monetary risk measures, such as weak and middle acceptance time consistentcy or rejection time consistentcy, are tied to this connection as well. In [140], the author proposed a general approach to time consistency for cash-additive risk measures by introducing so called 'test sets' or 'benchmark sets.' Each form of time consistency was associated to a benchmark set of random variables, and larger benchmark sets correspond to stronger forms of time consistency. For reader convenience, this concept is presented with more details in Appendix A.2.

Let us now present some insight on time consistency for dynamic performance measures (dPMs). The first study of time consistency of scale invariance measures is presented in [20], where the authors elevated the theory of coherent acceptability indices to dynamic setup in discrete time. It was pointed out that none of the forms of time consistency for risk measures is suitable for scale invariant maps. Recursive property similar to (3.2) or benchmark sets approach essentially can not be applied to scale invariant maps. Consequently, one of the main challenge was to find an appropriate form of time consistency of acceptability indices, that would be both financially reasonable and mathematically tractable. For the case of random variables, the proposed form of time consistency for a dynamic coherent acceptability index α reads as follows: for any $\mathcal{F}_t$-measurable random variables $m_t,\ n_t$, and any $t < T$, the following implications hold

$$\begin{aligned} \alpha_{t+1}(X) \geq m_t \quad &\Rightarrow \quad \alpha_t(X) \geq m_t, \\ \alpha_{t+1}(X) \leq n_t \quad &\Rightarrow \quad \alpha_t(X) \leq n_t. \end{aligned} \tag{3.3}$$

The financial interpretation is also clear – if tomorrow X is acceptable at least at level m_t, then today X is also acceptable at least at level m_t; similar interpretation holds true for the second part (3.3). It is fair to say, we think, that dynamic acceptability indices and their time consistency properties play a critical role in so called conic approach to valuation and hedging of financial contracts [16, 23, 126].

We recall that both risk measures and performance measures, in the nutshell, put preferences on the set of cashflows. While the corresponding forms of time consistency (3.1) and (3.3) for these classes of maps, as argued above, are different, we note that generally speaking both forms of time consistency are linking preferences between different times. The aim of this Section is to present a unified and flexible framework for time consistency of risk and performance measures, that integrates existing forms of time consistency as well as various connections between them. We consider a (large) class of maps that are postulated to satisfy only two properties - monotonicity and locality - which we call dynamic utility measures (dUMs), and we study time consistency of such maps. These two properties, in our opinion, have to be satisfied by any reasonable dynamic risk or performance measure. We introduce the notion of an update rule that is meant to link preferences between different times. The time consistency is defined in terms of an update rule. We provide various update rules that allow to recover several known forms of time consistency, and also such that allow to study some new forms of time consistency. When appropriate, for each form of time consistency we consider separately the case of (terminal) cashflows, referred in this paper as the case of random variables, and the case of dividend streams, referred to as the case of stochastic processes. For each type of time consistency we provide different equivalent formulations along with a discussion regarding financial interpretation and suitability of each rule. We also provide a comprehensive analysis of the connections between considered forms of time consistency. The numerous examples of specific dUMs and types of time consistency, that they follow will be presented in Chapter 4.

This Chapter is organized as follows. In Section 3.1 we set forth the main concepts of the paper – the notion of an update rule and the definition of time consistency of a dUM. We prove a general result about time consistency, that can be viewed as counterpart of dynamic programming principle (3.2), and that is used conveniently in the sequel.

Section 3.2 is devoted to various types of time consistency for random variables (i.e. when $\mathcal{X} \subseteq \bar{L}^0$). Each type of time consistency is discussed in a separate subsection. We start with the weakest form of time consistency – the weak time consistency, and we conclude with the notion of super/submartingale time consistency. We present some fundamental properties for each type of time consistency, and we establish some relationships between them.

Then, in Section 3.3 we briefly present the corresponding results for random processes, based on Section 3.2. Numerous examples both for random variables and stochastic processes will be presented later, in Section 4.

Section 3.4 will be devoted to a recursive construction, which allow to construct a strongly time consistent dUM form any given map (on finite time horizon),

This Chapter will be based on [18].

3.1 Definition of time consistency

In this section we introduce the time consistency of dynamic risk and performance measures, or more generally, the time consistency of dUMs introduced in the previous section.

We recall that these dUMs are defined on $\mathcal{X}$, where $\mathcal{X}$ either denotes the space of random variables (e.g. L^p, for $p \in \{0, 1, \infty\}$) or the space of stochastic processes (e.g. $\mathbb{V}^p$ for $p \in \{0, 1, \infty\}$), so, our study of time consistency is done relative to such spaces. While, for clarity, in this section we will only use spaces L^p and $\mathbb{V}^p$, the definition of time consistency can be easily adapted for other type of spaces, such as Orlicz hearts (as studied in [37]) or topological L^0-modules (see for instance [17]). Usually, the need to consider spaces smaller than L^0 or $\mathbb{V}^0$ is motivated by the aim to obtain robust representation of such measures, as explained in Section 2.2.3. For this, a certain topological structure is required (cf. Remark 2.4.11). On the other hand, 'time consistency' refers only to consistency of measurements in time, where no particular topological structure is needed, and thus most of the results obtained here hold true for $p = 0$.

Assume that $\{\varphi_t\}_{t\in\mathbb{T}}$ is a dUM on $\mathcal{X}$. For an arbitrary fixed $X \in \mathcal{X}$ and $t \in \mathbb{T}$ the value $\varphi_t(X)$ represents a quantification (measurement) of preferences about X at time t. Clearly, it is reasonable to require that any such quantification (measurement) methodology should be coherent as time passes. This is precisely the motivation behind the concepts of time consistency of dUMs.

There are various forms of time consistency proposed in the literature, some of them suitable for one class of measures, other for a different class of measures, without a unified approach to fit them all. For example, for dynamic convex (or coherent) risk measures time consistency is equivalent to dynamic programming principle (also called Bellman principle), which we have introduced in Section 2.3.1, or variations of it [2]. On the other hand, as shown in [20], dynamic programming principle essentially is not suited for scale invariant measures such as dynamic acceptability indices, and the authors introduce a new type of time consistency tailored for these measures and provide a robust representation of them. Nevertheless, in all these cases the time consistency property connects, in a coherent way, the measurements at different times.

Next, we will introduce the notion of update rule that serves as the main tool in relating the measurements of preferences at different times, and also, it is the main building block of our unified theory of time consistency property.

Definition 3.1.1 (Update rule)**.** We will call a family $\mu = \{\mu_{t,s}\}_{s>t}$, $s,t \in \mathbb{T}$, of maps

$$\mu_{t,s} : \bar{L}^0_s \times \mathcal{X} \to \bar{L}^0_t$$

an *update rule* if for any $s > t$, the map $\mu_{t,s}$ satisfies the following conditions:

1) (Locality) $\mathbb{1}_A \mu_{t,s}(m, X) = \mathbb{1}_A \mu_{t,s}(\mathbb{1}_A m, X)$;

2) (Monotonicity) if $m \geq m'$, then $\mu_{t,s}(m, X) \geq \mu_{t,s}(m', X)$;

for any $X \in \mathcal{X}$, $A \in \mathcal{F}_t$ and $m, m' \in \bar{L}^0_s$.

Remark 3.1.2. As we have mentioned, the update rule is responsible for updating preferences through time. This concept is tightly linked with maps, which are projections. The natural choice of an update rule is the conditional expectation operator, i.e. we can consider the update rule $\{\mu_{t,s}\}_{s>t}$, given by[1]

$$\mu_{t,s}(m, X) = E[m|\mathcal{F}_t]. \tag{3.4}$$

Note that this particular update rule does not depend on s and X. Update rule might be also responsible for discounting the preferences. Intuitively speaking, the risk of loss in the far future might be more preferred than the imminent risk of loss (see [44] for the more detailed explanation of this idea). For example, the update rule $\{\mu_{t,s}\}_{s>t}$ of the form

$$\mu_{t,s}(m, X) = \begin{cases} \alpha^{s-t} E[m|\mathcal{F}_t] & \text{on} \{E[m|\mathcal{F}_t] \geq 0\}, \\ \alpha^{t-s} E[m|\mathcal{F}_t] & \text{on } \{E[m|\mathcal{F}_t] < 0\}. \end{cases} \tag{3.5}$$

for a fixed $\alpha \in (0, 1)$ correspond to this concept. Note that 'discounting' proposed here has nothing to do with the ordinary discounting, as we act on discounted values already.

We are now ready to introduce the general definition of time consistency.

Definition 3.1.3.[2] Let μ be an update rule. We will say that the dUM $\{\varphi_t\}_{t\in\mathbb{T}}$ is *μ-acceptance time consistent* if

$$\varphi_s(X) \geq m_s \quad \Longrightarrow \quad \varphi_t(X) \geq \mu_{t,s}(m_s, X), \tag{3.6}$$

[1] We consider here the generalized conditional expectation defined in (2.6).

[2] We introduce the concept of time consistency only for dUMs, for transparency. However, the definition itself is suitable for any map acting from $\mathcal{X}$ to $\bar{L}^0$. For example, traditionally in the literature, the time consistency is defined for dynamic risk measures (which form a subclass of negatives of dUMs), and the above definition of time consistency will be appropriate, although one has to flip 'acceptance' with 'rejection'.

for all $s,t\in\mathbb{T}$, $s>t$, $X\in\mathcal{X}$ and $m_s\in\bar{L}^0_s$. Respectively, we will say that $\{\varphi_t\}_{t\in\mathbb{T}}$ is *μ-rejection time consistent* if

$$\varphi_s(X)\leq m_s \quad\Longrightarrow\quad \varphi_t(X)\leq \mu_{t,s}(m_s,X), \tag{3.7}$$

for all $s,t\in\mathbb{T}$, $s>t$, $X\in\mathcal{X}$ and $m_s\in\bar{L}^0_s$. If properties (3.6) and (3.7) are satisfied only for $s,t\in\mathbb{T}$, such that $s=t+1$, then we will say that $\{\varphi_t\}_{t\in\mathbb{T}}$ is *one step μ-acceptance time consistent* and *one step μ-rejection time consistent*, respectively.

Since dUMs are local (dL) and monotone (dMI), properties with clear financial interpretations, the update rules are naturally assumed to be local and monotone too.

We see that the first argument $m\in\bar{L}^0_s$ in $\mu_{t,s}$ serves as a benchmark to which the measurement $\varphi_s(X)$ is compared. The presence of the second argument, $X\in\mathcal{X}$, in $\mu_{t,s}$, allows the update rule to depend on the objects (the Xs), which the preferences are applied to. However, as we will see in next section, there are natural situations when the update rules are independent of $X\in\mathcal{X}$, and sometimes they do not even depend on the future times $s\in\mathbb{T}$.

Remark 3.1.4. With the update rule $\{\mu_{t,s}\}_{s>t}$ defined in (3.4), the concept of acceptance and rejection time consistency coincide with supermartingale and submartingale property, respectively. In other word for a given dUM, say $\{\varphi_t\}_{t\in\mathbb{T}}$, we ask if the property

$$\varphi_t(X)\geq E[\varphi_s(X)|\mathcal{F}_t] \quad (\text{resp.}\ \ \leq) \tag{3.8}$$

is satisfied for any $X\in\mathcal{X}$.[3]

Next, we define several particular classes of update rules, suited for our needs.

Definition 3.1.5 (Various types of update rules)**.** Let μ be an update rule. We will say that μ is:

1) *X-invariant*, if $\mu_{t,s}(m,X)=\mu_{t,s}(m,0)$;

2) *sX-invariant*, if there exists a family $\{\mu_t\}_{t\in\mathbb{T}}$ of maps $\mu_t:\bar{L}^0\to\bar{L}^0_t$, such that $\mu_{t,s}(m,X)=\mu_t(m)$;

3) *Projective*, if it is sX-invariant and $\mu_t(m_t)=m_t$;

for any $s,t\in\mathbb{T}$, $s>t$, $X\in\mathcal{X}$, $m\in\bar{L}^0_s$ and $m_t\in\bar{L}^0_t$.

Remark 3.1.6. If an update rule $\mu=\{\mu_{t,s}\}_{s>t}$ is sX-invariant, then it is enough to consider only the corresponding family $\{\mu_t\}_{t\in\mathbb{T}}$. Hence, with slight abuse of notation we shall write $\mu=\{\mu_t\}_{t\in\mathbb{T}}$ and call it an update rule as well.

Remark 3.1.7. Examples of update rules satisfying 1) and 3) are given by (3.5) and (3.4), respectively. The update rule, which satisfy 2), but not 3) can be constructed by substituting α^{t-s} with a constant in (3.5). Generally speaking update rules for stochastic processes will not satisfy 1) as the information about the process in the time interval (t,s) will affect $\mu_{t,s}$, see Section 3.3 for details.

The financial interpretation of acceptance time consistency is straightforward: if $X\in\mathcal{X}$ is accepted at some future time $s\in\mathbb{T}$, at least at level m, then today, at time $t\in\mathbb{T}$, it is accepted at least at level $\mu_{t,s}(m,X)$. Similarly for rejection time consistency. Essentially, the update rule μ translates the preference levels at time s to preference levels at time t. As it turns out, this simple and intuitive definition of time consistency, with appropriately chosen μ, will cover various

[3] See Proposition 3.1.8 for the proof of equivalence between (3.6) and (3.8) for this particular update rule.

cases of time consistency for risk and performance measures that can be found in the existing literature. Moreover, it will allow us to establish some fundamental properties of the dUMs and some important connections between different versions of time consistency.

Next, we will give an equivalent formulation of time consistency. While the proof of the equivalence is simple, the result itself will be conveniently used in the sequel. Moreover, it can be viewed as a counterpart of dynamic programming principle, which is an equivalent formulation of dynamic consistency for convex risk measures.

Proposition 3.1.8. *Let μ be an update rule and let $\{\varphi_t\}_t$ be a dUM. Then,*

1) $\{\varphi_t\}_{t\in\mathbb{T}}$ is μ-acceptance time consistent if and only if

$$\varphi_t(X) \geq \mu_{t,s}(\varphi_s(X), X), \tag{3.9}$$

for any $X \in \mathcal{X}$ and $s, t \in \mathbb{T}$, such that $s > t$.

2) $\{\varphi_t\}_{t\in\mathbb{T}}$ is μ-rejection time consistent if and only if

$$\varphi_t(X) \leq \mu_{t,s}(\varphi_s(X), X), \tag{3.10}$$

for any $X \in \mathcal{X}$ and $s, t \in \mathbb{T}$, such that $s > t$.

Proof. Let μ be an update rule.

1) The implication ($\Rightarrow$) follows immediately, by taking in the definition of acceptance time consistency $m_s = \varphi_s(X)$.

($\Leftarrow$) Assume that $\varphi_t(X) \geq \mu_{t,s}(\varphi_s(X), X)$, for any $s, t \in \mathbb{T}, s > t$, and $X \in \mathcal{X}$. Let $m_s \in \bar{L}^0_s$ be such that $\varphi_s(X) \geq m_s$. Using monotonicity of μ, we get $\varphi_t(X) \geq \mu_{t,s}(\varphi_s(X), X) \geq \mu_{t,s}(m_s, X)$.

2) The proof is similar to 1). $\square$

The financial interpretation of (3.9) is similar to that of (3.6): if in the future, at time s, we accept the cash-flow X at level $\varphi_s(X)$, then today, at time t, we should accept the same cash-flow at least at level $\mu_{t,s}(\varphi_t(X), X)$ – the update of the acceptance level of X from time s to time t. Analogous interpretation applies to rejection time consistency.

Remark 3.1.9. It is clear, and also naturally desired, that a monotone transformation of a dUM will not change the preference order of the underlying elements. We want to emphasize that a monotone transformation will also preserve the time consistency. In other words, the preference orders will be also preserved in time. Indeed, if $\{\varphi_t\}_{t\in\mathbb{T}}$ is μ-acceptance time consistent, and $g : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ is a strictly monotone function, then the family $\{g \circ \varphi_t\}_{t\in\mathbb{T}}$ is $\widetilde{\mu}$-acceptance time consistent, where the update rule $\widetilde{\mu}$ is defined by $\widetilde{\mu}_{t,s}(m, X) = g(\mu_{t,s}(g^{-1}(m), X))$, for $t, s \in \mathbb{T}$, $s > t$, $X \in \mathcal{X}$ and $m \in \bar{L}^0_s$.

Before moving to the concrete definitions of time consistency, we will give some general remarks about relationship between time consistency for random variables and time consistency for random processes.

In what follows, for the case of random variables, $\mathcal{X} = L^p$, we will only consider update rules that are X-invariant. Hence, as it will be clear later, the case of random variables can be viewed as a particular case of stochastic processes by considering cash-flows with only the terminal payoff, i.e. stochastic processes such that $V = 1_{\{T\}} V_T$ (for finite time horizon). Nevertheless, we treat this case separately for transparency.

In the present work, in the case of stochastic processes, we will focus on one step update rules, such that
$$\mu_{t,t+1}(m,V)=\mu_{t,t+1}(m,0)+f(V_t), \tag{3.11}$$
where $f:\bar{\mathbb{R}}\to\bar{\mathbb{R}}$ is a Borel measurable function, such that $f(0)=0$. We do this primarily to allow for a direct link between our results and the existing literature. We note, that any such one step update rule μ can be easily adapted to the case of random variables. Indeed, upon setting $\widetilde{\mu}_{t,t+1}(m):=\mu_{t,t+1}(m,0)$ we get a one step X-invariant update rule $\widetilde{\mu}$, which is suitable for random variables. Moreover, $\widetilde{\mu}$ will define the corresponding type of one step time consistency for random variables. Of course, this correspondence between update rule for processes and random variables is valid only for 'one step' setup.

Finally, we note that for update rules, which admit the so called nested composition property (cf. [131, 129] and references therein),
$$\mu_{t,s}(m,V)=\mu_{t,t+1}(\mu_{t+1,t+2}(\dots\mu_{s-2,s-1}(\mu_{s-1,s}(m,V),V)\dots V),V), \tag{3.12}$$
we have that μ-acceptance (resp. μ-rejection) time consistency is equivalent to one step μ-acceptance (resp. μ-rejection) time consistency.

This is another reason why we consider only one step update rules for stochastic processes, however one can consider more exotic forms of time consistency, within proposed framework, and derive numerous properties and relationships between them, a task that we will leave for further studies.

3.2 Selected types of time consistency for random variables

In this section we will analyze various types of time consistency, including some of those that have been studied in the literature, using the framework developed earlier in this paper. If $\mathcal{X}=L^p$, for $p\in\{0,1,\infty\}$, then the elements $X\in\mathcal{X}$ are interpreted as discounted terminal cash-flows.

3.2.1 Weak time consistency

The notion of weak time consistency was introduced in [140], and subsequently studied in [2, 7, 36, 57, 1, 38]. The idea is that if 'tomorrow', say at time s, we accept $X\in\mathcal{X}$ at level $m_s\in\mathcal{F}_s$, then 'today', say at time t, we would accept X at least at any level smaller or equal than m_s, adjusted by the information $\mathcal{F}_t$ available at time t (cf. (3.25)). Similarly, if tomorrow we reject X at level smaller than $m_s\in\mathcal{F}_s$, then today, we should also reject X at any level bigger than m_s, adapted to the flow of information $\mathcal{F}_t$. This suggests that the update rules should be taken as $\mathcal{F}_t$-conditional essential infimum and supremum, respectively. First, we will show that $\mathcal{F}_t$-conditional essential infimum and supremum are projective update rules.

Proposition 3.2.1. *The family $\mu^{\inf}:=\{\mu^{\inf}_t\}_{t\in\mathbb{T}}$ of maps $\mu^{\inf}_t:\bar{L}^0\to\bar{L}^0_t$ given by*
$$\mu^{\inf}_t(m):=\operatorname{ess\,inf}_t m,$$
is a projective[4] update rule. Moreover,
$$\mu^{\inf}_t(m)=\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}^1_t}E_{\mathbb{Q}}[m|\mathcal{F}_t], \tag{3.13}$$

[4]See Remark 3.1.6 for the comment about notation.

where $\mathcal{Q}_t^1$ *is defined in* (2.15). *Similar result is true for family* $\mu^{\sup} := \{\mu_t^{\sup}\}_{t\in\mathbb{T}}$*, defined by* $\mu_t^{\sup}(m) = \operatorname{ess\,sup}_t m$.

Proof. Monotonicity and locality of $\mu^{\inf}$ is a straightforward implication of Proposition 2.1.5. Thus, $\mu^{\inf}$ is sX-invariant update rule. The projectivity comes straight from the definition. Now, let a family $\mu = \{\mu_t\}_{t\in\mathbb{T}}$ of maps $\mu_t : \bar{L}^0 \to \bar{L}_t^0$ be given by

$$\mu_t(m) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}_t^1} E_{\mathbb{Q}}[m|\mathcal{F}_t]. \tag{3.14}$$

As any measure $\mathbb{Q} \ll \mathbb{P}$ could be associated with random variable $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$, we could write

$$\mu_t(m) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}_t^1} \frac{E[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} m|\mathcal{F}_t]}{E[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}|\mathcal{F}_t]} = \operatorname*{ess\,inf}_{Z\in P_t} E[Zm|\mathcal{F}_t], \tag{3.15}$$

where $P_t = \{Z \in L^1 \mid Z \geq 0,\ E[Z|\mathcal{F}_t] = 1\}$. Before proving (3.13), we will need to prove some facts about μ.

First, let us show that μ is sX-invariant update rule. Let $t \in \mathbb{T}$. Monotonicity is straightforward. Indeed, let $m, m' \in \bar{L}^0$ be such that $m \geq m'$. For any $Z \in P_t$, using the fact that $Z \geq 0$, we get $Zm \geq Zm'$. Thus, $E[Zm|\mathcal{F}_t] \geq E[Zm'|\mathcal{F}_t]$ and consequently

$$\operatorname*{ess\,inf}_{Z\in P_t} E[Zm|\mathcal{F}_t] \geq \operatorname*{ess\,inf}_{Z\in P_t} E[Zm'|\mathcal{F}_t].$$

Locality follows from the fact, for any $A \in \mathcal{F}_t$ and $m \in \bar{L}^0$, using Proposition 2.1.3 and convention $0 \cdot \pm\infty = 0$, we get

$$\begin{aligned}
\mathbb{1}_A \mu_t(m) &= \mathbb{1}_A \operatorname*{ess\,inf}_{Z\in P_t} E[Zm|\mathcal{F}_t] \\
&= \mathbb{1}_A \operatorname*{ess\,inf}_{Z\in P_t} (E[(\mathbb{1}_A Z)m|\mathcal{F}_t] + E[(\mathbb{1}_{A^c} Z)m|\mathcal{F}_t]) \\
&= \mathbb{1}_A \operatorname*{ess\,inf}_{Z\in P_t} E[(\mathbb{1}_A Z)m|\mathcal{F}_t] + \mathbb{1}_A \operatorname*{ess\,inf}_{Z\in P_t} E[(\mathbb{1}_{A^c} Z)m|\mathcal{F}_t] \\
&= \mathbb{1}_A \operatorname*{ess\,inf}_{Z\in P_t} E[Z(\mathbb{1}_A m)|\mathcal{F}_t] + \mathbb{1}_A \operatorname*{ess\,inf}_{Z\in P_t} \mathbb{1}_{A^c} E[Zm|\mathcal{F}_t] \\
&= \mathbb{1}_A \mu_t(\mathbb{1}_A m).
\end{aligned}$$

Note, that the third equality follows from the fact that $(\mathbb{1}_A Z)(\mathbb{1}_{A^c} Z') = 0$ for any $Z, Z' \in P_t$. Thus, μ is sX-invariant update rule.

Secondly, let us prove that we get

$$m \geq \mu_t(m), \tag{3.16}$$

for any $m \in \bar{L}^0$. Let $m \in L^0$. For $\alpha \in (0,1)$ let[5]

$$Z_\alpha := \mathbb{1}_{\{m \leq q_t^+(\alpha)\}} E[\mathbb{1}_{\{m \leq q_t^+(\alpha)\}}|\mathcal{F}_t]^{-1}. \tag{3.17}$$

where $q_t^+(\alpha)$ is $\mathcal{F}_t$-conditional (upper) α quantile of m, defined as

$$q_t^+(\alpha) := \operatorname{ess\,sup}\{Y \in L_t^0 \mid E[\mathbb{1}_{\{m\leq Y\}}|\mathcal{F}_t] \leq \alpha\}. \tag{3.18}$$

[5] In the risk measure framework, it might be seen as the risk minimazing scenario for conditional TV@R$_\alpha$.

For $\alpha \in (0,1)$, noticing that $Z_\alpha < \infty$, due to convention $0 \cdot \infty = 0$ and the fact that

$$\{E[\mathbb{1}_{\{m \le q_t^+(\alpha)\}}|\mathcal{F}_t] = 0\} \subseteq \{\mathbb{1}_{\{m \le q_t^+(\alpha)\}} = 0\} \cup B,$$

for some B, such that $\mathbb{P}[B] = 0$, we conclude that $Z_\alpha \in P_t$. Moreover, by the definition of $q_t^+(\alpha)$[6], there exists a sequence $Y_n \in L_t^0$, such that $Y_n \nearrow q_t^+(\alpha)$, and

$$E[\mathbb{1}_{\{m \le Y_n\}} \mid \mathcal{F}_t] \le \alpha.$$

Consequently, by monotone convergence theorem, we have

$$E[\mathbb{1}_{\{m \le q_t^+(\alpha)\}} \mid \mathcal{F}_t] \le \alpha.$$

Hence, we deduce

$$\mathbb{P}[m < q_t^+(\alpha)] = E[\mathbb{1}_{\{m < q_t^+(\alpha)\}}] \le E[E[\mathbb{1}_{\{m \le q_t^+(\alpha)\}}|\mathcal{F}_t]] \le E[\alpha] = \alpha,$$

which implies that

$$\mathbb{P}[m \ge q_t^+(\alpha)] \ge (1-\alpha). \tag{3.19}$$

On the other hand

$$\begin{aligned} \mathbb{1}_{\{m \ge q_t^+(\alpha)\}} m \ge \mathbb{1}_{\{m \ge q_t^+(\alpha)\}} q_t^+(\alpha) &= \mathbb{1}_{\{m \ge q_t^+(\alpha)\}} q_t^+(\alpha) E[Z_\alpha|\mathcal{F}_t] \\ &\ge \mathbb{1}_{\{m \ge q_t^+(\alpha)\}} E[Z_\alpha q_t^+(\alpha)|\mathcal{F}_t] \ge \mathbb{1}_{\{m \ge q_t^+(\alpha)\}} E[Z_\alpha m|\mathcal{F}_t], \end{aligned}$$

which combined with (3.19), implies that

$$P\Big[m \ge E[Z_\alpha m|\mathcal{F}_t]\Big] \ge 1-\alpha. \tag{3.20}$$

Hence, using (3.20), and the fact that

$$E[Z_\alpha m|\mathcal{F}_t] \ge \mu_t(m), \quad \alpha \in (0,1),$$

we get that

$$\mathbb{P}[m \ge \mu_t(m)] \ge 1-\alpha.$$

Letting $\alpha \to 0$, we conclude that (3.16) holds true for $m \in L^0$.

Now, assume that $m \in \bar{L}^0$, and let $A := \{E[\mathbb{1}_{\{m=-\infty\}}|\mathcal{F}_t] = 0\}$. Similar to the arguments above, we get

$$\mathbb{1}_A m \ge \mu_t(\mathbb{1}_A m). \tag{3.21}$$

Indeed, on $\{E[\mathbb{1}_{\{m=\infty\}}|\mathcal{F}_t] = 1\}$, inequality (3.21) is trivial and the set $\{E[\mathbb{1}_{\{m=\infty\}}|\mathcal{F}_t] < 1\}$ could be written as

$$\bigcup_{\alpha \in (0,1)} \{E[\mathbb{1}_{\{m=\infty\}}|\mathcal{F}_t] < 1-\alpha\}.$$

Next, on $\{E[\mathbb{1}_{\{m=\infty\}}|\mathcal{F}_t] < 1-\alpha\}$ we can set $Z_\alpha = 0$ in (3.17), which allow us to assume that $m \in L^0$ on that set. Finally, we let $\alpha \to 0$.

[6] Note that the family of random variables in (3.18) is upwards centered.

Since $\mu_t(0)=0$, and due to locality of μ_t, we deduce

$$\mathbb{1}_A m \geq \mu_t(\mathbb{1}_A m) = \mathbb{1}_A \mu_t(\mathbb{1}_A m) = \mathbb{1}_A \mu_t(m). \tag{3.22}$$

Moreover, taking $Z=1$ in (3.15), we get

$$\mathbb{1}_{A^c} m \geq \mathbb{1}_{A^c}(-\infty) = \mathbb{1}_{A^c} E[m|\mathcal{F}_t] \geq \mathbb{1}_{A^c}\mu_t(m). \tag{3.23}$$

Combining (3.22) and (3.23), we concludes the proof of (3.16) for all $m \in \bar{L}^0$.

Finally, we will show that μ_t defined as in (3.15) satisfies property 1) from Proposition 2.1.5, which will consequently imply equality (3.13). Let $m \in \bar{L}^0$ and $A \in \mathcal{F}_t$. From the fact that $m \geq \mu_t(m)$ we get

$$\operatorname*{ess\,inf}_{\omega\in A} m \geq \operatorname*{ess\,inf}_{\omega\in A} \mu_t(m).$$

On the other hand we know that $\mathbb{1}_A \operatorname{ess\,inf}_{\omega\in A} m \leq \mathbb{1}_A m$ and $\mathbb{1}_A \operatorname{ess\,inf}_{\omega\in A} m \in \bar{L}^0_t$. Thus, using Proposition 2.1.5, 2), we get

$$\begin{aligned}\operatorname*{ess\,inf}_{\omega\in A} m = \operatorname*{ess\,inf}_{\omega\in A}(\mathbb{1}_A \operatorname*{ess\,inf}_{\omega\in A} m) = \operatorname*{ess\,inf}_{\omega\in A}(\mathbb{1}_A \mu_t(\mathbb{1}_A \operatorname*{ess\,inf}_{\omega\in A} m)) \leq \\ \leq \operatorname*{ess\,inf}_{\omega\in A}(\mathbb{1}_A\mu_t(\mathbb{1}_A m)) = \operatorname*{ess\,inf}_{\omega\in A}(\mathbb{1}_A \mu_t(m)) = \operatorname*{ess\,inf}_{\omega\in A}\mu_t(m)\end{aligned}$$

which proves the equality. The proof for $\operatorname{ess\,sup}_t$ is similar and we omit it here. This concludes the proof. □

Recall that the case of random variables corresponds to $\mathcal{X}=L^p$, for a fixed $p\in\{0,1,\infty\}$. We proceed with the definition of weak acceptance and weak rejection time consistency (for random variables).

Definition 3.2.2 (Weak time consistency for random variables)**.** Let $\varphi=\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM. Then φ is said to be

- *Weakly acceptance time consistent* if it is $\mu^{\inf}$-acceptance time consistent,
- *Weakly rejection time consistent*, if it is $\mu^{\sup}$-rejection time consistent.

Definition 3.2.2 of time consistency is equivalent to many forms of time consistency studied in the current literature. Usually, the weak time consistency is considered for dynamic monetary risk measures on L^∞ (cf. [2] and references therein), to which we refer to as 'classical (benchmark) weak time consistency' (see Appendix A.2).

It was proved in [2] that in the classical weak time consistency framework, weak acceptance (respectively weak rejection) time consistency is equivalent to the statement that for any $X\in\mathcal{X}$ and $s>t$, we get

$$\varphi_s(X)\geq 0 \Rightarrow \varphi_t(X)\geq 0 \qquad \text{(resp. } \leq\text{)}. \tag{3.24}$$

This was the very starting point for our definition of weak acceptance (respectively weak rejection) time consistency, and the next proposition explains why so.

Proposition 3.2.3. *Let $\varphi=\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM. The following conditions are equivalent*

1) φ is weakly acceptance time consistent, i.e. for any $X \in \mathcal{X}$, $t,s \in \mathbb{T}$, $s > t$, and $m_s \in \bar{L}^0_s$,

$$\varphi_s(X) \geq m_s \Rightarrow \varphi_t(X) \geq \operatorname{ess\,inf}_t(m_s). \tag{3.25}$$

2) For any $X \in \mathcal{X}$, $s,t \in \mathbb{T}$, $s > t$, $\varphi_t(X) \geq \operatorname{ess\,inf}_t \varphi_s(X)$.

3) For any $X \in \mathcal{X}$, $s,t \in \mathbb{T}$, $s > t$, and $m_t \in \bar{L}^0_t$,

$$\varphi_s(X) \geq m_t \Rightarrow \varphi_t(X) \geq m_t.$$

If additionally $\{-\varphi_t\}_{t\in\mathbb{T}}$ is a dynamic monetary risk measure, then the above conditions are equivalent to

4) For any $X \in \mathcal{X}$ and $s,t \in \mathbb{T}$, $s > t$,

$$\varphi_s(X) \geq 0 \Rightarrow \varphi_t(X) \geq 0.$$

Similar result holds true for weak rejection time consistency.

Proof. We will only show the proof for acceptance consistency. The proof for rejection consistency is similar. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM.

1) $\Leftrightarrow$ 2). This is a direct application of Proposition 3.1.8.

1) $\Rightarrow$ 3). Assume that φ is weakly acceptance consistent, and let $m_t \in \bar{L}^0_t$ be such that $\varphi_s(X) \geq m_t$. Then, using Proposition 3.1.8, we get $\varphi_t(X) \geq \operatorname{ess\,inf}_t(\varphi_s(X)) \geq \operatorname{ess\,inf}_t(m_t) = m_t$, and hence 3) is proved.

3) $\Rightarrow$ 1). By the definition of conditional essential infimum, $\operatorname{ess\,inf}_t(\varphi_s(X)) \in \bar{L}^0_t$, for any $X \in \mathcal{X}$, and $t,s \in \mathcal{T}$. Moreover, by Proposition 2.1.5.(3), we have that $\varphi_s(X) \geq \operatorname{ess\,inf}_t(\varphi_s(X))$. Using 3) with $m_t = \operatorname{ess\,inf}_t(\varphi_s(X))$, we immediately obtain $\varphi_t(X) \geq \operatorname{ess\,inf}_t(\varphi_s(X))$. Due to Proposition 3.1.8 this concludes the proof.

3) $\Leftrightarrow$ 4). Clearly 3) $\Rightarrow$ 4). If additionally φ is a monetary risk measure, then in particular $-\varphi$ it is cash-additive. Hence, for any $m_t \in \bar{L}^0_t$ such that $\varphi_s(X) \geq m_t$, we have that $\varphi_s(X - m_t) \geq 0$, and since 4) holds true, we get that $\varphi_t(X - m_t) \geq 0$. Invoking one more time cash-additivity, we complete the proof. □

Property 3) in Proposition 3.2.3 was also suggested as the notion of (weak) acceptance and (weak) rejection time consistency in the context of scale invariant measures, called acceptability indices (cf. [14, 20]).

As next result shows, the weak time consistency is indeed one of the weakest forms of time consistency, being implied by any time consistency generated by a projective rule.

Proposition 3.2.4. *Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM and let μ be a projective update rule. If $\{\varphi_t\}_{t\in\mathbb{T}}$ is μ-acceptance (resp. μ-rejection) time consistent, then $\{\varphi_t\}_{t\in\mathbb{T}}$ is weakly acceptance (resp. weakly rejection) time consistent.*

Proof. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM, $\mu = \{\mu_t\}_{t\in\mathbb{T}}$ a projective update rule, and assume that $\{\varphi_t\}_{t\in\mathbb{T}}$ is μ-acceptance time consistent. Then, using Proposition 2.1.5, for any $t,s \in \mathbb{T}$, $s > t$, and any $X \in \mathcal{X}$, we get

$$\varphi_t(X) \geq \mu_t(\varphi_s(X)) \geq \mu_t(\operatorname{ess\,inf}_s(\varphi_s(X))) \geq \mu_t(\operatorname{ess\,inf}_t(\varphi_s(X))) = \operatorname{ess\,inf}_t(\varphi_s(X)).$$

The proof for rejection time consistency is similar. □

Remark 3.2.5. Recall that time consistency is preserved under monotone transformations, Remark 3.1.9. Thus, for any strictly monotone function $g : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$, if $\{\varphi_t\}_{t\in\mathbb{T}}$ is weakly acceptance (resp. weakly rejection) time consistent, then $\{g \circ \varphi_t\}_{t\in\mathbb{T}}$ also is weakly acceptance (resp. weakly rejection) time consistent.

3.2.2 Middle time consistency

Before we give the definition of middle acceptance/rejection time consistency, we need to show that any $\bar{L}^0$-extension of an dUM is an sX-invariant update rule, and we give necessary and sufficient conditions when this update rule is also projective. Moreover, we will use the notation from Section 2.2.2, i.e. for any dUM denoted by $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$, the maps $\varphi^- = \{\varphi_t^-\}_{t\in\mathbb{T}}$ and $\varphi^+ = \{\varphi_t^+\}_{t\in\mathbb{T}}$ will correspond to upper and lower $\bar{L}^0$-extensions of φ, respectively.

Proposition 3.2.6. *Any $\bar{L}^0$-extension $\widehat{\varphi}$ of a dUM φ is an sX-invariant update rule. Moreover, $\widehat{\varphi}$ is projective if and only if $\varphi_t(X) = X$, for $t \in \mathbb{T}$ and $X \in \mathcal{X} \cap \bar{L}_t^0$.*

Proof. The first part follows immediately from the definition of $\bar{L}^0$-extension. Clearly, projectivity of $\widehat{\varphi}$ implies that $\varphi(X) = X$, for $X \in \mathcal{X}_t$. To prove the opposite implication, it is enough to prove that φ^+ and φ^- are projective. Assume that φ is such that $\varphi_t(X) = X$, for $t \in \mathbb{T}$ and $X \in \mathcal{X}_t$. Let $X \in \bar{L}_t^0$. For any $n \in \mathbb{N}$, we get

$$\mathbb{1}_{\{n\geq X\geq -n\}}\varphi_t^+(X) = \mathbb{1}_{\{n\geq X\geq -n\}}\varphi_t^+(\mathbb{1}_{\{n\geq X\geq -n\}}X) = \mathbb{1}_{\{n\geq X\geq -n\}}\varphi_t(\mathbb{1}_{\{n\geq X\geq -n\}}X) = \mathbb{1}_{\{n\geq X\geq -n\}}X.$$

Thus, on set $\bigcup_{n\in\mathbb{N}}\{-n \leq X \leq n\} = \{-\infty < X < \infty\}$, we have

$$\varphi_t^+(X) = X, \qquad \text{for } X \in \bar{L}_t^0. \tag{3.26}$$

Next, for any $A \in \mathcal{F}_t$, such that $A \subseteq \{X = \infty\}$, we get $\mathcal{Y}_A^+(X) = \emptyset$, which implies

$$\mathbb{1}_{\{X=\infty\}}\varphi^+(X) = \infty.$$

Finally, for any $n \in \mathbb{R}$, using locality of φ_t^+ and the fact that $n \in \mathcal{X}_t$, we get

$$\mathbb{1}_{\{X=-\infty\}}\varphi_t^+(X) \leq \mathbb{1}_{\{X=-\infty\}}\varphi_t^+(\mathbb{1}_{\{X=-\infty\}}n) = \mathbb{1}_{\{X=-\infty\}}\varphi_t(n) = \mathbb{1}_{\{X=-\infty\}}n,$$

which implies $\mathbb{1}_{\{X=-\infty\}}\varphi^+(X) = -\infty$. Hence (3.26) holds true on entire space. The proof for φ^- is analogous. □

Let us start with the definition of middle acceptance and middle rejection time consistency.

Definition 3.2.7 (Middle time consistency for random variables)**.** Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM. Then φ is said to be

- *Middle acceptance time consistent* if it is φ^--acceptance time consistent.
- *Middle rejection time consistent*, if it is φ^+-rejection time consistent.

As in the case of weak time consistency, the notion of middle time consistency is usually presented for functions $\{-\varphi_t\}_{t\in\mathbb{T}}$ being dynamic monetary risk measures on L^∞ (cf. [2] and references therein). It is not difficult to prove (cf. [2]), that in L^∞ framework the middle acceptance (resp. middle rejection) time consistency is equivalent to the statement that

$$\varphi_t(X) \geq \varphi_t(\varphi_s(X)) \qquad (\text{resp. } \leq), \quad X \in \mathcal{X},\ s > t. \tag{3.27}$$

However, in case of a general domain of definition $\mathcal{X}$ of φ, we may have that $\varphi_s(X) \notin \mathcal{X}$ and, consequently, (3.27) cannot be used directly for time consistency. This is precisely the reason why we have introduced the $\bar{L}^0$-extensions. On the other hand, due to the fact that in Definition 3.2.7 the update rules are extensions, our concept of middle time consistency is stronger than the classical approach to middle time consistency, as shown in the next result.

Proposition 3.2.8. *Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM. The following two conditions are equivalent*

1) φ is middle acceptance time consistent, i.e. for any $X \in \mathcal{X}$, $s,t \in \mathbb{T}$, $s > t$, and $m_s \in \bar{L}^0_s$,

$$\varphi_s(X) \geq m_s \Rightarrow \varphi_t(X) \geq \varphi_t^-(m_s).$$

2) For any $X \in \mathcal{X}$, $s,t \in \mathbb{T}$, $s > t$,

$$\varphi_t(X) \geq \varphi_t^-(\varphi_s(X)).$$

If additionally $\{-\varphi_t\}_{t\in\mathbb{T}}$ is a dynamic monetary risk measure, then 1) or 2) implies

3) For any $X \in \mathcal{X}$, $s,t \in \mathbb{T}$, $s > t$, and $Y \in \mathcal{X} \cap \bar{L}^0_s$, we get

$$\varphi_s(X) \geq \varphi_s(Y) \Rightarrow \varphi_t(X) \geq \varphi_t(Y).$$

Analogous results are true for middle rejection time consistency,

The proof of the equivalence of 1) and 2) in Proposition 3.2.8 follows immediately from Proposition 3.1.8, and the proof that 1) implies 3) is straightforward upon taking $m_s = \varphi_s(Y)$.

Next, we will show that, in principle, middle acceptance time consistency is not suited for acceptability indices [20, 45].

Proposition 3.2.9. *Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM such that*

1) $\varphi_t(X) = \infty$, for any $t \in \mathbb{T}$ and $X \in \mathcal{X}$, such that $X \geq 0$ and $P[X > 0] > 0$;

2) there exists $X_0 \in \mathcal{X}$ and $t_1, t_2 \in \mathbb{T}$, $t_1 \neq t_2$, such that $0 < \varphi_{t_i}(X_0) < \infty$, for $i = 1, 2$.

Then, $\{\varphi_t\}_{t\in\mathbb{T}}$ is not middle acceptance time consistent.

Proof. Let us assume that φ satisfies 1), 2) and it is φ^--acceptance time consistent. Using Proposition 3.1.8 and the monotonicity of φ^-, we get

$$\infty > \varphi_{t_1}(X_0) \geq \varphi_{t_1}^-(\varphi_{t_2}(X_0)) \geq \varphi_{t_1}^-(\varphi_{t_2}(X_0) \wedge 1) = \varphi_{t_1}(\varphi_{t_2}(X_0) \wedge 1) = \infty,$$

which leads to contradiction. □

Remark 3.2.10. Properties 1) and 2) in Proposition 3.2.9 are characteristic for acceptability indices: the first property is related to 'arbitrage consistency' proposed in [45]; the second property is a technical assumption that eliminates degenerate cases. Thus, the concept of middle acceptance time consistency, and therefore (as seen in next section) the concept of strong time consistency, is not proper for such maps.

Remark 3.2.11. In general, middle acceptance/rejection time consistency does not imply weak acceptance/rejection time consistency. Indeed, let us consider $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$, such that $\varphi_t(X) = t$ (resp. $\varphi_t(X) = -t$) for all $X \in L^0$. Since $\varphi_t(0) = t \not\geq \operatorname{ess\,inf}_t \varphi_s(0) = s$ (resp. $-t \not\leq -s$), for $s > t$, we conclude that φ is not weakly acceptance (resp. weakly rejection) time consistent. On the other hand $\varphi_t(X) = \varphi_t(\varphi_s(X))$ for any $X \in L^0$, and hence φ is both middle acceptance and middle rejection time consistent.

3.2.3 Strong time consistency

The strong version of time consistency was one of the first one studied in the literature, in the context of dynamic (coherent and consequently convex) risk measures. There is an extensive literature on this subject (cf. [2, 7, 36, 57, 1, 36]. The key features of strong time consistency is its equivalence to Bellman's principle of optimality [7]. The definition that we will propose here will be slightly stronger (see Proposition 3.2.13), but nevertheless, the main idea will remain the same. Let us start with the definition of strong time consistency.

Definition 3.2.12 (Strong time consistency for random variables)**.** Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM. Then φ is said to be *strongly time consistent* if there exists $\widehat{\varphi}$, $\bar{L}^0$-extension of φ, such that the family φ is both $\widehat{\varphi}$-acceptance and $\widehat{\varphi}$-rejection time consistent.

Using (3.27), we have that strong time consistency for $\{-\varphi_t\}_{t\in\mathbb{T}}$ a dynamic monetary risk measure on L^∞ (see also [2] and references therein) is equivalent to the following property

$$\varphi_t(X) = \varphi_t(\varphi_s(X)), \quad \text{fro any} \quad X \in \mathcal{X},\ s > t, \tag{3.28}$$

known as Bellman's principle or dynamic programming principle. As mentioned in previous section, once the dUM is defined on larger space than L^∞, to make sense of dynamic programming principle, and thus strong time consistency, one needs to work with proper extensions of these function. Next key results show an alternative formulation for strong time consistency, that also has a clear financial interpretation.

Proposition 3.2.13. *Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM so that for any $t \in \mathbb{T}$, there exists $X \in \mathcal{X}$ such that $\varphi_t(X) = 0$. The following conditions are equivalent*

1) There exists update rule μ, such that μ is X-invariant and the family φ is both μ-acceptance and μ-rejection time consistent.

2) For any $X, Y \in \mathcal{X}$, $s, t \in \mathbb{T}$, $s > t$,

$$\varphi_s(X) = \varphi_s(Y) \Rightarrow \varphi_t(X) = \varphi_t(Y).$$

In particular 1) and 2) are satisfied if one of the following (equivalent) conditions hold

3) φ is strongly time consistent.

4) There exists $\widehat{\varphi}$, $\bar{L}^0$-extension of φ, such that for any $X \in \mathcal{X}$, $s,t \in \mathbb{T}$, $s > t$ we get

$$\varphi_t(X) = \widehat{\varphi}_t(\varphi_s(X)).$$

Proof. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM.
1) $\Rightarrow$ 2). Assume that μ is an X-invariant update rule, such that φ is both μ-acceptance and μ-rejection consistent. Then, by Theorem 3.1.8, $\varphi_t(X) = \mu_{t,s}(\varphi_s(X), 0)$, for any $t \in \mathbb{T}$ and $X \in \mathcal{X}$. Let $s,t \in \mathbb{T}$ and $X,Y \in \mathcal{X}$ be such that $s > t$ and $\varphi_s(X) = \varphi_s(Y)$. From the above, and by monotonicity of μ, we have

$$\varphi_t(X) = \mu_{t,s}(\varphi_s(X), 0) = \mu_{t,s}(\varphi_s(Y), 0) = \varphi_t(Y).$$

2) $\Rightarrow$ 1). Let $t,s \in \mathbb{T}$ be such that $s > t$, and consider the following set

$$\mathcal{X}_{\varphi_s} = \{X \in \bar{L}^0 \mid X = \varphi_s(Y) \text{ for some } Y \in \mathcal{X}\}.$$

From 2), for any $X,Y \in \mathcal{X}$, such that $\varphi_s(X) = \varphi_s(Y)$, we get $\varphi_t(X) = \varphi_t(Y)$. Thus, there exists a map $\phi_{t,s} : \mathcal{X}_{\varphi_s} \to \bar{L}^0_t$ such that

$$\phi_{t,s}(\varphi_s(X)) = \varphi_t(X), \quad X \in \mathcal{X}.$$

Next, since there exists $Z \in \mathcal{X}$, such that $\varphi_s(Z) = 0$, using locality of φ, we get that for any $X \in \mathcal{X}_{\varphi_s}$, $A \in \mathcal{F}_t$, there exist $Y \in \mathcal{X}$, so that[7]

$$\mathbb{1}_A X = \mathbb{1}_A\varphi_s(Y) = \mathbb{1}_A\varphi_s(\mathbb{1}_A Y) + \mathbb{1}_{A^c}\varphi_s(\mathbb{1}_{A^c} Z) = (\mathbb{1}_A + \mathbb{1}_{A^c})\varphi_s(\mathbb{1}_A Y + \mathbb{1}_{A^c} Z) = \varphi_s(\mathbb{1}_A Y + \mathbb{1}_{A^c} Z).$$

Thus, $\mathbb{1}_A X \in \mathcal{X}_{\varphi_s}$, for any $A \in \mathcal{F}_t$, $X \in \mathcal{X}_{\varphi_s}$. Hence, from 2) and locality of φ, for any $X,Y \in \mathcal{X}_{\varphi_s}$, $A \in \mathcal{F}_t$, we get

(A) $X \geq Y \Rightarrow \phi_{t,s}(X) \geq \phi_{t,s}(Y)$;

(B) $\mathbb{1}_A\phi_{t,s}(X) = \mathbb{1}_A\phi_{t,s}(\mathbb{1}_A X)$.

In other words, $\phi_{t,s}$ is local and monotone on $\mathcal{X}_{\varphi_s} \subseteq \bar{L}^0_s$. In view of Corollary 2.2.14, there exists an extension of $\phi_{t,s}$, say $\widehat{\phi}_{t,s} : \bar{L}^0_s \to \bar{L}^0_t$, which is local and monotone on $\bar{L}^0_s$. Finally, we take $\mu_{t,s} : \bar{L}^0_s \times \mathcal{X} \to \bar{L}^0_t$ defined by

$$\mu_{t,s}(m, X) := \widehat{\phi}_{t,s}(m), \qquad X \in \mathcal{X}, m \in \bar{L}^0_s.$$

Clearly the family $\mu_{t,s}$ is an X-invariant update rule, and thus, by Proposition 3.1.8, φ is both μ-acceptance and μ-rejection time consistent.

The proof of the second part of Proposition 3.2.13 is immediate. Clearly, 3) $\Rightarrow$ 1) and 3) $\Leftrightarrow$ 4), due to Proposition 3.1.8. $\square$

[7] Note that $\mathbb{1}_B\varphi_s(\mathbb{1}_B Z) = 0$ for any $B \in \mathcal{F}_t$ (as $\mathcal{F}_t \subseteq \mathcal{F}_s$).

Remark 3.2.14. Property 2) from Proposition 3.2.13 is what is referred in the existing literature as strong time consistency (see Appendix A.2). Note that strong time consistency introduced in Definition 3.2.12 is stronger than property 2) from Proposition 3.2.13. In particular, the update rule considered in Definition 3.2.12 is sX-invariant, while property 2) guarantees existence of update rule, which is just X-invariant.

Remark 3.2.15. On infinite time horizon the class of strongly time consistent dRMs coincided with the class of maps known as *Dynamic Entropic Risk Measures* (see Section 4.1.1 or [103] for details). On the other hand, on finite time horizon, the class of strongly time consistent dRMs is much richer. The typical scheme to obtain a strongly time consistent dRM from any dRM, is to use so called *recursive construction*, which we will introduce in Section 3.4.

Remark 3.2.16. Strong time consistency if crucial property in the theory of *Markov Risk Measures*. They provide theoretical framework for (risk-averse) dynamic programming equations for Markov Decision Processess (cf. [131, 129, 132, 133] and references therein).

3.2.4 Submartingales, supermartingales and robust expectations

The definition of projective update rule is strictly connected to the definition of so called (conditional) non-linear expectation (see for instance [49] for definition and related properties of non-linear expectation). In [125, 114], the authors made an important connections between non-linear expectations and dynamic risk measures. It was also shown (see, for instance, [33, 34] for details) that among dynamic convex risk measures, the dynamic coherent risk measures are the only ones which satisfy Jensen's inequality for dynamic maps; a property critically important in our framework, as it leads to projective update rules for which time consistency is invariant under concave transformations (see Proposition 3.2.18). One particularly important case is obtained by using as an update rule equal to the standard expectation operator. Finally, we want to mention that this type of time consistency in L^∞ framework, was studied in [57, Section 5] and is related to the definition of supermartingale and submartingale property.

Definition 3.2.17 (Supermartingale and Submartingale time consistency for random variables)**.** Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM and let $\mu = \{\mu_t\}_{t\in\mathbb{T}}$ be given by $\mu_t(m) = E[m|\mathcal{F}_t]$ (for $m \in \bar{L}^0$). Then φ is said to be

- *Supermartingale time consistent* if it is μ-acceptance time consistent, i.e. for any $X \in \mathcal{X}$, and $m_s \in \mathcal{F}_s$, we have
$$\varphi_s(X) \geq m_s \;\Rightarrow\; \varphi_t(X) \geq E[m_s|\mathcal{F}_t].$$
- *Submartingale time consistent* if it is μ-rejection time consistent, i.e. for any $X \in \mathcal{X}$, and $m_s \in \mathcal{F}_s$, we have
$$\varphi_s(X) \leq m_s \;\Rightarrow\; \varphi_t(X) \leq E[m_s|\mathcal{F}_t].$$

Next result is devoted to a more general class of updates rules, and hence concepts of time consistency, for which we do not give a specific name. The case of super/sub-martingale time consistency will correspond to the particular case of determining sets $D_t = \{1\}$.

Proposition 3.2.18. *Let $\{D_t\}_{t\in\mathbb{T}}$ be a determining family of random variables*[8] *and let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM. Consider the family of maps $\phi=\{\phi_t\}_{t\in\mathbb{T}}$ and $\phi'=\{\phi'_t\}_{t\in\mathbb{T}}$, $\phi_t,\phi'_t:\bar{L}^0\to\bar{L}^0_t$, given by*

$$\phi_t(m)=\operatorname*{ess\,inf}_{Z\in D_t}E[Zm|\mathcal{F}_t],\qquad \phi'_t(m)=\operatorname*{ess\,sup}_{Z\in D_t}E[Zm|\mathcal{F}_t]. \tag{3.29}$$

Then, the following statements hold true:

1) the families ϕ and ϕ' are projective update rules;

2) if $\{\varphi_t\}_{t\in\mathbb{T}}$ is ϕ-acceptance time consistent, then $\{g\circ\varphi_t\}_{t\in\mathbb{T}}$ is also ϕ-acceptance time consistent, for any increasing, and concave function $g:\bar{\mathbb{R}}\to\mathbb{R}$.

3) if $\{\varphi_t\}_{t\in\mathbb{T}}$ is ϕ'-rejection time consistent, then $\{g\circ\varphi_t\}_{t\in\mathbb{T}}$ is also ϕ'-rejection time consistent, for any increasing, and convex function $g:\bar{\mathbb{R}}\to\mathbb{R}$.

Proof. Let us consider $\{\phi_t\}_{t\in\mathbb{T}}$ and $\{\phi'_t\}_{t\in\mathbb{T}}$ as given in (3.29).
1) The proof of monotonicity and locality is straightforward. Finally, for any $t\in\mathbb{T}$, $\mathbb{Q}\in\mathcal{D}_t$ and $m\in\bar{L}^0_t$, we immediately get

$$E_{\mathbb{Q}}[m|\mathcal{F}_t]=1_{\{m\geq 0\}}mE_{\mathbb{Q}}[1|\mathcal{F}_t]+1_{\{m<0\}}(-m)E_{\mathbb{Q}}[-1|\mathcal{F}_t]=m,$$

and thus, $\phi_t(m)=\phi'_t(m)=m$, for any $m\in\bar{L}^0_t$. Hence, $\{\phi_t\}_{t\in\mathbb{T}}$ is projective.
2) Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM which is ϕ-rejection time consistent, and $g:\bar{\mathbb{R}}\to\bar{\mathbb{R}}$ be an increasing, concave function. Then, for any $X\in\mathcal{X}$, we get

$$g(\varphi_t(X))\geq g(\phi_t(\varphi_s(X))=g(\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t}E_{\mathbb{Q}}[\varphi_s(X)|\mathcal{F}_t])=\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t}g(E_{\mathbb{Q}}[\varphi_s(X)|\mathcal{F}_t]). \tag{3.30}$$

Next, by Jensen's inequality, we deduce

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t}g(E_{\mathbb{Q}}[\varphi_s(X)|\mathcal{F}_t])\geq\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t}E_{\mathbb{Q}}[g(\varphi_s(X))|\mathcal{F}_t]=\phi_t(g(\varphi_s(X))). \tag{3.31}$$

Combining (3.30) and (3.31), ϕ-acceptance time consistency of $\{g\circ\varphi_t\}_{t\in\mathbb{T}}$ follows.
3) The proof is analogues to 2). ☐

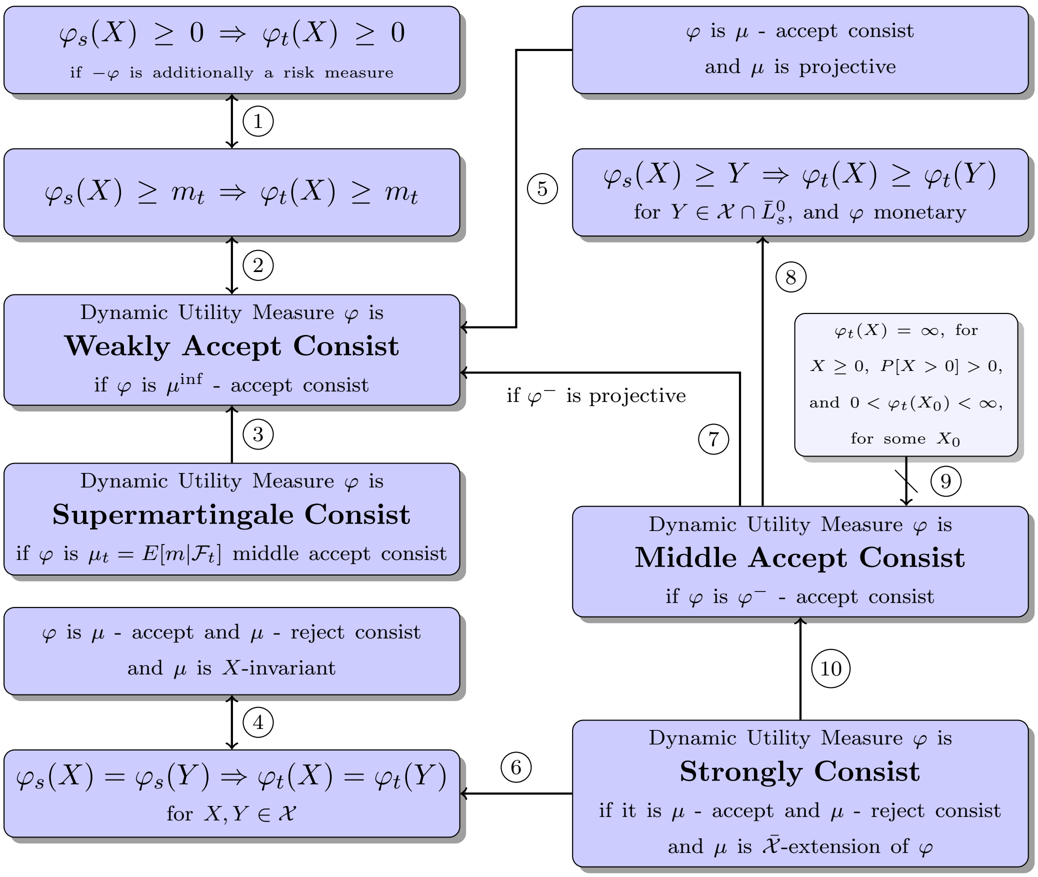
Remark 3.2.19. It could be easily shown from (3.13) that for any determining family of sets we get $\phi_t(m)\geq\operatorname{ess\,inf}_t m$. Thus, a dUM that is acceptance time consistent with respect to the update rule ϕ_t is also weakly acceptance time consistent. In particular, any supermartingale consistent dUM is also weakly acceptance time consistent. Similar statement holds true for rejection consistency.

3.2.5 Summary

The main goal of this section was to develop a unified framework for time consistency of dUMs that, in particular, comprises various types of time consistency for dynamic risk measures and dynamic performance measures known in the existing literature. The obtained results are summarised in the Chartflow 3.1. For convenience, we label (by circled numbers) each arrow (implication or equivalence) in the flowcharts, and we relate the labels to the relevant result from the sectio, along with comments on converse implications whenever appropriate.

[8] i.e. it is a non-empty family of random variables , such that for any $Z\in D_t$ we get $E[Z|\mathcal{F}_t]=1$, D_t is uniformly integrable, L^1-closed and $\mathcal{F}_t$-convex, for any $t\in\mathbb{T}$.

Figure 3.1: *Summary of results for acceptance time consistency for random variables*



(1) Proposition 3.2.3, 4)

(1) Proposition 3.2.3, 3)

(3) Remark 3.2.19 and Proposition 3.2.4. The converse implication is not true in general, see Section 4.1.1.

(4) Proposition 3.2.13, 1), 2)

(5) Proposition 3.2.4. Generally speaking the converse implication is not true. See for instance Section 4.1.1: negative of Dynamic Entropic Risk Measure with $\gamma < 0$ is weakly acceptance time consistent, but it is not supermaringale time consistent, i.e. it is not acceptance time consistent with respect to the projective update rule $\mu_t = E_t[m|\mathcal{F}_t]$.

(6) Proposition 3.2.13, 3), 4). The converse implication is not true in general. As a counterexample, consider $\varphi_t(X) = tE[X]$.

(7) Proposition 3.2.4, and see also (5). In general, middle acceptance time consistency does not imply weak acceptance time consistency, see Remark 3.2.11.

(8) Proposition 3.2.8, 3)

(9) Proposition 3.2.9

(10) Proposition 2.2.13. The converse implication is not true in general, see [116, Example 4.1].

3.3 Selected types of time consistency for stochastic processes

In this Section we assume that $\mathcal{X} = \mathbb{V}^p$, for $p \in \{0, 1, \infty\}$. The elements of $\mathcal{X}$, are interpreted as discounted dividend processes. It needs to be remarked, that all concepts developed for $\mathcal{X} = \mathbb{V}^p$ can be easily adapted to the case of cumulative discounted value processes (see Section 4.3.1 for details). While we preserve the same name for time consistency as in the case of random variables, the update rules for stochastic processes will differ significantly.

Usually, the case of stochastic processes is more intricate. If φ is a dUM, and $V \in \mathbb{V}^p$, then in order to compare $\varphi_t(V)$ and $\varphi_s(V)$, for $s > t$, one also needs to take into account the cash-flows between times t and s, so the update rule is not X-invariant in general.

In this Section we will briefly present the types of time consistency, which could be regarded as counterparts of the corresponding types from Section 3.2. We will not present the detailed proofs and comments, as they coincide with the previous case. It is worth mentioning, that in this section we will focus on one-step time consistency. See (3.11) and comments below, for the explanation of this approach. We will also present one new type of time consistency, which we will call semi-weak time consistency.

3.3.1 Weak time consistency

In this subsection we assume that $\mathcal{X} = \mathbb{V}^p$, for a fixed $p \in \{0, 1, \infty\}$, i.e. we consider the case of adapted stochastic processes.

Definition 3.3.1 (Weak time consistency for stochastic processes)**.** Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM (for stochastic processes). We say that φ is

- *Weakly acceptance time consistent* if it is one step μ-acceptance time consistent, where the update rule is given by
$$\mu_{t,t+1}(m, V) = \operatorname{ess\,inf}_t(m) + V_t.$$
- *Weakly rejection time consistent*, if it is one step μ-acceptance time consistent, where
$$\mu_{t,t+1}(m, V) = \operatorname{ess\,sup}_t(m) + V_t.$$

Similarly to Proposition 3.2.3, we have the following result.

Proposition 3.3.2. *Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM on $\mathbb{V}^p$. The following conditions are equivalent*

1) φ is weakly acceptance time consistent, i.e. for any $V \in \mathcal{X}$, $t < T$ $(t \in \mathbb{T})$ and $m_{t+1} \in \bar{L}^0_{t+1}$,
$$\varphi_{t+1}(V) \geq m_{t+1} \Rightarrow \varphi_t(X) \geq \operatorname{ess\,inf}_t(m_{t+1}) + V_t.$$

2) For all $V \in \mathcal{X}$, $t \in \mathbb{T}$, $t < T$,
$$\varphi_t(V) \geq \operatorname{ess\,inf}_t(\varphi_{t+1}(V)) + V_t.$$

3) For all $V \in \mathcal{X}$, $t \in \mathbb{T}$, $t < T$ and $m_t \in \bar{L}^0_t$,

$$\varphi_{t+1}(V) \geq m_t \Rightarrow \varphi_t(V) - V_t \geq m_t.$$

If additionally $\{-\varphi_t\}_{t\in\mathbb{T}}$ is a dynamic monetary risk measure, then the above conditions are equivalent to

3) For all $V \in \mathcal{X}$ and $t \in \mathbb{T}$, $t < T$,

$$\varphi_{t+1}(V) \geq 0 \Rightarrow \varphi_t(V) - V_t \geq 0.$$

Analogous results hold true for weak rejection time consistency.

Proof. We will only show the proof for weakly acceptance consistency. The proof for rejection consistency is similar. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM.

1) $\Leftrightarrow$ 2). This is a direct implication of Proposition 3.1.8.

1) $\Rightarrow$ 3). Let $m_t \in \bar{L}^0_t$ be such that $\varphi_{t+1}(V) \geq m_t$. Using the monotonicity of $\operatorname{ess\,inf}_t$, we have

$$\varphi_t(V) \geq \operatorname{ess\,inf}_t(\varphi_{t+1}(V)) + V_t \geq \operatorname{ess\,inf}_t(m_t) + V_t = m_t + V_t,$$

which concludes the proof.

3) $\Rightarrow$ 1). By Proposition 2.1.5, we get

$$\varphi_{t+1}(V) \geq \operatorname{ess\,inf}_t(\varphi_{t+1}(V)),$$

for any $V \in \mathcal{X}$, and $\operatorname{ess\,inf}_t(\varphi_{t+1}(X)) \in \bar{L}^0_t$. Using 3) with $m_t = \operatorname{ess\,inf}_t(\varphi_{t+1}(X))$ we immediately obtain

$$\varphi_t(V) \geq \operatorname{ess\,inf}_t(\varphi_{t+1}(V)) + V_t$$

and using 2) the weakly acceptance time consistency of φ follows.

3) $\Rightarrow$ 4) is obvious (take $m_t = 0$).

4) $\Rightarrow$ 3) Let us now assume that $\{\varphi_t\}_{t\in\mathbb{T}}$ is a negative of dynamic risk measure. For given $m_t \in \bar{L}^0_t$ it is enough to apply 3) to the process $V \pm 1_{t+1}m_t$, and 4) follows.

□

As mentioned earlier, the update rule, and consequently weak time consistency for stochastic processes, depends also on the value of the process (the dividend paid) at time t. If tomorrow, at time $t+1$, we accept $X \in \mathcal{X}$ at level greater than $m_{t+1} \in \mathcal{F}_{t+1}$, then today at time t, we will accept X at least at level $\operatorname{ess\,inf}_t m_{t+1}$ (i.e. the worst level of m_{t+1} adapted to the information $\mathcal{F}_t$) plus the dividend V_t received today.

Finally, we present the counterpart of Proposition 3.2.4 for the case of stochastic processes.

Proposition 3.3.3. *Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM on $\mathbb{V}^p$ and let ϕ be a projective update rule. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be one step μ-acceptance (resp. one step μ-rejection) time consistent, where μ is given by*

$$\mu_{t,t+1}(m, V) = \phi_t(m + V_t), \quad m \in \bar{L}^0_{t+1},\ V \in \mathcal{X}.$$

Then, $\{\varphi_t\}_{t\in\mathbb{T}}$ is weakly acceptance (resp. weakly rejection) time consistent.

The proof of Proposition 3.3.3 is analogous to the proof of Proposition 3.2.4, and we omit it.

3.3.2 Semi-weak time consistency

In this section we introduce the concept of semi-weak time consistency for stochastic processes. As it turns out, for the case of random variables semi-weak time consistency coincides with the definition of weak time consistency, hence omitted before. Thus, we take $\mathcal{X} = \mathbb{V}^p$, for a fixed $p \in \{0, 1, \infty\}$. As it was shown [20], none of the existing, at that time, forms of time consistency were suitable for scale-invariant maps, such as acceptability indices. In fact, even the weak acceptance and the weak rejection time consistency for stochastic processes (as defined in the present paper) are too strong in case of scale-invariant maps. Because of that we need even a weaker notion of time consistency, which we will refer to as semi-weak acceptance and semi-weak rejection time consistency. The notion of semi-weak time consistency for stochastic processes, introduced next, is suited precisely for such maps, and we refer the reader to [20] for a detailed discussion on time consistency for scale invariant measures and their dual representations[9].

Definition 3.3.4 (Semi-weak time consistency for stochastic processes)**.** Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM (for processes). Then φ is said to be:

- *Semi-weakly acceptance time consistent* if it is one step μ-acceptance time consistent, where the update rule is given by
$$\mu_{t,t+1}(m, V) = 1_{\{V_t\geq 0\}}\mu_t^{\inf}(m) + 1_{\{V_t<0\}}(-\infty).$$
- *Semi-weakly rejection time consistent* if it is one step μ'-rejection time consistent, where the update rule is given by
$$\mu'_{t,t+1}(m, V) = 1_{\{V_t\leq 0\}}\mu_t^{\sup}(m) + 1_{\{V_t>0\}}(+\infty).$$

It is straightforward to check that weak acceptance/rejection time consistency for stochastic processes always implies semi-weak acceptance/rejection time consistency.

Next, we will show that the definition of semi-weak time consistency is indeed equivalent to time consistency introduced in [20], that was later studied in [14, 16].

Proposition 3.3.5. *Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM on $\mathbb{V}^p$. The following conditions are equivalent*

1) φ is semi-weakly acceptance time consistent, i.e. for all $V \in \mathcal{X}$, $t \in \mathbb{T}$, $t < T$, and $m_t \in \bar{L}^0_t$,
$$\varphi_{t+1}(V) \geq m_{t+1} \Rightarrow \varphi_t(V) \geq \mathbb{1}_{\{V_t\geq 0\}} \operatorname{ess\,inf}_t(m_{t+1}) + 1_{\{V_t<0\}}(-\infty).$$

2) For all $V \in \mathcal{X}$ and $t \in \mathbb{T}$, $t < T$, $\varphi_t(V) \geq \mathbb{1}_{\{V_t\geq 0\}} \operatorname{ess\,inf}_t(\varphi_{t+1}(V)) + 1_{\{V_t<0\}}(-\infty)$.

3) For all $V \in \mathcal{X}$, $t \in \mathbb{T}$, $t < T$, and $m_t \in \bar{L}^0_t$, such that $V_t \geq 0$ and $\varphi_{t+1}(V) \geq m_t$, then $\varphi_t(V) \geq m_t$.

Similar result is true for semi-weak rejection time consistency.

[9] In [20] the authors combined both semi-weak acceptance and rejection time consistency into one single definition and call it time consistency.

Proof. We will only show the proof for acceptance consistency. The proof for rejection consistency is similar. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM.

1) $\Leftrightarrow$ 2). This is a direct implication of Proposition 3.1.8.

1) $\Rightarrow$ 3). Assume that φ is semi-weakly acceptance consistent. Let $V \in \mathcal{X}$ and $m_t \in \bar{L}^0_t$ be such that $\varphi_{t+1}(V) \geq m_t$ and $V_t \geq 0$. Then, using Proposition 3.1.8, we get

$$\varphi_t(V) \geq \mu_{t,t+1}(\varphi_{t+1}(V), V) = \mathbb{1}_{\{V_t\geq 0\}}\mu_t^{\inf}(\varphi_{t+1}(V)) \geq \mu_t^{\inf}(m_t) = \operatorname{ess\,inf}_t(m_t) = m_t,$$

and hence 3) is proved.

3) $\Rightarrow$ 2). Let $V \in \mathcal{X}$. We must show that

$$\varphi_t(V) \geq 1_{\{V_t\geq 0\}}\mu_t^{\inf}(\varphi_{t+1}(V)) + 1_{\{V_t<0\}}(-\infty). \tag{3.32}$$

On the set $\{V_t < 0\}$ inequality (3.32) is trivial. We know that

$$(\mathbb{1}_{\{V_t\geq 0\}} \cdot_t V)_t \geq 0 \quad \text{and} \quad \varphi_{t+1}(\mathbb{1}_{\{V_t\geq 0\}} \cdot_t V) \geq \operatorname{ess\,inf}_t \varphi_{t+1}(\mathbb{1}_{\{V_t\geq 0\}} \cdot_t V)$$

Thus, for $m_t = \operatorname{ess\,inf}_t \varphi_{t+1}(\mathbb{1}_{\{V_t\geq 0\}} \cdot_t V)$, using locality of φ and $\mu^{\inf}$ as well as 3), we get

$$\mathbb{1}_{\{V_t\geq 0\}}\varphi_t(V) = \mathbb{1}_{\{V_t\geq 0\}}\varphi_t(\mathbb{1}_{\{V_t\geq 0\}} \cdot_t V) \geq \mathbb{1}_{\{V_t\geq 0\}}m_t = \mathbb{1}_{\{V_t\geq 0\}}\mu_t^{\inf}(\varphi_{t+1}(V)).$$

and hence (3.32) is proved on the set $\{V_t \geq 0\}$. This conclude the proof of 2). □

Property 3) in Proposition 3.3.5 illustrates best the financial meaning of semi-weak acceptance time consistency: if tomorrow we accept the dividend stream $V \in \mathcal{X}$ at level m_t, and if we get a positive dividend V_t paid today at time t, then today we accept the cash-flow V at least at level m_t as well. Similar interpretation is valid for semi-weak rejection time consistency.

In the next section we will see (Propositions 4.2.3 and 4.2.4) that semi-weak time consistency appears naturally, when we study the connection between cash additive and scale invariant maps.

3.3.3 Middle time consistency

In this section we will adapt the middle time consistency to the case of stochastic processes, and we start with the definition of one step $\bar{L}^0$-extensions.

As before, for the case of stochastic processes we take $\mathcal{X} = \mathbb{V}^p$, for a fixed $p \in \{0, 1, \infty\}$. In what follows we will also make use of notation $\mathbb{T}' = \{0, 1, \ldots, T-1\}$.[10]

In this subsection, for a dUM $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$, we denote by $\widetilde{\varphi} = \{\widetilde{\varphi}_t\}_{t\in\mathbb{T}'}$ a family of maps $\widetilde{\varphi}_t : L^p_{t+1} \to \bar{L}^0_t$ given by

$$\widetilde{\varphi}_t(X) := \varphi_t(1_{\{t+1\}}X). \tag{3.33}$$

Since φ is monotone and local on $\mathbb{V}^p$, then, clearly, $\widetilde{\varphi}_t$ is local and monotone on L^p_{t+1}. Next, similar to the previous section, for any $t \in \mathbb{T}'$, we consider the extension of $\widetilde{\varphi}_t$ to $\bar{L}^0_{t+1}$, preserving locality and monotonicity (see Corollary 2.2.14). Note that formally $\widetilde{\varphi}$ is not a dUM, since the domain of the definition depends on $t \in \mathbb{T}'$, however, with slight abuse of notation, we will call such extension

[10] For infinite time horizon we get $\mathbb{T}' = \mathbb{T}$.

one step $\bar{L}^0$*-extension* of $\widetilde{\varphi}$. For any $\widetilde{\varphi}_t$ and $t \in \mathbb{T}'$, we consider the maps $\widetilde{\varphi}_t^+ : \bar{L}^0_{t+1} \to \bar{L}^0_t$ and $\widetilde{\varphi}_t^- : \bar{L}^0_{t+1} \to \bar{L}^0_t$ defined as in (2.10) and (2.11), with the sets $\mathcal{Y}^+_A(X)$ and $\mathcal{Y}^-_A(X)$ there replaced by

$$\mathcal{Y}^+_{t,A}(X) := \{Y \in L^p_{t+1} \mid \mathbb{1}_A Y \geq \mathbb{1}_A X\}, \qquad \mathcal{Y}^-_{t,A}(X) := \{Y \in L^p_{t+1} \mid \mathbb{1}_A Y \leq \mathbb{1}_A X\},$$

for any $X \in \bar{L}^0_{t+1}$, We will call $\widetilde{\varphi}^+$ and $\widetilde{\varphi}^-$ *upper and lower one step* $\bar{L}^0$*-extensions* of $\widetilde{\varphi}$, respectively. Now, we are ready to present the definition of middle acceptance and middle rejection time consistency for processes.

Definition 3.3.6 (Middle time consistency for stochastic processes)**.** Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM (for stochastic processes). Then φ is said to be

- *Middle acceptance time consistent* if it is one step μ-acceptance time consistent, where the update rule is given by
$$\mu_{t,t+1}(m, V) = \widetilde{\varphi}_t^-(m + V_t).$$
- *Middle rejection time consistent* if it is one step μ-rejection time consistent, where the update rule is given by
$$\mu_{t,t+1}(m, V) = \widetilde{\varphi}_t^+(m + V_t).$$

Proposition 3.3.7. *Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM on $\mathbb{V}^p$. The following conditions are equivalent*

1) φ is middle acceptance time consistent, i.e. for any $V \in \mathcal{X}$, $t \in \mathbb{T}'$ and $m_{t+1} \in \bar{L}^0_{t+1}$,

$$\varphi_{t+1}(V) \geq m_{t+1} \Rightarrow \varphi_t(V) \geq \widetilde{\varphi}_t^-(m_{t+1} + V_t).$$

2) For any $V \in \mathcal{X}$ and $t \in \mathbb{T}'$, $\varphi_t(V) \geq \widetilde{\varphi}_t^-(\varphi_{t+1}(V) + V_t)$.

If additionally $\{-\varphi_t\}_{t\in\mathbb{T}}$ is a dynamic monetary risk measure, then 1) or 2) implies

3) For any $V, V' \in \mathcal{X}$, and $t \in \mathbb{T}'$, we get

$$\varphi_{t+1}(V) \geq \varphi_{t+1}(1_{\{t+1\}}V'_{t+1}) \Rightarrow \varphi_t(V) \geq \widetilde{\varphi}_t^-(V'_{t+1} + V_t).$$

Analogous results are true for middle rejection time consistency.

The first part of Proposition 3.3.7 is a straightforward implication of Proposition 3.1.8. Since for cash additive measures $\varphi_{t+1}(1_{\{t+1\}}V'_{t+1}) = V'_{t+1}$, then, by taking $m_{t+1} = V'_{t+1}$ in 1), the second part follows immediately.

3.3.4 Strong time consistency

In this subsection we will use notation similar to the case of middle acceptance and middle rejection time consistency from Section 3.3.3.

Definition 3.3.8 (Strong time consistency for stochastic processes)**.** Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM (for stochastic processes). Then φ is said to be *strongly time consistent* if there exists $\widehat{\varphi}$, an one step $\bar{L}^0$-extension of $\widetilde{\varphi}$, such that φ is both one step μ-acceptance and one step μ-rejection time consistent with respect to

$$\mu_{t,t+1}(m, V) = \widehat{\varphi}_t(m + V_t).$$

Proposition 3.3.9. *Let $\varphi=\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM on $\mathbb{V}^p$. Assume that φ is independent of the past, and for any $t\in\mathbb{T}$, there exists $V\in\mathcal{X}$ such that $\varphi_t(V)=0$. The following two conditions are equivalent*

1) There exists an update rule μ, such that: for all $t\in\mathbb{T}'$, $m\in\bar{L}^0_t$, and $V,V'\in\mathcal{X}$, so that $V_t=V'_t$, we have $\mu_{t,t+1}(m,V)=\mu_{t,t+1}(m,V')$; the family φ is both one step μ-acceptance and one step μ-rejection time consistent.

2) For any $V,V'\in\mathcal{X}$, and $t\in\mathbb{T}'$,

$$V_t=V'_t \text{ and } \varphi_{t+1}(V)=\varphi_{t+1}(V')\Rightarrow\varphi_t(V)=\varphi_t(V').$$

In particular 1) and 2) are satisfied if one of the following (equivalent) conditions hold

3) φ is strongly time consistent.

4) There exists $\widehat{\varphi}$, one step $\bar{L}^0$-extension of $\widetilde{\varphi}$, such that for any $V\in\mathcal{X}$ and $t\in\mathbb{T}$ ($t<T$), we get $\varphi_t(V)=\widehat{\varphi}_t(\varphi_{t+1}(V)+V_t)$.

Proof. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM, which is independent of the past.

1) $\Rightarrow$ 2). Assume that μ is an update rule, fulfilling condition from 1), such that φ is both μ-acceptance and μ-rejection consistent. Then, by Proposition 3.1.8, $\varphi_t(X)=\mu_{t,t+1}(\varphi_{t+1}(X),Y)$, for any $t\in\mathbb{T}$ ($t<T$), $X\in\mathcal{X}$ and $Y\in\mathcal{X}$, such that $X_t=Y_t$. Let $t\in\mathbb{T}$ ($t<T$) and $X,Y\in\mathcal{X}$ be such that $X_t=Y_t$ and $\varphi_{t+1}(X)\geq\varphi_{t+1}(Y)$. From the above, and by monotonicity of μ, we have

$$\varphi_t(X)=\mu_{t,t+1}(\varphi_{t+1}(X),X)=\mu_{t,t+1}(\varphi_{t+1}(X),Y)\geq\mu_{t,t+1}(\varphi_{t+1}(Y),Y)=\varphi_t(Y).$$

2) $\Rightarrow$ 1). Let $t\in\mathbb{T}$ be such that $t<T$ and consider the following set

$$\mathcal{X}_{\varphi_{t+1}}=\{X\in\bar{L}^0\mid X=\varphi_{t+1}(Y)\text{ for some }Y\in\mathcal{X}\}.$$

From 2), for any $X,Y\in\mathcal{X}$, such that $\varphi_{t+1}(X)=\varphi_{t+1}(Y)$ and $X_t=Y_t$, we get $\varphi_t(X)=\varphi_t(Y)$. Thus, using independence of the past of φ, there exists a map $\phi_{t,t+1}:\mathcal{X}_{\varphi_{t+1}}\times L^p_t\to\bar{L}^0_t$ such that

$$\phi_{t,t+1}(\varphi_{t+1}(X),Y_t)=\varphi_t(X-1_{\{t\}}(X_t-Y_t)),\quad X\in\mathcal{X}.$$

Next, since there exists $Z\in\mathcal{X}$, such that $\varphi_{t+1}(Z)=0$, using locality of φ, we get that for any $X\in\mathcal{X}_{\varphi_{t+1}}$, $A\in\mathcal{F}_t$, there exist $Y\in\mathcal{X}$, so that

$$\mathbb{1}_AX=\mathbb{1}_A\varphi_{t+1}(Y)=\mathbb{1}_A\varphi_{t+1}(\mathbb{1}_A\cdot_{t+1}Y)+\mathbb{1}_{A^c}\varphi_{t+1}(\mathbb{1}_{A^c}\cdot_{t+1}Z)=\varphi_{t+1}(\mathbb{1}_A\cdot_{t+1}Y+\mathbb{1}_{A^c}\cdot_{t+1}Z).$$

Thus, $\mathbb{1}_AX\in\mathcal{X}_{\varphi_{t+1}}$, for any $A\in\mathcal{F}_t$, $X\in\mathcal{X}_{\varphi_{t+1}}$. Hence, from 2) and locality of φ, for any $X,X'\in\mathcal{X}_{\varphi_{t+1}}$, $Y_t\in L^p_t$ and $A\in\mathcal{F}_t$, we get

(A) $X\geq X'\Rightarrow\phi_{t,t+1}(X,Y_t)\geq\phi_{t,t+1}(X',Y_t)$;

(B) $\mathbb{1}_A\phi_{t,t+1}(X,Y_t)=\mathbb{1}_A\phi_{t,t+1}(\mathbb{1}_AX,Y_t)$.

In other words, for any fixed $Y_t \in L_t^p$, $\phi_{t,t+1}(\cdot, Y_t)$ is local and monotone on $\mathcal{X}_{\varphi_{t+1}} \subseteq \bar{L}^0_{t+1}$. In view of Corollary 2.2.14, for any fixed $Y_t \in L_t^p$ there exists an extension (to $\bar{L}^0_{t+1}$) of $\phi_{t,t+1}(\cdot, Y_t)$, say $\widehat{\phi}_{t,t+1}(\cdot, Y_t)$, which is local and monotone on $\bar{L}^0_{t+1}$. Finally, we take $\mu_{t,t+1} : \bar{L}^0_{t+1} \times \mathcal{X} \to \bar{L}^0_t$ defined by

$$\mu_{t,t+1}(m, X) := \widehat{\phi}_{t,t+1}(m, X_t), \qquad X \in \mathcal{X}, m \in \bar{L}^0_{t+1}.$$

Clearly the family $\mu_{t,t+1}$ is a (one step) update rule. Moreover, we get

$$\mu_{t,t+1}(m, X) = \mu_{t,t+1}(m, X'),$$

for $m \in \bar{L}^0_{t+1}$ and $X, X' \in \mathcal{X}$, such that $X_t = X'_t$. Finally, by Proposition 3.1.8, φ is both μ-acceptance and μ-rejection time consistent, as

$$\varphi_t(X) = \varphi_t(X - 1_{\{t\}}(X_t - X_t)) = \phi_{t,t+1}(\varphi_{t+1}(X), X_t) = \mu_{t,t+1}(\varphi_{t+1}(X), X).$$

The proof of the second part of Proposition 3.3.9 is immediate. Clearly, 3) $\Rightarrow$ 1) and 3) $\Leftrightarrow$ 4), due to Proposition 3.1.8. $\square$

3.3.5 Submartingales, supermartingales and robust expectations

The sub/super-martingale time consistency is defined similarly, by considering one step update rules of the form $\mu_{t,t+1}(m, V) = E[m|\mathcal{F}_t] + V_t$. Similar to Proposition 3.2.18, we have that time consistency property generated by updates rules of the form $\mu_{t,t+1}(m, V) = \phi_t(m+V_t)$ are invariant under concave/convex transformations.

Proposition 3.3.10. *Let $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ be a dUM (for processes). Let a one step update rule $\mu = \{\mu_t\}_{t\in\mathbb{T}}$ be given by $\mu_{t,t+1}(m, V) = \phi_t(m + V_t)$, for $\{\phi_t\}_{t\in\mathbb{T}}$ defined in (3.29). Then*

1) if φ is μ-acceptance time consistent, then $g \circ \varphi = \{g \circ \varphi_t\}_{t\in\mathbb{T}}$ also is μ-acceptance time consistent, for any increasing, and concave function $g : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$.

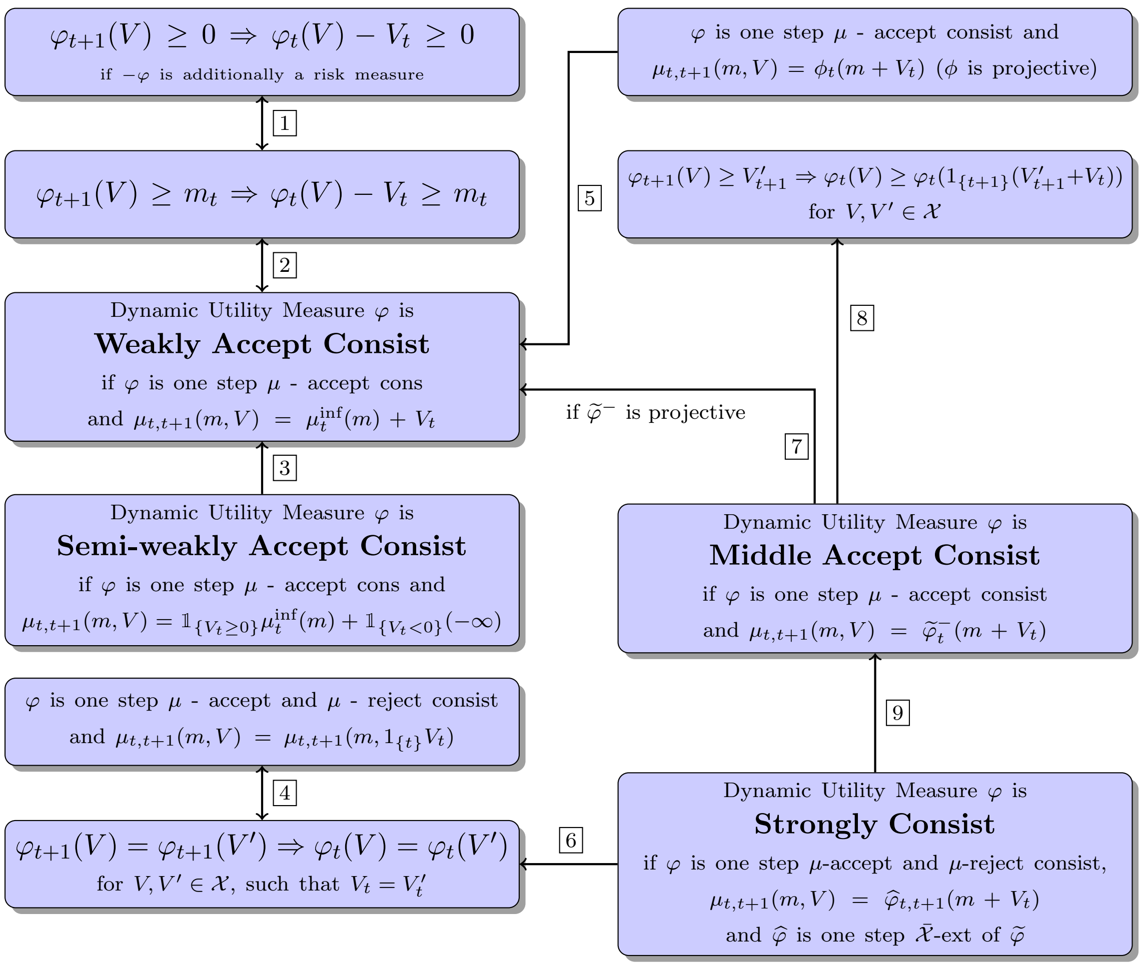
2) if φ is μ-rejection time consistent, then $g \circ \varphi = \{g \circ \varphi_t\}_{t\in\mathbb{T}}$ also is μ-rejection time consistent, for any increasing, and convex function $g : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$.

The proof of Proposition 3.3.10 is analogous to the proof of Proposition 3.2.18.

3.3.6 Summary

The main goal of this section was to develop a unified framework for time consistency of dUMs for stochastic processes that, in particular, comprises various types of time consistency for dynamic risk measures and dynamic performance measures known in the existing literature. The obtained results are summarised in the Chartflow 3.2. For convenience, we label (by rectangled numbers) each arrow (implication or equivalence) in the flowcharts, and we relate the labels to the relevant result from this section, along with comments on converse implications whenever appropriate.

Figure 3.2: *Summary of results for acceptance time consistency for stochastic processes*



$\boxed{1}$ Proposition 3.3.2, 4)

$\boxed{2}$ Proposition 3.3.2, 3)

$\boxed{3}$ Proposition 3.3.5, 3)

$\boxed{4}$ Proposition 3.3.9, 1), 2)

$\boxed{5}$ Proposition 3.3.3

$\boxed{6}$ Proposition 3.3.9, 3), 4)

$\boxed{7}$ Proposition 3.3.3, and see also 5.

$\boxed{8}$ Proposition 3.3.7, 3)

$\boxed{9}$ Proposition 2.2.13.

Remark 3.3.11. The converse implications in Flowchart 3.2 do not hold true in general, and one can use the same counterexamples as in the case of random variables.

3.4 Recursive construction on finite time horizon

As we have explained in the previous section, strong time consistency could be considered as a form of Bellman's principle, which is very convenient, when we want to conduct the dynamic portfolio optimisation. In risk measure theory on L^∞ a common tool used to construct a strongly-time consistent dRM (from any dRM) is so called *recursive construction*, introduced in [36, Section 4.2] and studied e.g. in [2, Section 4.4] or [40].[11] One could also show, that any strongly time consistent dRM could be written as a composition of (one-step) conditional RMs, so there exists duality between those two approaches [40]. For some specific families, the dynamic risk measures obtained using recursive construction are referred to (especially in markov coherent risk measure theory) as *multiperiod* or *composite dRMs* and the recursive construction is called *nested composition* (see e.g. [132, 129, 131, 130]).

On L^∞, given a dRM (which negative is denoted by $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$), we can construct a mapping $\widetilde{\varphi} = \{\widetilde{\varphi}_t\}_{t\in\mathbb{T}}$, defined recursively, where $\widetilde{\varphi}_T(X) = X$ and where for $(t = T-1, \ldots, 0)$, we set

$$\widetilde{\varphi}_t(X) = \varphi_t(\widetilde{\varphi}_{t+1}(X)). \tag{3.34}$$

It is easy to see that $\widetilde{\varphi}$ will be a strongly time consistent dRM (see [2, Prop. 1.16]).

Unfortunately, for L^p, when $p \neq \infty$, this approach cannot be always used directly, as we might get $\widetilde{\varphi}_t(X) \notin L^p$ for some $t \in \mathbb{T}$. Nevertheless, one could overcome this difficulty, considering the subset of L^p or extend dRM as in Subsection 2.2.2 and then use construction (3.34). One of many ways of doing this, is to use Robust representation and simply consider the extended conditional expectation operator and conditional penalty function (see e.g. [40, 44]).

Let us now show how to obtain such construction for representable coherent dRMs (see [40] for the class of representable convex dRMs). We know that any representable coherent dRM is associated with the family of measures $\{\mathcal{D}_t\}_{t\in\mathbb{T}}$, through robust representation.

Definition 3.4.1 (Determining family of measures)**.** We will say that $\mathcal{D} = \{\mathcal{D}_t\}_{t\in\mathbb{T}}$ is a determining family of measures, if $\mathcal{D}_t \subseteq \mathcal{Q}^1_t$, $\mathcal{D}_t \neq \emptyset$, $\mathcal{D}_t$ is uniformly integrable, L^1-closed and $\mathcal{F}_t$-convex, for any $t \in \mathbb{T}$.[12]

We refer to [43, Section 2] for a discussion about determining families. It is straightforward to check, that any determining family define a representable coherent dRM. Given a determining family, we can also construct a dRM recursively both for random variables as well as for stochastic processes, as will be shown in (3.35) and (5.37). We will refer to such dRM as $\mathcal{D}$-composite dRM and to it's negative as $\mathcal{D}$-composite dUM, which we will now define.

Definition 3.4.2 ($\mathcal{D}$-composite dUM for random variables)**.** Let $\mathcal{D} = \{\mathcal{D}_t\}_{t\in\mathbb{T}}$ be a determining family of measures. We will call a family $\{\varphi_t\}_{t\in\mathbb{T}}$ of mappings $\varphi_t : \bar{L}^0 \to \bar{L}^0_t$ a *$\mathcal{D}$-composite dUM* (for random variables), if $\{\varphi_t\}_{t\in\mathbb{T}}$ is defined as:

$$\begin{aligned} \varphi_T(X) &:= X \\ \varphi_t(X) &:= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t} E_{\mathbb{Q}}[\varphi_{t+1}(X)|\mathcal{F}_t]. \end{aligned} \tag{3.35}$$

[11] Please note that if time horizon is infinite, then the class of strongly time consistent dRMs coinccides with the class of Dynamic Entropic Risk Measures [103]. See Remark 4.1.4.

[12] in the sense, that the corresponding sets of Radon-Nikodem derivatives (i.e. random variables) admit those properties; $\mathcal{F}_t$-convex, i.e. for any $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{D}_t$ and $\lambda \in L^0_t$ such that $0 \leq \lambda \leq 1$, we get $\mathbb{Q}_3 \in \mathcal{D}_t$, where $\mathbb{Q}_3$ is such that $\frac{d\mathbb{Q}_3}{d\mathbb{P}} = \lambda\frac{d\mathbb{Q}_1}{d\mathbb{P}} + (1-\lambda)\frac{d\mathbb{Q}_2}{d\mathbb{P}}$.

Remark 3.4.3. The idea of construction (3.35) coincides with the construction (3.34). Note that on L^∞, the essential idea is to require the map $\varphi_t(X)$ to admit representation

$$\varphi_t(X) = \varphi_t(\varphi_{t+1}(\ldots\varphi_{T-1}(\varphi_T(X))\ldots))$$

for ay $X \in L^\infty$ and $t \in \mathbb{T}$, which implies strong time consistency of the corresponding dUM $\{\varphi_t\}_{t\in\mathbb{T}}$. On a bigger space, we can consider the space of all Xs, which admit the above representation, which define the biggest space on which $\{\varphi_t\}_{t\in\mathbb{T}}$ is strongly time consistent – the reason we will consider spaces defined in (3.39) and (3.40). See [43] for details.

Similarly, one could define a $\mathcal{D}$*-composite dUM* for stochastic processes.

Definition 3.4.4 ($\mathcal{D}$-composite dUM for stochastic processes)**.** Let $\mathcal{D} = \{\mathcal{D}_t\}_{t\in\mathbb{T}}$ be a determining family of measures. We will call a family $\{\varphi_t\}_{t\in\mathbb{T}}$ of mappings $\varphi_t : \bar{\mathbb{V}}^0 \to \bar{L}^0_t$ a $\mathcal{D}$*-composite dUM* (for stochastic processes), if $\{\varphi_t\}_{t\in\mathbb{T}}$ is defined as:

$$\begin{aligned} \varphi_T(V) &:= V_T \\ \varphi_t(V) &:= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t} E_{\mathbb{Q}}[V_t + \varphi_{t+1}(V)|\mathcal{F}_t]. \end{aligned} \tag{3.36}$$

The $\mathcal{D}$-composite dRMs and dUMs admit simpler representation for random variables and stochastic processes, which admit additional integrability conditions (see [43] for details). Given the determining family $\mathcal{D} = \{\mathcal{D}_t\}_{t\in\mathbb{T}}$, let $\{\mathcal{D}_{t,t+1}\}_{t\in\mathbb{T}}$ be a family of sets of measures given by

$$\mathcal{D}_{T,T+1} := \mathcal{D}^{\nu}_T, \tag{3.37}$$

$$\mathcal{D}_{t,t+1} := \{\mathbb{Q} \in \mathcal{Q}^1_t \mid \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = E[\frac{\mathrm{d}\mathbb{Q}'}{\mathrm{d}\mathbb{P}}|\mathcal{F}_{t+1}],\ \text{for } \mathbb{Q}' \in \mathcal{D}_t\} \quad (t < T). \tag{3.38}$$

Please note that $\{\mathcal{D}_{t,t+1}\}_{t\in\mathbb{T}}$ define the same composite dUM and the family $\{\mathcal{D}_{t,t+1}\}_{t\in\mathbb{T}}$ is also a determining family. Next, we define new families of measures

$$\widetilde{\mathcal{D}} := \{\mathbb{Q} \in \mathcal{Q}^1_t| \ \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \prod_{s=0}^{T} \frac{\mathrm{d}\mathbb{Q}_s}{\mathrm{d}\mathbb{P}},\ \text{and } \{\mathbb{Q}_s\}_{s=t}^T \text{ is such that } \mathbb{Q}_s \in \mathcal{D}_{s,s+1}\},$$

$$\bar{\mathcal{D}} := \{\mathbb{Q} \in \mathcal{Q}^1_t| \ \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \prod_{s=0}^{T} \frac{\mathrm{d}\mathbb{Q}_s}{\mathrm{d}\mathbb{P}},\ \text{and } \{\mathbb{Q}_s\}_{s=0}^T \text{ is such that } \mathbb{Q}_s \in \mathcal{D}_{s,s+1} \cup \{1\}\}.$$

Please note that $\mathcal{D}_{t,t+1} \subseteq \bar{\mathcal{D}}$ for any $t \in \mathbb{T}$. The *strong* $L^1_s(\mathcal{D})$*-space* and *weak* $L^1_s(\mathcal{D})$*-space* for a given set of measures (see [41, Section 2.2] for more details) are defined by

$$L^1_s(\mathcal{D}) := \Big\{X \in L^0 : \lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{D}} E_{\mathbb{Q}}\big[\mathbb{1}_{\{|X|>n\}}|X|\ \big] = 0\Big\} \tag{3.39}$$

$$L^1_w(\mathcal{D}) := \Big\{X \in L^0 : |\inf_{\mathbb{Q}\in\mathcal{D}} E_{\mathbb{Q}}[\pm X]| < \infty\Big\} \tag{3.40}$$

Note that if $\mathcal{D} = \{\mathbb{Q}\}$, then $L^1_s(\mathcal{D}) = L^1_w(\mathcal{D}) = L^1(\Omega, \mathcal{F}, \mathbb{Q})$, which justifies this notation. Moreover, we might get $L^1_s(\mathcal{D}) \neq L^1_w(\mathcal{D})$ (see [42] for example). Nevertheless, for many families of dUMs those two spaces coincide.

Proposition 3.4.5. *Let $\mathcal{D}$ denote a determining family of measures and let φ denote the corresponding $\mathcal{D}$-composite dUM defined for stochastic processes. For $V \in \mathbb{V}^0$, such that $V_t \in L^1_s(\bar{\mathcal{D}})$ for any $t \in \mathbb{T}$, we get*

$$\varphi_t(V) = \operatorname*{ess\,inf}_{\mathbb{Q} \in \widetilde{\mathcal{D}}} E_{\mathbb{Q}}\Big[\sum_{i=t}^{T} V_i \Big| \mathcal{F}_t\Big].$$

Moreover, there exists a minimiser for any $V \in \mathbb{V}^0$, such that $V_t \in L^1_s(\bar{\mathcal{D}})$ for any $t \in \mathbb{T}$, i.e. for any $t \in \mathbb{T}$, we get $\varphi_t(V) = E_{\mathbb{Q}^*}\Big[\sum_{i=t}^{T} V_i \Big| \mathcal{F}_t\Big]$ for some $\mathbb{Q}^* \in \widetilde{\mathcal{D}}$.

The proof of Proposition 3.4.5 could be found in [43, Prop. 2.1].

Chapter 4

Selected families of dynamic risk and performance measures

In this Chapter we will introduce three families of dUMs and show some recognisable representatives from each family.

Firstly, we will introduce the family of convex (and thus coherent) dRMs. This family of (negatives of) dUMs has attracted significant attention in the literature recently. The need to understand how to measure the risk, what is the risk and finally, how one can influence the risk (e.g. through dynamic control) naturally lead to the class of convex dRMs. The axiomatic approach to RMs initiated in [6] has attracted significant attention in Mathematical Finance (cf. [77, 7, 78] and references therein for literature overview), as the properties like convexity (CV), monotonicity (MD) or normalization (N) have a natural financial interpretation (see Remark 2.2.2). Apart from selected basic facts, which we will use in the next Chapter, we will introduce three important families of convex dUMs, often used in stochastic control problems, due to their traceability.

Secondly, we will introduce the class of dynamic performance measures, which additionally satisfy quasi-concavity (QCC), namely dynamic acceptability indices. This class of maps was introduced in [45], and studied (also for the dynamic case) e.g. in [126, 15, 20]. One can show very tight connection between dRMs and dynamic acceptability indices, which justifies the importance of this class. It it used to quantify the performance of a financial position. This class might measure the degree of arbitrage consistency in the market, compare financial positions or present the ratio between risk and reward (see e.g. [65, 9, 39] for details). Very often when we deal with stochastic control problem with risk constraints, it is convenient to transform it to the problem with single objective function. Usually such function is an acceptability index, as will be explained in Section 5.2.

Finally, we will introduce the class of Dynamic limit growth indices. This class of maps is designed to measure the long-run performance of a financial portfolio. Importance of measurement of the long run growth of a portfolio is widely recognized among financial practitioners, and has been extensively discussed in the literature (see for instance [5, 71], and references therein). Here, we shall focus on measures that quantify the tradeoff between portfolio growth and the risk associated with it, appropriately normalized in time. Among several such possible measures, the one which has attracted the most attention, is the so called Risk Sensitive Criterion [142, 21, 22]. While, this class might be in fact considered as a subclass of dynamic acceptability indices, we treat this case separately, as they are introduced in an entirely different framework.

4.1 Dynamic convex and coherent risk measures

The family of convex dRMs is (see page 12 for definition) is perhaps the most important class of monetary dRMs as it allow us to use tools from convex analysis (e.g. in portfolio optimisation [3]) and in particular provide the robust representation of the form

$$\rho(X) = -\inf_{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P})}[E_{\mathbb{Q}}[X] + \alpha^{\min}(\mathbb{Q})], \tag{4.1}$$

for some penalty function $\alpha^{\min} : \mathcal{M}_1(\mathbb{P}) \to \mathbb{R}\cup\{\infty\}$ (see Subsection 2.2.3 for details and the dynamic equivalent). We will now present some subclasses of convex RMs, which are often used in stochastic control theory and show some of their's properties. Nevertheless, we will not present here overview of the whole convex risk measurement theory, as this is not a main topic of this thesis. For a general good brief survey about convex RMs see e.g. [75].

In Subsection 4.1.1, we will introduce the family of *Dynamic entropic risk measures*, which are convex, but not coherent. This family is widely used in finance and other fields of applied mathematics (cf. [46] and references therein).

Next, in Subsection 4.1.2 we will define *Dynamic Tail Value at Risk*, which represents a class of coherent dRMs. Let us now explain, why this family of maps plays a crucial role in the coherent framework. In general, we know that a representable risk measure is coherent, if the penalty function only takes values in the set $\{0,\infty\}$. Thus, every representable coherent dRM $\varphi = \{\varphi_t\}_{t\in\mathbb{T}}$ could be expressed as

$$\rho_t(X) = -\inf_{\mathbb{Q}\in\mathcal{Q}_t} E_{\mathbb{Q}}[X|\mathcal{F}_t], \tag{4.2}$$

for $\{\mathcal{Q}_t\}_{t\in\mathbb{T}}$, such that $\mathcal{Q}_t \subseteq \mathcal{Q}_t^1$. Of course, different families $\{\mathcal{Q}_t\}_{t\in\mathbb{T}}$ could define the same coherent dRM (for a given $\mathcal{X}$), but usually one could define the largest family of sets, for which the representation (4.2) will hold (its called the determining family for a coherent dRM [41]). One of the most interesting results in the theory of coherent RMs is so called Kusuoka's Theorem, which shows that the family of *Tail Value at Risk* maps could be used as a building blocks for any coherent dRM, which admits law invariance (LI)[1]. Let us present this theorem for L^∞ (similar result could be obtained for the general conditional case on L^p [52]; there exists also a generalisation of this theorem for convex dRMs, admitting (LI) [84]).

Theorem 4.1.1 (Kusuoka's Theorem)**.** *Let $\mathcal{X} = L^\infty$ and let ρ be a coherent RM. Then ρ is (LI) if and only if there exists a (compact, convex) set $\mathcal{M}$ of probability measures on $(0,1]$, such that*

$$\rho(X) = -\inf_{\nu\in\mathcal{M}}\int_0^1 \rho^\alpha(X)\nu(\mathrm{d}\alpha), \tag{4.3}$$

where $\rho^\alpha(X) = -\operatorname{ess\,inf}_{\mathbb{Q}\in\mathcal{D}^\alpha} E_{\mathbb{Q}}[X]$ and $\mathcal{D}^\alpha = \{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P}) : \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \leq \alpha^{-1}\}$ for $p\in(0,1]$.

The family of coherent RMs $\{\rho^\alpha\}_{\alpha\in(0,1]}$, which appears in Theorem 4.1.1 is precisely the family of Tail Value at Risk RMs.

Finally, in Subsection 4.1.3 we will introduce the class of *Weighted Value at Risk* dRMs [41], sometimes also called *Spectral dRMs* [4]. In the static framework, this family is obtained, choosing

[1] For a general non-atomic Ω, we take the risk measure known as *Exptected Shortfall* as a building block. Nevertheless, in our framework those two families coincide. See [4] for details.

a singleton in (4.3), i.e. $\mathcal{M} = \{\nu\}$ for some fixed probability measure ν on $(0,1]$. It is worth noticing that this class of maps coincides with coherent risk measures which admit law invariance (LI) and comonotonicity[2]. The name *Spectral risk measure* is justified by different representation for this class of maps. Any Weighted Value at Risk RM admits equality

$$\rho(X) := -\int_0^1 q^p(X)\phi(p)\,\mathrm{d}p, \tag{4.4}$$

for some function ϕ, where $q^p(X)$ denotes the (upper) α-quantile of X and ϕ is an admissible risk spectrum, i.e. $\phi : [0,1) \to [0,\infty)$ is a right-continous, decreasing function such that $\int_0^1 \phi(p)dp = 1$ (cf. [41] and references therein). The relation between risk spectrum ϕ and probability measure ν is expressed through equation

$$\phi(t) = \int_{(t,1]} \frac{1}{s}\nu(ds).$$

For other interesting families of convex dRMs see e.g. [75, Section 4].

4.1.1 Dynamic Entropic Risk Measure

Entropic Risk Measure is a classical convex risk measure, which attracted a lot of attention in the risk measure literature [57, 141, 46, 103]. Let $\mathcal{X} = L^p$, for $p \in [1,\infty]$.

Definition 4.1.2 (Dynamic entropic risk measure)**.** A *Dynamic entropic risk measure* is a family $\rho^\gamma = \{\rho_t^\gamma\}_{t\in\mathbb{T}}$ of mappings $\rho_t^\gamma : \mathcal{X} \to \bar{L}_t^0$, indexed by $\gamma \in \mathbb{R}$, and defined by

$$\rho_t^\gamma(X) = \begin{cases} -\frac{1}{\gamma}\ln E[\exp(\gamma X)|\mathcal{F}_t] & \text{if } \gamma \neq 0, \\ -E[X|\mathcal{F}_t] & \text{if } \gamma = 0. \end{cases} \tag{4.5}$$

It is straightforward to check that for any $\gamma \in \mathbb{R}$, the map ρ^γ is dRM [103]. Moreover, if $\gamma \leq 0$, then ρ^γ is convex (dCV). As we will be working in the concave framework (see Remark 2.2.7) we will use $\varphi^\gamma = \{\varphi_t^\gamma\}_{t\in\mathbb{T}}$ to denote the negative of dynamic entropic risk measure, i.e.

$$\varphi_t^\gamma(X) = -\rho_t^\gamma(X).$$

We will refer to φ^γ as *Dynamic entropic utility measure*. Let us now recall some basic facts about these maps.

Proposition 4.1.3. *Let $\mathcal{X} = L^1$ and let φ^γ denote a dynamic entropic utility measure. Then*

1) $\{\varphi_t^\gamma\}_{t\in\mathbb{T}}$ is concave (dCC) *if $\gamma \leq 0$ and convex* (dCV) *if $\gamma \geq 0$.*

2) $\{\varphi_t^\gamma\}_{t\in\mathbb{T}}$ is dCE.

3) $\{\varphi_t^\gamma\}_{t\in\mathbb{T}}$ is strongly time consistent.

4) $\{\varphi_t^\gamma\}_{t\in\mathbb{T}}$ is increasing with γ.

5) $\{\varphi_t^\gamma\}_{t\in\mathbb{T}}$ is supermartingale time consistent in L^1 if and only if $\gamma \geq 0$.

[2]i.e. $\rho(X+Y) = \rho(X) + \rho(Y)$ for comonotone X, Y, that is random variables for which $(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0$, $\mathbb{P}(\mathrm{d}\omega) \otimes \mathbb{P}(\mathrm{d}\omega')$-a.s.

6) $\{\varphi_t^\gamma\}_{t\in\mathbb{T}}$ *is submartingale time consistent in* L^1 *if and only if* $\gamma \leq 0$.

For the proof of 1) and 2), see e.g. [103]; the proof in [103] is given for the case of L^∞, but can be adapted to the case of L^1. Moreover, the function defining φ^γ as (dCE) is given by

$$U^\gamma(x) := \begin{cases} \frac{1-\exp(\gamma x)}{1-\exp(\gamma)} & \text{if } \gamma \neq 0, \\ x & \text{if } \gamma = 0. \end{cases}$$

For the proof of 3), we first need to recall that the dynamic entropic utility measure is upper semi-continuous (USC) in L^1 (cf. [13, 37]), and then refer to [17]. For the proof of 4), we need to recall that the robust representation of dynamic entropic risk measures holds in the L^1 framework [37, 3], and then refer to [103]. Properties 5) and 6) follow directly from property 4), combined with dynamic programming reformulation of property 3); see [2] and [57, Proposition 6], where the proofs are done for the case of L^∞, but can be adapted to the case of L^1.

Remark 4.1.4. For $\mathcal{X} = L^p$ ($p \in [1,\infty]$) and infinite time horizon, one could show that the entropic dRMs are the only dRMs, which admit strong time consistency. It follows from the fact, that the class of entropic dRMs coincides with the class of dCEs which additionally admit counter cash-additivity (dCCA). See [103] for details.

Proposition 4.1.5. *Let* $\mathcal{X} = L^1$[3] *and let* φ^γ *denote dynamic entropic utility measure. For any* $\gamma < 0$*, the map* φ^γ *is representable and admits representation*

$$\varphi_t^\gamma(X) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P})} E_{\mathbb{Q}}[X|\mathcal{F}_t] - \frac{1}{\gamma}H_t(\mathbb{Q}|\mathbb{P}), \tag{4.6}$$

where $H_t(\mathbb{Q}|\mathbb{P})$ *corresponds to the conditional relative entropy of* $\mathbb{Q}$ *w.r.t.* $\mathbb{P}$*, i.e.*

$$H_t(\mathbb{Q}|\mathbb{P}) = \begin{cases} E\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\ln\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t\right] & \textit{if } \mathbb{Q} \ll \mathbb{P}, \\ +\infty & \textit{otherwise.} \end{cases} \tag{4.7}$$

The proof of Proposition 4.1.5 is a direct implication of *the variational principle*. See e.g. [3, Example 2.5] or [51, 77, 57] for more detailed proofs.

Remark 4.1.6. The (conditional) relative entropy introduced in (4.7), also called *Kullback-Leibler divergence*, might be considered as a (non-symmetric) measure of distance between two probability measures $\mathbb{Q}$ and $\mathbb{P}$, which describes the information lost when $\mathbb{Q}$ is used to approximate $\mathbb{P}$. It has strong connections to physics, as entropy can be described in terms of the *Energy dispersal*. See [51] for the connections to stochastic dynamic games and [46] for a general comment about applications to finance and economy.

Proposition 4.1.7. *Let* $\mathcal{X} = L^1$ *and let* φ^γ *denote dynamic entropic utility measure. Let* $\gamma < 0$*. For any* $t \in \mathbb{T}$ *and* $X \in L^1$*, such that* $\mu_0^\gamma(X) \in \mathbb{R}$ *and* $\mu_0^\gamma(X\ln X) \in \mathbb{R}$*, the essential supremum in* (4.6) *is attained for measure* $\mathbb{Q}_X$*, such that*

$$\frac{d\mathbb{Q}_X}{d\mathbb{P}} = \frac{e^{\gamma X}}{E[e^{\gamma X}|\mathcal{F}_t]} = e^{\gamma[X-\varphi_t^\gamma(X)]}. \tag{4.8}$$

[3] One could also consider the space $\{X \in L^0 \mid X\ln X \in L^1\}$. See [51] for details.

The proof of Proposition 4.1.7 is a direct extension of the proof from [51, Proposition 2.3], which was done for the static case. See also [57, Section 4], for the conditional case in L^∞ framework.

Remark 4.1.8. The transformation $\mathbb{P} \to \mathbb{Q}_X$ introduced in (4.8) is usually called Esscher transformation (see e.g. Gerber [85]). See also [77, Section 3.2] for more details about exponential utility and relative entropy and [46] for a general comment about problems, which involve minimisation of the relative entropy, using Esscher transformation.

One could generalise the Dynamic Entropic Risk Measure introduced in (4.5) by making the risk aversion parameter a non-constant adapted process, i.e. we could consider the dynamic risk measure given by

$$\varphi_t^{\gamma_t}(X) = \begin{cases} \frac{1}{\gamma_t} \ln E[\exp(\gamma_t X)|\mathcal{F}_t] & \text{if } \gamma_t \neq 0, \\ E[X|\mathcal{F}_t] & \text{if } \gamma_t = 0. \end{cases} \tag{4.9}$$

where $\{\gamma_t\}_{t\in\mathbb{T}}$ is such that $\gamma_t \in L_t^\infty$ and $X \in \mathcal{X} = L^\infty$, $t \in \mathbb{T}$. Noting that the map introduced in (4.5) is increasing with γ, it could be easily shown (see [2] for the idea of the proof) that $\{\varphi_t^{\gamma_t}\}_{t\in\mathbb{T}}$ is strongly time consistent, if and only if $\{\gamma_t\}_{t\in\mathbb{T}}$ is a constant process, middle acceptance time consistent if and only if $\{\gamma_t\}_{t\in\mathbb{T}}$ is a non-increasing process (i.e. $\gamma_{t+1} \leq \gamma_t$ for $t \in \mathbb{T}$, $t < T$) and middle rejection time consistent if and only if $\{\gamma_t\}_{t\in\mathbb{T}}$ is non-decreasing.

4.1.2 Dynamic TV@R

Let $\mathcal{X} = L^0$. The static Tail Value at Risk is a classic example of a coherent RM. In the literature, sometimes other names are used for this class of maps, such as *Tail Value at Risk*, *Average Value at Risk* or *Expected Shortfall*. While the definitions coincide for random variables with continuous distribution, they slightly differ in the general case [77].

Tail Value at Risk could be regarded as the modification of Value at Risk, when we consider the conditional expectation, instead of a simple quantile (see [93] for details).

Definition 4.1.9. A *Tail Value at Risk* (TV@R) is a map $\rho^\alpha : \mathcal{X} \to \bar{\mathbb{R}}$, indexed by $\alpha \in (0,1]$, and defined by

$$\rho^\alpha(X) = -\inf_{\mathbb{Q}\in\mathcal{D}^\alpha} E_{\mathbb{Q}}[X]. \tag{4.10}$$

where $\mathcal{D}^\alpha := \{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) : \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \alpha^{-1}\}$.

From the definition, we get that TV@R is a representable coherent RM for any $\alpha \in (0,1]$. Moreover, if X has the continuous distribution, then we get

$$\rho^\alpha(X) := -E[X|X \leq q_\alpha(X)], \tag{4.11}$$

where $\alpha \in (0,1]$ and $q_\alpha(X)$ denotes the α-quantile of X. This representation explains the name Tail Value at Risk. See [41] for a through discussion about properties of (static) TV@R.

The dynamic version of TV@R could be obtained, modifying the set $\mathcal{D}^\alpha$. For $\alpha \in (0,1]$ let $\{\mathcal{D}_t^\alpha\}_{t\in\mathbb{T}}$ be defined by

$$\mathcal{D}_t^\alpha := \{\mathbb{Q} \in \mathcal{Q}_t^1 : \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \alpha^{-1}\}. \tag{4.12}$$

Definition 4.1.10. A *Dynamic Tail Value at Risk* (dTV@R) is a family $\{\rho_t^\alpha\}_{t\in\mathbb{T}}$ of mappings $\rho_t^\alpha : \mathcal{X} \to \bar{L}_t^0$, indexed by $\alpha \in (0,1)$, and defined by

$$\rho_t^\alpha(X) = -\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha} E_{\mathbb{Q}}[X|\mathcal{F}_t]. \tag{4.13}$$

It is straightforward to check that for any $\alpha \in (0,1)$, the map $\{\rho_t^\alpha\}_{t\in\mathbb{T}}$ is dRM [41]. Working in concave framework, throughout this Subsection we will use $\varphi^\alpha = \{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ to denote negative of $\{\rho_t^\alpha\}_{t\in\mathbb{T}}$, i.e.

$$\varphi_t^\alpha(X) = -\rho_t^\alpha(X). \tag{4.14}$$

Proposition 4.1.11. *Let $\mathcal{X} = L^0$ and let $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ denote negative of dTV@R. Then*

1) $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ is subadditive (dSPA) and positively homogeneous (dPH).[4]

2) $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ is increasing with α.

3) $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ is submartingale time consistent in L^0.[5]

4) $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ is not weakly acceptance time consistent in L^0.

The proof of 1) is straightforward (see e.g. [41]). Indeed for $\alpha \in (0,1]$, $t \in \mathbb{T}$, $\beta \geq 0$ ($\beta \in \bar{L}_t^0$) and $X, Y \in L^0$, using Proposition 2.1.3, we get

$$\begin{aligned}\varphi_t^\alpha(X+Y) &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha} E_{\mathbb{Q}}[X+Y|\mathcal{F}_t] \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha}[E_{\mathbb{Q}}[X|\mathcal{F}_t] + E_{\mathbb{Q}}[X|\mathcal{F}_t]] \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha} E_{\mathbb{Q}}[X|\mathcal{F}_t] + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha} E_{\mathbb{Q}}[Y|\mathcal{F}_t] = \varphi_t^\alpha(X) + \varphi_t^\alpha(Y),\end{aligned}$$

and

$$\varphi_t^\alpha(\beta X) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha} E_{\mathbb{Q}}[\beta X|\mathcal{F}_t] = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha} \beta E_{\mathbb{Q}}[X|\mathcal{F}_t] = \beta\varphi_t^\alpha(X).$$

Next, 2) is a simple implication of the fact, that for $\alpha_1 > \alpha_2$ and $t \in \mathbb{T}$, we get $\mathcal{D}_t^{\alpha_1} \subseteq \mathcal{D}_t^{\alpha_2}$.

To prove 3), it enough to note that for $t, s \in \mathbb{T}$, such that $s > t$, we get $\mathcal{D}_s^\alpha \subseteq \mathcal{D}_t^\alpha$. Because of that we get[6]

$$\varphi_t^\alpha(X) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^\alpha} E_{\mathbb{Q}}[X|\mathcal{F}_t] \leq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_s^\alpha} E_{\mathbb{Q}}[X|\mathcal{F}_t] \leq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_s^\alpha} E[E_{\mathbb{Q}}[X|\mathcal{F}_s]|\mathcal{F}_t]. \tag{4.15}$$

Now, using the fact that $\mathcal{D}_s^\alpha$ is L^1-closed (see [41] for details), for any $X \in L^0$, there exist $\mathbb{Q}_X \in \mathcal{D}_s^\alpha$ such that $\varphi_s^\alpha(X) = E_{\mathbb{Q}_X}[X|\mathcal{F}_s]$. This implies

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_s^\alpha} E[E_{\mathbb{Q}}[X|\mathcal{F}_s]|\mathcal{F}_t] = E[E_{\mathbb{Q}_X}[X|\mathcal{F}_s]|\mathcal{F}_t] = E[\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_s^\alpha} E_{\mathbb{Q}}[X|\mathcal{F}_s]|\mathcal{F}_t] = E[\varphi_s^\alpha(X)|\mathcal{F}_t]. \tag{4.16}$$

Combining (4.15) and (4.16) we obtain submartingale time-consistency.

For the proof of 4) it is enough to consider the counterexample in L^∞ framework (taken from [7]). Take a 2-step discrete dynamics with 3 paths in each step and consider

$$\Omega = \{[uu], [um], [ud], [du], [dm], [dd]\}$$

where $\mathbb{P}$ is uniform on Ω and $\mathcal{F}_1$ is generated by $[u\cdot]$ and $[d\cdot]$. Take

$$X([uu]) = -10,\ X([um]) = 12,\ X([ud]) = 14,\ X([du]) = -20,\ X([dm]) = 22, X([dd]) = 22.$$

[4] i.e. ρ^α is a coherent dRM.

[5] Note, that this implies that negative of dynamic TV@R is weakly rejection time consistent.

[6] Note that for any $Z := \frac{d\mathbb{Q}}{d\mathbb{P}}$ we get $E[ZX|\mathcal{F}_t] \leq E[E[ZX|\mathcal{F}_s]|\mathcal{F}_t]$.

Then, for $\alpha = \frac{2}{3}$ we get

$$\varphi_0^\alpha(X) = \frac{-20-10+12+14}{4} = -1$$

$$\varphi_1^\alpha(X)([u\cdot]) = \frac{-10+12}{2} = 1, \quad \varphi_1^\alpha(X)([d\cdot]) = \frac{-20+22}{2} = 1,$$

which implies $-1 = \varphi_0^\alpha(X) \not\geq \operatorname{ess\,inf}_0(\varphi_1^\alpha(X)) = 1$. It is also worth mentioning that $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ is not middle rejection time consistent.[7]

The lack of weak acceptance time consistency (and strong time consistency) is one of the main drawbacks of using dTV@R. Nevertheless, using so called recursive construction (see [2] for details) for a finite time horizon $\mathbb{T} = \{0, 1, \dots, T\}$ ($T \in \mathbb{N}$) and L^∞ one could define a new (coherent) dRM, which will be strongly time consistent. For a fixed $\alpha \in (0, 1)$ it will take the form:

$$\widetilde{\rho}_t^\alpha(X) = -\operatorname*{ess\,inf}_{Q\in\widetilde{\mathcal{D}}_t^\alpha} E_{\mathbb{Q}}[X|\mathcal{F}_t], \tag{4.17}$$

for

$$\widetilde{\mathcal{D}}_t^\alpha := \Big\{\mathbb{Q} \in \mathcal{Q}_t^1 \Big| \ \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \prod_{s=t}^{T-1} \frac{\mathrm{d}\mathbb{Q}_s}{\mathrm{d}\mathbb{P}}, \text{ where } \mathbb{Q}_s \in \mathcal{D}_s^\alpha, \ \frac{\mathrm{d}\mathbb{Q}_s}{\mathrm{d}\mathbb{P}} \in L_{s+1}^1\Big\}$$

$$= \Big\{\mathbb{Q} \in \mathcal{Q}_t^1 \Big| \ \frac{Z_{s+1}^Q}{Z_s^Q} \leq \alpha, \text{ for } s = t, \dots, T-1, \text{ where } Z_s^Q = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\big|_{\mathcal{F}_t}\Big\}$$

See [40, Example 2.3.1] or [2, Example 36] for details.

Moreover, one could generalise the family of dTV@R maps, allowing risk-averse parameter to be non constant, i.e. we can consider a process $\{a_t\}_{t\in\mathbb{T}}$, where $a_t \in L_t^0$ and $0 < a_t < 1$, instead of $a \in (0, 1)$ in (4.13) or (4.17) (again see [2, Example 36] for details).

4.1.3 Dynamic WV@R

In this subsection let us give some comment about a special family of coherent dRMs, namely Dynamic Weighted Value at Risk, which includes also the family of TV@Rs. The class of dWV@Rs appears to be very convenient and analytically traceable (where portfolio optimisation problems are considered, the reason we introduce those mappings (see e.g. [44, 43] for details).

Let us start, by recalling the definition of the (static) Weighted Value at Risk RM for $\mathcal{X} = L^0$

Definition 4.1.12. We say that $\rho^\nu : \mathcal{X} \to \bar{\mathbb{R}}$ is a *Weighted Value at Risk*, if

$$\rho^\nu(X) := \int_0^1 \rho^\alpha(X)\nu(\mathrm{d}\alpha), \tag{4.18}$$

where ν is a probability measure on $(0, 1]$ and $\{\rho^\alpha\}_{\alpha\in(0,1]}$ is a family of TV@Rs.

To omit various technical problems, we will generalise (4.18) to the dynamic case using robust representation. One could show (see e.g. [41]), that the map defined in (4.18) could be rewritten as

$$\rho^\nu(X) = -\inf_{\mathbb{Q}\in\mathcal{D}^\nu} E_{\mathbb{Q}}[X]$$

[7] Take a 2-step discrete dynamics with 3 paths in each step and consider $\mathcal{F}_1 = \sigma(\{1,2,3\},\{4,5,6\},\{7,8,9\})$, $\mathcal{F} = \mathcal{F}_2 = \sigma(\{1\},\dots,\{9\})$, $X(\omega) = -\omega$, for $\omega \in \{1,2,\dots,9\}$. Then $-8 = \varphi_0^{1/3}(X) \geq \varphi_0^{1/3}(\varphi_1^{1/3}(X)) = -9$.

for

$$\mathcal{D}^{\nu} := \{\mathbb{Q} \ll \mathbb{P} \mid E[(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} - x)^{+}] \leq \Phi_{\nu}(x),\ \forall x \in \mathbb{R}_{+}\}, \tag{4.19}$$

where

$$\Phi_{\nu}(x) := \sup_{y\in[0,1]} \Big[\int_0^y \int_{[z,1]} \lambda^{-1}\nu(\mathrm{d}\lambda)\,\mathrm{d}z - xy \Big].$$

We are now ready to provide the definition of *Dynamic Weighted Value at Risk*.

Definition 4.1.13. We will call $\{\rho_t^{\nu}\}_{t\in\mathbb{T}}$ a *Dynamic Weighted Value at Risk* if

$$\rho_t^{\nu}(X) = -\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t^{\nu}} E_{\mathbb{Q}}[X|\mathcal{F}_t], \tag{4.20}$$

where ν is a probability measure on $(0,1]$ and

$$\mathcal{D}_t^{\nu} := \{\mathbb{Q} \in \mathcal{Q}_t^1 \mid E[(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} - x)^{+}|\mathcal{F}_t] \leq \Phi_{\nu}(x),\ \forall x \in \mathbb{R}_{+}\}. \tag{4.21}$$

As always we will use notation $\{\varphi_t^{\nu}\}_{t\in\mathbb{T}}$ to denote the negative of a dynamic Weighted Value at Risk. We also know that the static TV@R is law invariant (LI), so there exists a functional $\widetilde{\varphi}^{\nu}$, defined on distributions, such that $\varphi^{\nu}(X) = \widetilde{\varphi}^{\nu}(Law(X))$. Let us now recall some basic properties of this class of maps.

Proposition 4.1.14. *Let ν be a probability measure on $[0,1)$. Then*

1. *For any $X \in L^0$ and $t \in \mathbb{T}$ we get $\varphi_t^{\nu}(X) = \widetilde{\varphi}^{\nu}(Law(X|\mathcal{F}_t))$.*
2. *$\{\varphi_t^{\nu}\}_{t\in\mathbb{T}}$ is weakly rejection time consistent in L^0.*
3. *$\{\varphi_t^{\nu}\}_{t\in\mathbb{T}}$ is not weakly acceptance time consistent in L^0.*

For the proof of 1), see [43, Lemma 2.2]. The proof of 2) is a straightforward implication of the fact that for $s > t$, we get $\mathcal{D}_s^{\nu} \subseteq \mathcal{D}_t^{\nu}$ and the fact that $\{\varphi_t^{\nu}\}_{t\in\mathbb{T}}$ is local, i.e. $\mathbb{Q}$ could be defined locally in (4.20)*. The counterexample for 3) could be constructed using the idea from [7, Section 5.2].*

4.2 Dynamic acceptability indices

The family of Acceptability Indices was introduced in [45], and studied (also for the dynamic case) e.g. in [126, 15, 20]. Let us introduce a family of *regular* acceptability indices, which will allow us to show the tight connection between coherent RMs and acceptability indices.

Definition 4.2.1 (Regular acceptability index)**.** Let α be an acceptability index.[8] We will say that α is *regular* if it satisfies:

1) *Nonnegativity*, i.e. $\alpha(X) \geq 0$ for all $X \in \mathcal{X}$;

2) *Non-degeneracy*, i.e.

[8] i.e. α is adapted (A), translation invariant (TI), monotone increasing (MI), scale invariant (SI) and quasi-concave (QCC). See page 13.

(a) $\alpha(X) = 0$, for some $X \in \mathcal{X}$;
(b) $\alpha(X) = \infty$, for some $X \in \mathcal{X}$;

Similarly, if $\alpha = \{\alpha_t\}_{t\in\mathbb{T}}$ is a dynamic acceptability index, then we will say, that α is *regular* if for any $t \in \mathbb{T}$, the map α_t satisfies 1) and 2).

Let us recall now the duality theorem from [45] for random variables on L^∞.

Theorem 4.2.2. *Let $\mathcal{X} = L^\infty$. A map $\alpha : \mathcal{X} \to [0,\infty]$ is a regular acceptability index satisfying Fatou property (FP) if and only if there exists a family of subsets $\{\mathcal{D}_x\}_{x\in\mathbb{R}_+}$ of $\mathcal{M}_1(\mathbb{P})$ such that $\mathcal{D}_x \subseteq \mathcal{D}_y$ for $x \leq y$ and*[9]

$$\alpha(X) = \sup\Big\{x \in \mathbb{R}_+ \mid \inf_{\mathbb{Q}\in\mathcal{D}_x} E_{\mathbb{Q}}[X] \geq 0\Big\}. \tag{4.22}$$

For the proof of Theorem 4.2.2 see [45, Theorem 1].

One could easily see that for each $x \in \mathbb{R}^+$ in Theorem 4.2.2, the map $\varphi^x(X) := \inf_{\mathbb{Q}\in\mathcal{D}_x} E_{\mathbb{Q}}[X]$ corresponds to negative of a representable coherent RM. Moreover, as for $x \leq y$, we get $\mathcal{D}_x \subseteq \mathcal{D}_y$, we know that the family $\{\varphi^x\}_{x\in\mathbb{R}_+}$ should be decreasing, i.e. $\varphi^x(X) \geq \varphi^y(X)$ for any $X \in \mathcal{X}$ and $x \leq y$. Theorem 4.2.2 could be generalized to conditional case [15] as well as to the space of stochastic processes [20]. We don't present the results here, as they require many technical assumptions and are not the main topic of this thesis. Let us alone mention that for any decreasing[10] family of dUMs $\{\varphi^x_t\}_{t\in\mathbb{T}}$ (indexed by $x \in \mathbb{R}_+$; typically coherent dRMs) satisfying certain technical properties and

$$\mathbb{X} := \{(x_n)_{n\in\mathbb{N}} \mid x_n \in \mathbb{R}_+,\ x_0 = 0,\ x_{n+1} > x_n\},$$

the dynamic map $\alpha = \{\alpha_t\}_{t\in\mathbb{T}}$, where $\alpha_t \colon \mathcal{X} \to \bar{L}^0_t$ given by

$$\alpha_t(X) = \operatorname*{ess\,sup}_{(x_n)\in\mathbb{X}}\{\sum_{n=0}^{\infty} \mathbb{1}_{\{\varphi^{x_n}_t(X)\geq 0\}}(x_{n+1} - x_n)\} \qquad (X \in \mathcal{X}), \tag{4.23}$$

is a dynamic acceptability index[11] (see [15, 20] for details). With slight abuse of notation, we could also write that α is such that

$$\alpha_t(X) = \sup\{x \in \mathbb{R}_+ : \varphi^x_t(X) \geq 0\} \qquad (X \in \mathcal{X}), \tag{4.24}$$

understanding (4.24) ω-by-ω, and taking it's $\mathcal{F}_t$-measurable version (see [20] for details). Conversly, if α is a dynamic acceptability index, then we could recover the decreasing family of dUMs defining for each $x \in \mathbb{R}_+$ the map $\varphi_t \colon \mathcal{X} \to \bar{L}^0_t$ by

$$\varphi^x_t(X) = \inf\{c \in \mathbb{R} : \alpha_t(X - c1_{\{t\}}) \leq x\} \qquad (X \in \mathcal{X}), \tag{4.25}$$

where (4.25) is understood in the same way as (4.24) (again, see [20] for details).

Now, let us show, how transformations (4.24) and (4.25) preserve time consistency. The next two results will give an important (dual) connection between cash additive measures and scale invariant measures.

[9] If (for a given $x \in \mathbb{R}_+$) the set $\mathcal{D}_x$ is empty, we will write $\inf_{\mathbb{Q}\in\mathcal{D}_x} E_{\mathbb{Q}}[X] = \infty$. On the other hand if (for a given $X \in \mathcal{X}$), there exists no $x \in \mathbb{R}_+$, for which the inequality in (4.22) is attained, we will write $\alpha(X) = 0$. In other words, we use convention $\inf\emptyset = \infty$ and $\sup\emptyset = 0$.

[10] i.e. $\varphi^x_t(X) \leq \varphi^y_t(X)$ for all $X \in \mathcal{X}$, $t \in \mathbb{T}$ and $x, y \in \mathbb{R}_+$, such that $x \leq y$.

[11] Please note that for any $X \in \mathcal{X}$ and $x \in \mathbb{R}_+$, we get $\{\varphi^x_t(X) \geq 0\} \in \mathcal{F}_t$ so $\alpha_t(X) \in \bar{L}^0_t$. Moreover, $\alpha_t(X) \geq 0$ for any $X \in \mathcal{X}$, due to the convention $\inf\emptyset = \infty$ and $\sup\emptyset = 0$, similar to the one in (4.22).

Proposition 4.2.3. *For $x \in \mathbb{R}_+$, let $\{\varphi_t^x\}_{t\in\mathbb{T}}$ be a decreasing family of dUMs. Moreover, let us assume that for each $x \in \mathbb{R}_+$, $\{\varphi_t^x\}_{t\in\mathbb{T}}$ is weakly acceptance (resp. weakly rejection) time consistent. Then the family $\{\alpha_t\}_{t\in\mathbb{T}}$ of maps $\alpha_t : \mathcal{X} \to \bar{L}_t^0$ defined in (4.24) is a semi-weakly acceptance (resp. semi-weakly rejection) time consistent dUM.*

Proof. The proof of locality and monotonicity of (4.24) is straightforward (see [20] for details). Let us assume that $\{\varphi_t^x\}_{t\in\mathbb{T}}$ is weakly acceptance time consistent. Using Proposition 3.1.8 we get

$$\begin{aligned}
1_{\{V_t\geq 0\}}\alpha_t(V) &= 1_{\{V_t\geq 0\}}\Big(\sup\{x\in\mathbb{R}_+ : 1_{\{V_t\geq 0\}}\varphi_t^x(V)\geq 0\}\Big)\\
&\geq 1_{\{V_t\geq 0\}}\Big(\sup\{x\in\mathbb{R}_+ : 1_{\{V_t\geq 0\}}[\operatorname{ess\,inf}_t \varphi_{t+1}^x(V)+V_t]\geq 0\}\Big)\\
&\geq 1_{\{V_t\geq 0\}}\Big(\sup\{x\in\mathbb{R}_+ : 1_{\{V_t\geq 0\}}\operatorname{ess\,inf}_t \varphi_{t+1}^x(V)\geq 0\}\Big)\\
&= 1_{\{V_t\geq 0\}}\operatorname{ess\,inf}_t\Big(\sup\{x\in\mathbb{R}_+ : 1_{\{V_t\geq 0\}}\varphi_{t+1}^x(V)\geq 0\}\Big)\\
&= 1_{\{V_t\geq 0\}}\operatorname{ess\,inf}_t \alpha_{t+1}(V)
\end{aligned}$$

This leads to inequality

$$\alpha_t(V)\geq 1_{\{V_t\geq 0\}}\operatorname{ess\,inf}_t \alpha_{t+1}(V)+1_{\{V_t<0\}}(-\infty),$$

which, by Proposition 3.1.8, is equivalent to semi-weak rejection time consistency. The proof of weak acceptance time consistency is similar. ☐

Proposition 4.2.4. *Let $\{\alpha_t\}_{t\in\mathbb{T}}$ be a dUM, which is independent of the past and translation invariant. Moreover, let us assume that $\{\alpha_t\}_{t\in\mathbb{T}}$ is semi-weakly acceptance (resp. semi-weakly rejection) time consistent. Then for any $x\in\mathbb{R}_+$ the family $\{\varphi_t^x\}_{t\in\mathbb{T}}$ defined in (4.25) is a weakly rejection (resp. weakly acceptance) time consistent dUM.*

Proof. The proof of locality and monotonicity of (4.25) is straightforward (see [20] for details). Let us prove weak acceptance time consistency. Let us assume that $\{\alpha_t\}_{t\in\mathbb{T}}$ is semi-weakly acceptance time consistent. Using Proposition 3.1.8 we get

$$\begin{aligned}
\varphi_t^x(V) &= \inf\{c\in\mathbb{R} : \alpha_t(V-c1_{\{t\}})\leq x\}\\
&= \inf\{c\in\mathbb{R} : \alpha_t(V-c1_{\{t+1\}})\leq x\}\\
&= \inf\{c\in\mathbb{R} : \alpha_t(V-c1_{\{t+1\}}-V_t1_{\{t\}})\leq x\}+V_t\\
&\geq \inf\{c\in\mathbb{R} : 1_{\{0\geq 0\}}\operatorname{ess\,inf}_t \alpha_{t+1}(V-c1_{\{t+1\}}-V_t1_{\{t\}})+1_{\{0<0\}}(-\infty)\leq x\}+V_t\\
&= \inf\{c\in\mathbb{R} : \operatorname{ess\,inf}_t \alpha_{t+1}(V-c1_{\{t+1\}})\leq x\}+V_t\\
&= \operatorname{ess\,inf}_t\big(\inf\{c\in\mathbb{R} : \alpha_{t+1}(V-c1_{\{t+1\}})\leq x\}\big)+V_t\\
&= \operatorname{ess\,inf}_t \varphi_{t+1}^x(V)+V_t
\end{aligned}$$

Which, using Proposition 3.1.8, is equivalent to weak acceptance time consistency. The proof of rejection time consistency is similar. ☐

This type of dual representation, i.e. (4.24)–(4.25), first appeared in [45] where the authors studied static (one period of time) scale invariant measures. Subsequently, in [20], the authors extended these results to the case of stochastic processes with special emphasis on time consistency property. In contrast to [20], we consider an arbitrary probability space, not just a finite one.

4.2.1 Dynamic TV@R Acceptability Index for Processes

Tail Value at Risk Acceptability Index was introduced in [45], as a scale invariance measure of performance for the case of random variables. Using [20], we extend this notion to the case of stochastic processes. Let $\mathcal{X} = \mathbb{V}^0$, and for a fixed $\alpha \in (0,1]$ we consider the sets $\{\mathcal{D}_t^\alpha\}_{t\in\mathbb{T}}$ defined as in (4.12).

Definition 4.2.5. A *Dynamic Tail Value at Risk Acceptability Index* (Dynamic TV@R Acceptability Index) is a family $\{\alpha_t\}_{t\in\mathbb{T}}$ of mappings $\alpha_t : \mathcal{X} \to \bar{L}_t^0$, given by

$$\alpha_t(V) = \sup\{x \in \mathbb{R}_+ : \rho_t^x(V) \geq 0\}, \tag{4.26}$$

where for $x \in \mathbb{R}_+$, we define $\rho^x = \{\rho_t^x\}_{t\in\mathbb{T}}$ as

$$\rho_t^x(V) = \operatorname*{ess\,inf}_{Z\in D_t^{g(x)}} E[Z\sum_{i=t}^{T} V_i|\mathcal{F}_t], \quad V \in \mathbb{V},\ t \in \mathcal{T}, \tag{4.27}$$

for (a distortion function) $g(x) = \frac{1}{1+x}$, $x \in \mathbb{R}^+$.

Proposition 4.2.6. *Let $\mathcal{X} = \mathbb{V}^0$ and let $\{\alpha_t\}_{t\in\mathbb{T}}$ denote a Dynamic* TV@R *Accep. Index. Then*

1. *$\{\alpha_t\}_{t\in\mathbb{T}}$ is a regular dynamic acceptability index.*
2. *$\{\alpha_t\}_{t\in\mathbb{T}}$ is semi-weakly rejection time consistent.*

For the proof of 1) one could notice that it is easy to show that ρ^x is an increasing (with respect to x) family of (negatives of) dynamic coherent risk measures for processes (see [45] and [20] for details). Hence, the map $\{\alpha_t\}_{t\in\mathbb{T}}$ given by (4.26) is an acceptability index for processes (again, see [45] and [20]). For the proof of 2), we can use similar arguments as in Section 4.1.2, to conclude that ρ^x is weakly rejection time consistent, for any fixed $x \in \mathbb{R}_+$. Hence, by Proposition 4.2.3 we obtain that α is semi-weakly acceptance time consistent.

On the other hand $\{\alpha_t\}_{t\in\mathbb{T}}$ is not semi-weakly acceptance time consistent. Indeed, following similar reasoning as in the proof of duality from [20] and using Proposition 4.2.4, we get that if α is semi-weakly acceptance time consistent, then $\{\rho_t^x\}_{t\in\mathbb{T}}$ is weakly acceptance time consistency, for any $x \in \mathbb{R}_+$. This leads to a contradiction, since these maps are not weakly acceptance time consistent, as stated in Example 4.1.2.

4.2.2 Dynamic RAROC for processes

Risk Adjusted Return On Capital (RAROC) is a popular measure of scale invariant measure of performance equal (see for instance [45] for static RAROC, and [20] for its extension to dynamic setup). We consider the space $\mathcal{X} = \mathbb{V}^1$, we fix $\alpha \in (0,1)$ and set $\mathbb{T} = \{0,1,\ldots,T\}$.

Definition 4.2.7. A *Dynamic Coherent Risk-Adjusted Return on Capital* (Dynamic RAROC) is a family $\{\alpha_t\}_{t\in\mathbb{T}}$ of mappings $\alpha_t : \mathcal{X} \to \bar{L}_t^0$, given by

$$\alpha_t(V) := \begin{cases} \frac{E[\sum_{i=t}^T V_i|\mathcal{F}_t]}{\rho_t^\alpha(V)} & \text{if } E[\sum_{i=t}^T V_i|\mathcal{F}_t] > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.28}$$

where $\rho_t(V) = -\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}_t} E_{\mathbb{Q}}[(\sum_{i=t}^T V_i)|\mathcal{F}_t]$ for some family $\{\mathcal{D}_t\}_{t\in\mathbb{T}}$, such that $\mathcal{D}_t \subseteq \mathcal{Q}_t^1$. We use the convention $\alpha_t(V) = +\infty$, if $\rho_t(V) \leq 0$.

Proposition 4.2.8. *Let $\mathcal{X} = \mathbb{V}^1$ and let $\{\alpha_t\}_{t\in\mathbb{T}}$ denote a Dynamic RAROC, where the corresponding family $\{\mathcal{D}_t\}_{t\in\mathbb{T}}$ is such that $\mathcal{D}_s \subseteq \mathcal{D}_t$ for $s > t$. Then*

1) $\{\alpha_t\}_{t\in\mathbb{T}}$ is a regular dynamic acceptability index.

2) $\{\alpha_t\}_{t\in\mathbb{T}}$ admits representation $\alpha_t(V) = \sup\{x \in \mathbb{R}_+ : \varphi_t^x(V) \geq 0\}$, where $\{\varphi_t^x\}_{t\in\mathbb{T}}$ is given by

$$\varphi_t^x(V) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{B}_t^x} E_{\mathbb{Q}}[(\sum_{i=t}^{T} V_i)|\mathcal{F}_t],$$

for family $\mathcal{B}_t^x = \{\mathbb{Q} \in \mathcal{Q}_t^1 : \mathbb{Q} = \frac{1}{1+x}\mathbb{P} + \frac{x}{1+x}\mathbb{Q}_1,$ for some $\mathbb{Q}_1 \in \mathcal{D}_t\}$.

3) $\{\alpha_t\}_{t\in\mathbb{T}}$ is semi-weakly acceptance time consistent.

The proof of 1) and 2) for the static case could be found in [45, Section 3.4] and could be easily converted to the dynamic case (see [20, 15] for the idea of the proof). The proof of 3) follows from Proposition 4.2.3. It is enough to note that for any $x \in \mathbb{R}_+$, the map $\{\phi_t^x\}_{t\in\mathbb{T}}$ is a weakly acceptance time consistent dUM.[12] It is worth mentioning that $\{\alpha_t\}_{t\in\mathbb{T}}$ might be not semi-weakly rejection time consistent, see [20, Example 6.5] for a simple counterexample.

4.2.3 Dynamic GLR for processess

Dynamic Gain Loss Ratio (dynamic GLR) is another popular measure of performance, which essentially overcomes the deficiencies of Sharpe Ratio by penalizing for positive returns, and is equal to the ratio of expected return over expected losses. For various properties and dual representations of dynamic GLR see for instance [20, 17]. Let $\mathbb{T} = \{0, 1, \dots, T\}$ and $\mathcal{X} = \mathbb{V}^1$.

Definition 4.2.9. A *Dynamic Gain Loss Ratio* (Dynamic GLR) is a family $\{\alpha_t\}_{t\in\mathbb{T}}$ of mappings $\alpha_t : \mathcal{X} \to \bar{L}_t^0$, given by

$$\alpha_t(V) := \begin{cases} \frac{E[\sum_{i=t}^T V_i|\mathcal{F}_t]}{E[(\sum_{i=t}^T V_i)^-|\mathcal{F}_t]} & \text{if } E[\sum_{i=t}^T V_i|\mathcal{F}_t] > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.29}$$

Proposition 4.2.10. *Let $\mathcal{X} = \mathbb{V}^1$ and let $\{\alpha_t\}_{t\in\mathbb{T}}$ denote a Dynamic GLR. Then*

1) $\{\alpha_t\}_{t\in\mathbb{T}}$ is a regular dynamic acceptability index.

2) $\{\alpha_t\}_{t\in\mathbb{T}}$ is semi-weakly acceptance time consistent.

3) $\{\alpha_t\}_{t\in\mathbb{T}}$ is semi-weakly rejection time consistent.

The proof of 1) could be found in [20]. To prove 2), we notice that due to Proposition 3.1.8, we only need to prove that

$$\alpha_t(V) \geq \mathbb{1}_{\{V_t\geq 0\}} \operatorname{ess\,inf}_t(\alpha_{t+1}(V)) + \mathbb{1}_{\{V_t<0\}}(-\infty). \tag{4.30}$$

[12] See 3) from Proposition 4.1.11, for the idea of the proof of weak acceptance time consistency.

Next, on the set $\{V_t < 0\}$ the inequality (4.30) is trivial. Because α_t is non-negative and local, without loss of generality we could assume that $\operatorname{ess\,inf}_t(\alpha_{t+1}(V)) > 0$. Moreover,

$$\alpha_{t+1}(V) \geq \operatorname{ess\,inf}_t(\alpha_{t+1}(V)),$$

which implies

$$E[\sum_{i=t+1}^{T} V_i|\mathcal{F}_{t+1}] \geq \operatorname{ess\,inf}_t(\alpha_{t+1}(V)) \cdot E[(\sum_{i=t+1}^{T} V_i)^-|\mathcal{F}_{t+1}]. \tag{4.31}$$

Using (4.31) we obtain

$$\begin{aligned}
\mathbb{1}_{\{V_t\geq 0\}}E[\sum_{i=t}^{T} V_i|\mathcal{F}_t] &\geq \mathbb{1}_{\{V_t\geq 0\}}E[E[\sum_{i=t+1}^{T} V_i|\mathcal{F}_{t+1}]|\mathcal{F}_t] \\
&\geq \mathbb{1}_{\{V_t\geq 0\}} \operatorname{ess\,inf}_t(\alpha_{t+1}(V)) \cdot E[\mathbb{1}_{\{V_t\geq 0\}}E[(\sum_{i=t+1}^{T} V_i)^-|\mathcal{F}_{t+1}]|\mathcal{F}_t] \\
&\geq \mathbb{1}_{\{V_t\geq 0\}} \operatorname{ess\,inf}_t(\alpha_{t+1}(V)) \cdot E[(\sum_{i=t}^{T} V_i)^-|\mathcal{F}_t]. \qquad (4.32)
\end{aligned}$$

Combining (4.32) with the fact that from $\operatorname{ess\,inf}_t(\alpha_{t+1}(V)) > 0$, we get $\mathbb{1}_{\{V_t\geq 0\}}E[\sum_{i=t}^{T} V_t|\mathcal{F}_t] > 0$, we get 2). The proof of 3) will analogous. See also [20] for more detailed proof of 2) and 3).

4.3 Dynamic limit growth indices

This section is based on [18]. If not stated otherwise, in this subsection we will assume that $\mathcal{X} = \mathbb{V}^p_\tau$ for $p \in \{0\} \cup [1, \infty]$. One should look at $\mathcal{X}$ as the cumulative value process of some portfolio. Let us present a main object of study in this subsection, which we will name Dynamic Limit Growth Indices.

Definition 4.3.1. A *Dynamic Limit Growth Index* (dLGI) is a family $\{\varphi_t\}_{t\in\mathbb{T}}$ of maps $\varphi_t : \mathcal{X} \to \bar{L}^0_t$ such that

$$\varphi_t(V) = \liminf_{T\to\infty} \frac{\mu_t(\ln \frac{V_T}{V_t})}{T}, \tag{4.33}$$

where $\mu_t : \widehat{L}^0 \to \bar{L}^0_t$, and $\{\mu_t\}_{t\in\mathbb{T}}$ is local (dL) and monotone (dMI). Additionally, we will say that dLGI is *risk seeking*, if $\{\mu_t\}_{t\in\mathbb{T}}$ is such that $\mu_t(X) = \mu_t(X^+)$ for $t \in \mathbb{T}$ and $X \in \widehat{L}^0$.

We will often refer to $\{\mu_t\}_{t\in\mathbb{T}}$ as a family of mappings that defines dLGI. The maps introduced in Definition 4.33 have a natural financial interpretation. The cumulative log-return over the period (t, T) is a common way to measure the process growth. Because it is a random variable, we use a tUM, say μ_t, which represents our preferences (at time t). Finally we divide the outcome by T to normalize it in time. Taking the liminf as T goes to infinity allows us to measure the long-time efficiency of our value process. We use liminf because we want to measure the actual (worst case) efficiency of our portfolio. It also makes this measure more robust (at least to losses). Also, note that risk seeking dLGI ignores the losses in the sense that it substitutes all losses (negative log returns) by 0.

We want to use dLGI to assess performance of value processes: the greater the value of dLGI the better the performance of the portfolio. This is in line with the theory of performance measures. In particular, we are interested in identifying conditions under which dLGIs are quasi-concave dPMs. Towards this end, we provide Proposition 4.3.2 that give sufficient and necessary conditions for dLGI to be quasi-concave dPM.

Proposition 4.3.2. *Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be a dLGI defined in terms of $\{\mu_t\}_{t\in\mathbb{T}}$. Then, $\{\varphi_t\}_{t\in\mathbb{T}}$ is quasi-concave dPM if and only if for any $t\in\mathbb{T}$, and any $V\in\mathcal{X}$,*

$$\liminf_{T\to\infty}\frac{\mu_t(\ln\frac{V_T}{V_t})}{T}=\liminf_{T\to\infty}\frac{\mu_t(\ln V_T)}{T}. \tag{4.34}$$

Proof. Let $\{\varphi_t\}_{t\in\mathbb{T}}$ be dLGI generated by $\{\mu_t\}_{t\in\mathbb{T}}$. We know that $\{\mu_t\}_{t\in\mathbb{T}}$ is (dL) and (dMI).
($\Leftarrow$) Let $\{\mu_t\}_{t\in\mathbb{T}}$ satisfy (4.34), and we will show that $\{\varphi_t\}_{t\in\mathbb{T}}$ is a quasi-concave dPM.

(dMI) is straightforward. Let $V,V'\in\mathcal{X}$, such that $V\geq V'$. We will show that $\varphi_t(V)\geq\varphi_t(V')$ for any $t\in\mathbb{T}$. Consider $t,T\in\mathbb{T}$, such that $T\geq t$. Since $V_T\geq V'_T$, we have that $\ln V_T\geq\ln V'_T$, and consequently $\frac{\mu_t(\ln V_T)}{T}\geq\frac{\mu_t(\ln V'_T)}{T}$, for any $T\geq t$. Hence,

$$\liminf_{T\to\infty}\frac{\mu_t(\ln V_T)}{T}\geq\liminf_{T\to\infty}\frac{\mu_t(\ln V'_T)}{T}.$$

Next we prove (dL). Let us fix $t\in\mathbb{T}$ and $A\in\mathcal{F}_t$. For $T\geq t$, using (tL) of μ_t and the convention $0\cdot\infty=0$, we deduce

$$\begin{aligned}\mathbf{1}_A\varphi_t(\mathbf{1}_A\cdot_t V)&=\mathbf{1}_A\liminf_{T\to\infty}\frac{\mu_t(\ln\mathbf{1}_AV_T)}{T}=\liminf_{T\to\infty}\frac{\mathbf{1}_A\mu_t(\ln\mathbf{1}_AV_T)}{T}\\&=\liminf_{T\to\infty}\frac{\mathbf{1}_A\mu_t(\mathbf{1}_A\ln\mathbf{1}_AV_T)}{T}=\liminf_{T\to\infty}\frac{\mathbf{1}_A\mu_t(\mathbf{1}_A\ln V_T+\mathbf{1}_A\ln\mathbf{1}_A)}{T}\\&=\liminf_{T\to\infty}\frac{\mathbf{1}_A\mu_t(\mathbf{1}_A\ln V_T)}{T}=\liminf_{T\to\infty}\frac{\mathbf{1}_A\mu_t(\ln V_T)}{T}=\mathbf{1}_A\varphi_t(V).\end{aligned}$$

Finally, let us prove (dQCC). Let $t\in\mathbb{T}$, $V,V'\in\mathcal{X}$ and $\lambda\in\mathcal{X}_t$, $0\leq\lambda\leq1$. Without loss of generality, using the fact that μ_t is (tL), we assume that $0<\lambda<1$. Since log is monotone, and $V,V'\geq0$, we get

$$\begin{aligned}\varphi_t(\lambda\cdot_t V+(1-\lambda)\cdot_t V')&=\liminf_{T\to\infty}\frac{\mu_t(\ln[\lambda V_T+(1-\lambda)V'_T])}{T}\\&\geq\liminf_{T\to\infty}\Big[\min\Big\{\frac{\mu_t(\ln\lambda V_T)}{T},\frac{\mu_t(\ln(1-\lambda)V'_T)}{T}\Big\}\Big]\\&=\min\Big(\liminf_{T\to\infty}\frac{\mu_t(\ln V_T)}{T},\liminf_{T\to\infty}\frac{\mu_t(\ln V'_T)}{T}\Big)\\&=\varphi_t(V)\wedge\varphi_t(V'),\end{aligned}$$

which completes this part of the proof.
($\Rightarrow$) Assume that $\{\varphi_t\}_{t\in\mathbb{T}}$ is a quasi-concave dPM. Let $t\in\mathbb{T}$, $V\in\mathcal{X}$, and define $V'_s=V_s$ for $s\neq t$, and $V'_t=\min(1,V_t)$. Note that $V'\in\mathcal{X}$, and $V\geq V'$. As φ_t is (MI), we get

$$\liminf_{T\to\infty}\frac{\mu_t(\ln\frac{V_T}{V_t})}{T}\geq\liminf_{T\to\infty}\frac{\mu_t(\ln\frac{V'_T}{V'_t})}{T},$$

using (tL) of μ_t, we continue

$$1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(\ln\frac{V_T}{V_t})}{T}\geq 1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(1_{\{V_t\geq 1\}}\ln\frac{V'_T}{V'_t})}{T}.$$

Next, since $V'_t=1$ on the set $\{V_t\geq 1\}$, we have

$$1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(\ln\frac{V_T}{V_t})}{T}\geq 1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(1_{\{V_t\geq 1\}}\ln V'_T)}{T},$$

and since $V_T=V'_T$ for $T>t$, we finally conclude

$$1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(\ln\frac{V_T}{V_t})}{T}\geq 1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(\ln V_T)}{T}.$$

Note that $1_{\{V_t\geq 1\}}\ln\frac{V_T}{V_t}\leq 1_{\{V_t\geq 1\}}\ln V_T$ for $T>t$. By (MI) of μ_t, we get

$$1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(\ln\frac{V_T}{V_t})}{T}\leq 1_{\{V_t\geq 1\}}\liminf_{T\to\infty}\frac{\mu_t(\ln V_T)}{T}.$$

Combining the above inequalities, we have that equality (4.34) holds true on set $\{V_t\geq 1\}$. The proof for the set $\{V_t<1\}$ is similar. ☐

Relation (4.34) says that the value of the dLGI at time t is independent of the value of the process V at time t. As mentioned above, the purpose of dLGI is to measure the long term growth of V, which intuitively should not depend on the current state.

Remark 4.3.3. For $\mathcal{X}=\mathbb{V}^0_\tau$, an equivalent formulation of condition (4.34) is to require that for any $t\in\mathbb{T}$, $m\in L^0_t$ and $\{X_T\}_{T\in\mathbb{N}}$ such that $X_T\in\widehat{L}^0$, we have that

$$\liminf_{T\to\infty}\frac{\mu_t(X_T+m)}{T}=\liminf_{T\to\infty}\frac{\mu_t(X_T)}{T}.$$

In particular, this will be satisfied if there exists a family of maps $f_t:L^0_t\to L^0_t$ such that for all $X\in\widehat{L}^0$, $|\mu_t(X+m)-\mu_t(X)|\leq f_t(m)$ on the set $\{\mu_t(X)\neq\pm\infty\}$, and $\mu_t(X+m)=\mu_t(X)$ on $\{\mu_t(X)=\pm\infty\}$. For example, if μ_t is cash additive (tCA) then $f_t(m)=|m|$ (see also Proposition 4.3.6).

Corollary 4.3.4. *Let $\{\mu_t\}_{t\in\mathbb{T}}$ be local* (dL) *and monotone* (dMI)*, and let $\varphi=\{\varphi_t\}_{t\in\mathbb{T}}$ be a dLGI generated by $\{\mu_t\}_{t\in\mathbb{T}}$. Then*

1) The family φ is adapted (dA)*, local* (dL)*, scale invariant* (dSI) *and independent of the past* (dIP)*.*

2) If $\{\mu_t\}_{t\in\mathbb{T}}$ satisfies (4.34)*, then φ is quasi-concave dPM, i.e. φ is monotone* (dMI)*, quasi-concave* (dQCC) *and translation invariant* (dTI)*.*

Remark 4.3.5. By Corollary 4.3.4, any dLGI that is generated by $\{\mu_t\}_{t\in\mathbb{T}}$ which admits representation (4.34) fulfils all core conditions of dPMs introduced in [20] (except of time consistency and positiveness), and for static case introduced in [45], which were the object of study in the previous subsection (See Subsection 4.2). Thus, dLGI can be seen as a dynamic measure of performance of a given value process and in fact a dynamic acceptability index. Similar remark applies to dLGIs defined as $[\varphi_t(V)]^+$.Nevertheless, it should be mentioned that this class of maps is not normalized in the sense of [45].[13]

Next we will show that dLGIs that are also quasi-concave dPMs could be easily generated through dRMs or dCEs, as shown in the next two propositions.

Proposition 4.3.6. *For any dRM $\{\rho_t\}_{t\in\mathbb{T}}$ defined on $\widehat{L}^0$, the family $\{-\rho_t\}_{t\in\mathbb{T}}$ is local (dL), monotone (dMI) (hence generates a dLGI) and satisfies condition (4.34). Moreover, let $\{\widetilde{\rho}_t\}_{t\in\mathbb{T}}$ be given by $\widetilde{\rho}_t(X)=\rho_t(X^+)$. Then $\{-\widetilde{\rho}_t\}_{t\in\mathbb{T}}$ is also local, monotone and satisfies condition (4.34).*

Proof. Let $\{\rho_t\}_{t\in\mathbb{T}}$ be dRM defined on $\widehat{L}^0$. (MI) and (tL) of $\{-\rho_t\}_{t\in\mathbb{T}}$ follow directly from the definition of dRM. Let us fix $t\in\mathbb{T}$. First we will prove that condition (4.34) is satisfied by $\{-\rho_t\}_{t\in\mathbb{T}}$. For $V\in\mathcal{X}$, we have

$$\liminf_{T\to\infty}\frac{-\rho_t(\ln\frac{V_T}{V_t})}{T}=\liminf_{T\to\infty}\frac{-\rho_t(\ln V_T)-\ln V_t}{T}=\liminf_{T\to\infty}\frac{-\rho_t(\ln V_T)}{T}.$$

The above equality is straightforward on set $\{V_t>0\}$, since $\frac{\ln V_t}{T}\to 0$, $T\to\infty$. On the set $\{V_t=0\}$, we have that $\mathbb{1}_{\{V_t=0\}}V_T=0$, and by (tL) and (N) of $-\rho_t$, we get that both sides are equal to $(-\infty)$.

Next we will show that (4.34) also holds true for $\widetilde{\rho}$, given by $\widetilde{\rho}_t(X)=\rho_t(X^+)$. Let $V\in\mathcal{X}$. On the $\mathcal{F}_t$-measurable set $\{V_t=0\}$ both sides of (4.34) are equal to 0. Due to this, and (tL) of ρ_t, we can assume that $\mathbb{P}[V_t>0]=1$. Then, it is easy to note that

$$\begin{aligned}\liminf_{T\to\infty}\frac{-\rho_t([\ln\frac{V_T}{V_t}]^+)}{T}&=\liminf_{T\to\infty}\frac{-\rho_t(\mathbb{1}_{\{V_T>V_t\}}\ln\frac{V_T}{V_t})}{T}\\&=\liminf_{T\to\infty}\frac{-\rho_t(\mathbb{1}_{\{V_T>V_t\}}\ln V_T-\mathbb{1}_{\{V_T>V_t\}}\ln V_t)}{T}.\end{aligned}$$

Also, one can easily deduce the following inequalities

$$\mathbb{1}_{\{V_T>1\}}\ln V_T-2|\ln V_t|\leq\mathbb{1}_{\{V_T>V_t\}}\ln V_T-\mathbb{1}_{\{V_T>V_t\}}\ln V_t\leq\mathbb{1}_{\{V_T>1\}}\ln V_T+|\ln V_t|.$$

From the above, and monotonicity of dRM, we get

$$\liminf_{T\to\infty}\frac{-\rho_t([\ln V_T]^+-2|\ln V_t|)}{T}\leq\liminf_{T\to\infty}\frac{-\rho_t([\ln\frac{V_T}{V_t}]^+)}{T}\leq\liminf_{T\to\infty}\frac{-\rho_t([\ln V_T]^++2|\ln V_t|)}{T}.$$

Since $-\rho_t$ is (tCA), continue

$$\liminf_{T\to\infty}\frac{-\rho_t([\ln V_T]^+\pm 2|\ln V_t|)}{T}=\liminf_{T\to\infty}\frac{-\rho_t([\ln V_T]^+)\pm 2|\ln V_t|}{T}=\liminf_{T\to\infty}\frac{-\rho_t([\ln V_T]^+)}{T},$$

which concludes the proof. ∎

[13] i.e. $\varphi_t(V)=\infty$, if $V\geq 0$ and $\varphi_t(V)=0$, if $V<0$.

Proposition 4.3.7. *Let $\{\mu_t\}_{t\in\mathbb{T}}$ be dCE defined on $\widehat{L}^0$. Then, $\{\mu_t\}_{t\in\mathbb{T}}$ is local* (dL) *and monotone* (dMI) *(and hence generates a dLGI). Moreover, if additionally u from* (2.9) *is bi-Lipschitz on $\mathbb{R}$ (i.e. u and u^{-1} are Lipschitz), then $\{\mu_t\}_{t\in\mathbb{T}}$ satisfies* (4.34)*.*

Proof. Let $\{\mu_t\}_{t\in\mathbb{T}}$ be a dCE, with u being a continuous an increasing function. Clearly $\{\mu_t\}_{t\in\mathbb{T}}$ is (dA).

(dMI) is straightforward. Let us fix $t\in\mathbb{T}$. Let $X,Y\in\widehat{L}^0$, $X\geq Y$. Because u is increasing transform we get $u(X)\geq u(Y)$, and $E[u(X)|\mathcal{F}_t]\geq E[u(Y)|\mathcal{F}_t]$. Now, u^{-1} is also an increasing function, so $u^{-1}(E[u(X)|\mathcal{F}_t])\geq u^{-1}(E[u(Y)|\mathcal{F}_t])$.

Next we prove (dL). Note that any deterministic function, in particular u and u^{-1}, is local. Thus, for any $t\in\mathbb{T}$ and $A\in\mathcal{F}_t$, we have

$$\begin{aligned}\mathbb{1}_A\mu_t(X)&=\mathbb{1}_A u^{-1}(E[u(X)|\mathcal{F}_t])=\mathbb{1}_A u^{-1}(\mathbb{1}_A E[u(X)|\mathcal{F}_t])\\&=\mathbb{1}_A u^{-1}(E[1_A u(X)|\mathcal{F}_t])=\mathbb{1}_A u^{-1}(E[u(1_A X)|\mathcal{F}_t])\\&=\mathbb{1}_A\mu_t(1_A X),\end{aligned}$$

which proves that μ_t satisfies (tL).

Finally we will prove the second part of the Proposition 4.3.7. Let u be a bi-Lipschitz function with $L_u\in\mathbb{R}$ and $L_{u^{-1}}\in\mathbb{R}$ being the corresponding Lipschitz constants. Consider $t\in\mathbb{T}$ and $V\in\mathcal{X}$. On $\mathcal{F}_t$-measurable set $\{V_t=0\}$, $\mathbb{1}_{\{V_t=0\}}V_T=0$, and hence both sides of (4.34) are equal to $-\infty$.

From now on we make a (reasonable) assumption that $\mathbb{P}[V_t>0]>0$, which due to (tL) of μ_t, allows us to assume that $\mathbb{P}[V_t>0]=1$.

First we prove that for a fixed $T\in\mathbb{T}$, we get

$$\{u^{-1}(E[u(\ln V_T)|\mathcal{F}_t])=-\infty\}=\{u^{-1}(E[u(\ln\frac{V_T}{V_t})|\mathcal{F}_t])=-\infty\}.\tag{4.35}$$

As u is strictly increasing we know that (4.35) is equivalent to

$$\{E[u(\ln V_T)|\mathcal{F}_t]=u(-\infty)\}=\{E[u(\ln\frac{V_T}{V_t})|\mathcal{F}_t]=u(-\infty)\}.\tag{4.36}$$

Next we consider two cases: a) $u(-\infty)>-\infty$ and b) $u(-\infty)=-\infty$.

Case a) It is clear that the set $\{E[1_{\{V_T=0\}}|\mathcal{F}_t]=1\}$ is the subset of both sets in (4.36). Thus, it is sufficient to show that

$$P\Big[\{E[u(\ln V_T)|\mathcal{F}_t]=u(-\infty)\}\cap\{E[1_{\{V_T>0\}}|\mathcal{F}_t]>0\}\Big]=0\tag{4.37}$$

and

$$P\Big[\{E[u(\ln\frac{V_T}{V_t})|\mathcal{F}_t]=u(-\infty)\}\cap\{E[1_{\{V_T>0\}}|\mathcal{F}_t]>0\}\Big]=0.\tag{4.38}$$

Let us prove (4.37). Let

$$B:=\{E[u(\ln V_T)|\mathcal{F}_t]=u(-\infty)\}\cap\{E[1_{\{V_T>0\}}|\mathcal{F}_t]>0\}.$$

Note that $B\in\mathcal{F}_t$. On the contrary let us assume that $\mathbb{P}[B]>0$. Then

$$\mathbb{P}[\{V_T>0\}\cap B]=E[1_B E[1_{\{V_T>0\}}|\mathcal{F}_t]]>0.$$

Because $\{V_T > 0\} \cap B = \bigcup_{n\in\mathbb{N}} \{V_T > \frac{1}{n}\} \cap B$, we know that there exists $n_0 \in \mathbb{N}$, such that $\mathbb{P}[\{V_T > \frac{1}{n_0}\} \cap B] > 0$. Using that we obtain

$$\begin{aligned} E[1_B E[u(\ln V_T)|\mathcal{F}_t]] &= E[1_B E[1_{\{V_T > \frac{1}{n_0}\}} u(\ln V_T) + 1_{\{V_T \le \frac{1}{n_0}\}} u(\ln V_T)|\mathcal{F}_t]] \\ &\ge E[1_B E[1_{\{V_T > \frac{1}{n_0}\}} u(\ln \frac{1}{n_0}) + 1_{\{V_T \le \frac{1}{n_0}\}} u(-\infty)|\mathcal{F}_t]] \\ &= E[1_{B\cap\{V_T > \frac{1}{n_0}\}} u(\ln \frac{1}{n_0}) + 1_{B\cap\{V_T \le \frac{1}{n_0}\}} u(-\infty)] \\ &> E[1_B u(-\infty)]. \end{aligned} \tag{4.39}$$

Inequality (4.39) jointly with the definition of B leads to contradiction with the assumption that $P(B) > 0$, which verifies that (4.37) is true. The proof of (4.38) is analogous, since $P(V_t > 0) = 1$.
Case b) It is enough to show that

$$\{E[u(\ln V_T)|\mathcal{F}_t] = -\infty\} = \{E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t] = -\infty\}. \tag{4.40}$$

Now, because u is Lipschitz and $V_t > 0$, then, on the set $\{V_T > 0\}$ we get

$$u(\ln V_T) - L_u |\ln V_t| \le u(\ln \frac{V_T}{V_t}) \le u(\ln V_T) + L_u |\ln V_t|. \tag{4.41}$$

In addition, the above inequalities obviously hold true on the set $\{V_T = 0\}$, as on this set we have $u(\ln V_T) = u(\ln \frac{V_T}{V_t}) = u(-\infty) = -\infty$. Consequently,

$$E[u(\ln V_T)|\mathcal{F}_t] - L_u |\ln V_t| \le E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t] \le E[u(\ln V_T)|\mathcal{F}_t] + L_u |\ln V_t|. \tag{4.42}$$

Analogously, we obtain

$$E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t] - L_u |\ln V_t| \le E[u(\ln V_T)|\mathcal{F}_t] \le E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t] + L_u |\ln V_t|. \tag{4.43}$$

Combining (4.42) and (4.43), we obtain equality (4.40). So, (4.35) has been demonstrated.

Next, noting that $V_T < \infty$, and applying similar reasoning as in the proof of (4.35), one can show that

$$\{u^{-1}(E[u(\ln V_T)|\mathcal{F}_t]) = +\infty\} = \{u^{-1}(E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t]) = +\infty\}. \tag{4.44}$$

Now, let

$$K_T^- := \{u^{-1}(E[u(\ln V_T)|\mathcal{F}_t]) = -\infty\}, \quad K_T^+ := \{u^{-1}(E[u(\ln V_T)|\mathcal{F}_t]) = \infty\}, \quad T \in \mathbb{T}.$$

Combining (4.35) and (4.44) we obtain $\mu_t(\ln V_T) = \mu_t(\ln \frac{V_T}{V_t})$, on $\mathcal{F}_t$-measurable set $K_T^- \cup K_T^+$. On the set $(K_T^- \cup K_T^+)^c$ we get $|\mu_t(\ln V_T)| < \infty$ and $|\mu_t(\ln \frac{V_T}{V_t})| < \infty$. Moreover, since u is strictly increasing we also get $|E[u(\ln V_T)|\mathcal{F}_t]| < \infty$ and $|E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t]| < \infty$. Thus, using the fact that u is bi-Lipschitz, then, on set $(K_T^- \cup K_T^+)^c$, we get

$$\begin{aligned} |u^{-1}(E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t]) - u^{-1}(E[u(\ln V_T)|\mathcal{F}_t])| &\le L_{u^{-1}} |E[u(\ln \frac{V_T}{V_t})|\mathcal{F}_t] - E[u(\ln V_T)|\mathcal{F}_t]| \\ &\le L_{u^{-1}} L_u |\ln V_t|. \end{aligned} \tag{4.45}$$

We are now finally ready to prove the main statement. Let

$$K^- := \{\omega \in \Omega : \sum_{T\in\mathbb{T}} 1_{K_T^-}(\omega) < \infty\}, \quad K^+ := \{\omega \in \Omega : \sum_{T\in\mathbb{T}} 1_{(K_T^+)^c}(\omega) = \infty\}.$$

Using (4.45), on the set $K^- \cap K^+$ we obtain

$$\liminf_{T\to\infty} \frac{|\mu_t(\ln \frac{V_T}{V_t}) - \mu_t(\ln V_T)|}{T} \leq \liminf_{T\to\infty} \frac{L_u L_{u^{-1}} |\ln V_t|}{T} = 0.$$

which proves the equality (4.34) on this set. Using (4.35) we get the equality (4.34) on $(K^-)^c$; similarly, using (4.44) we get (4.34) on $(K^+)^c$. This completes the proof. ☐

Corollary 4.3.8. *By Proposition 4.3.6 and Proposition 4.3.2, any dLGI generated by $\mu_t = -\rho_t$, $t \in \mathbb{T}$, with $\{\rho_t\}_{t\in\mathbb{T}}$ being dRM, is quasi-concave dPM (for processes).*

4.3.1 Dynamic Risk Sensitive Criterion

Dynamic analog of Risk Sensitive Criterion [21], that we study in this section, is one of the most notable examples of dLGI. For simplicity we will assume that $\mathcal{X} = \mathbb{V}^0_\tau$.

Definition 4.3.9. A *Dynamic Risk Sensitive Criterion* is a family $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ of maps $\varphi^\gamma_t : \mathcal{X} \to \bar{L}^0_t$, indexed by $\gamma \in \mathbb{R}$, and defined by

$$\varphi^\gamma_t(V) = \begin{cases} \liminf_{T\to\infty} \frac{1}{T}\frac{1}{\gamma} \ln E[V_T^\gamma|\mathcal{F}_t] & \text{if } \gamma \neq 0, \\ \liminf_{T\to\infty} \frac{1}{T} E[\ln V_T|\mathcal{F}_t] & \text{if } \gamma = 0. \end{cases} \tag{4.46}$$

Remark 4.3.10. It is well known (cf. [61], and references therein) that for some processes V that are Markovian, the value of $\varphi^\gamma_t(V)$ is constant (independent of t in particular). In such cases of course, the analysis carried below trivialises. For example, let $V \in \mathcal{X}$ be such that $V_0 > 0$ and $V_t = V_0 \exp(\sum_{i=1}^t X_i)$, where $\{X_t\}_{t\in\mathbb{T}}$ is adapted, X_t is independent of $\mathcal{F}_{t-1}$ and $X_t \sim \mathcal{N}(0,1)$. In this case, $\varphi^\gamma_t(V) \equiv 0$. See also Subsection 5.1. Nevertheless, the class of processes V, for which $\varphi^\gamma_t(V)$ is a non-constant process, is quite rich; see e.g. (4.66) and (4.65).

We say that the Dynamic Risk Sensitive Criterion is risk-averse if $\gamma < 0$, risk neutral if $\gamma = 0$, and risk-seeking if $\gamma > 0$. Please note that with $t = 0$ we get the standard definition of (static) Risk Sensitive Criterion [21]; in particular, when $\gamma = 0$, the Risk Sensitive Criterion is called the Kelly criterion. We are now ready to present the main result of this Subsection. Arguably, properties 5) and 6) stated in Theorem 4.3.11 are the most interesting ones.

Theorem 4.3.11. *Let $\gamma \in \mathbb{R}$ and let $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ be a Dynamic Risk Sensitive Criterion. Then*

1) $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ is dLGI generated by $\{-\rho^\gamma_t\}_{t\in\mathbb{T}}$[14];

2) $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ is dPM, which admits (QCC);

3) $[\varphi^\gamma_t(V)]^+$ is a risk-seeking dLGI if and only if $\gamma > 0$;

4) $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ is increasing with γ, in $\mathbb{V}^1_{\ln}$;

[14] $\{\rho^\gamma_t\}_{t\in\mathbb{T}}$ denotes dynamic entropic risk measure with parameter $\gamma \in \mathbb{R}$, see Subsection 4.1.1 for details.

5) $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ *is supermartingale time consistent in* $\mathbb{V}^1_{\ln}$ *if and only if* $\gamma > 0$*;*

6) $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ *is submartingale time consistent in* $\mathbb{V}^1_{\ln}$ *if and only if* $\gamma < 0$*.*

Proof. For a fixed $\gamma \in \mathbb{R}$, let $\{\varphi^\gamma_t\}_{t\in\mathbb{T}}$ be a Dynamic Risk Sensitive Criterion.
1) It is enough to show that

$$\varphi^\gamma_t(V) = \liminf_{T\to\infty} \frac{-\rho^\gamma_t(\ln \frac{V_T}{V_t})}{T}, \quad t \in \mathbb{T},\ V \in \mathcal{X}. \tag{4.47}$$

Note that on $\mathcal{F}_t$-measurable set $\{V_t = 0\}$, $\mathbb{1}_{\{V_t=0\}}V_T = 0$, and hence both sides of (4.47) are equal to $-\infty$. Thus, due to (tL) of μ_t, it is enough to consider the case $\mathbb{P}[V_t > 0] = 1$.

For fixed $V \in \mathcal{X}$ and $t \in \mathbb{T}$ we have

$$\liminf_{T\to\infty} \frac{-\rho^\gamma_t(\ln \frac{V_T}{V_t})}{T} = \liminf_{T\to\infty} \frac{\ln E[\exp(\gamma \ln \frac{V_T}{V_t})|\mathcal{F}_t]}{\gamma T} = \liminf_{T\to\infty} \Big[\frac{1}{T}\frac{1}{\gamma} \ln E[V^\gamma_T|\mathcal{F}_t] - \frac{1}{T}\ln V_t\Big] = \varphi^\gamma_t(V).$$

For $\gamma = 0$, we immediately get

$$\liminf_{T\to\infty} \frac{-\rho^0_t(\ln \frac{V_T}{V_t})}{T} = \liminf_{T\to\infty} \Big[\frac{E[\ln V_T|\mathcal{F}_t]}{T} - \frac{\ln V_t}{T}\Big] = \liminf_{T\to\infty} \frac{1}{T} E[\ln V_T|\mathcal{F}_t] = \varphi^0_t(V).$$

2) It is an immediate result of Corollary 4.3.8 and 1), since $\{\rho^\gamma_t\}_{t\in\mathbb{T}}$ is dRM.
3) ($\Leftarrow$) It is enough to show that for $\gamma > 0$ we have (see Proposition 4.3.6)

$$\big[\varphi^\gamma_t(V)\big]^+ = \liminf_{T\to\infty} \frac{-\rho^\gamma_t([\ln \frac{V_T}{V_t}]^+)}{T}. \tag{4.48}$$

As in the previous case, without loss of generality, we can assume that $\mathbb{P}[V_t > 0] = 1$. For every $t \in \mathbb{T}$ and $V \in \mathcal{X}$, we deduce

$$\begin{aligned}
\liminf_{T\to\infty} \frac{-\rho^\gamma_t([\ln \frac{V_T}{V_t}]^+)}{T} &= \liminf_{T\to\infty} \frac{\ln E[\exp(\gamma [\ln \frac{V_T}{V_t}]^+)|\mathcal{F}_t]}{\gamma T} = \liminf_{T\to\infty} \frac{1}{T}\frac{1}{\gamma} \ln E\Big[\max\big(\frac{V_T}{V_t},1\big)^\gamma|\mathcal{F}_t\Big] \\
&= \liminf_{T\to\infty} \frac{1}{T}\frac{1}{\gamma} \ln E\Big[\frac{\max(V_T,V_t)^\gamma}{V^\gamma_t}|\mathcal{F}_t\Big] = \liminf_{T\to\infty} \Big[\frac{1}{T}\frac{1}{\gamma} \ln E[\max(V_T,V_t)^\gamma|\mathcal{F}_t] - \frac{1}{T}\ln V_t\Big] \\
&= \liminf_{T\to\infty} \frac{1}{T}\frac{1}{\gamma} \ln E[\max(V_T,V_t)^\gamma|\mathcal{F}_t].
\end{aligned} \tag{4.49}$$

Using the above, and the fact that $V_T \le \max(V_T, V_t)$, and $-\rho^\gamma_t([\ln \frac{V_T}{V_t}]^+) \ge 0$, for all $V \in \mathcal{X}$, we have the following inequality

$$\Big[\liminf_{T\to\infty} \frac{1}{T}\frac{1}{\gamma} \ln E[V^\gamma_T|\mathcal{F}_t]\Big]^+ \le \liminf_{T\to\infty} \frac{-\rho^\gamma_t([\ln \frac{V_T}{V_t}]^+)}{T}. \tag{4.50}$$

Next, we will prove the converse inequality. Without loss of generality, using locality, and the fact that the function $[\cdot]^+$ is non-negative, we could assume that

$$\liminf_{T\to\infty} \frac{-\rho^\gamma_t([\ln \frac{V_T}{V_t}]^+)}{T} > 0. \tag{4.51}$$

Let $X_T := E[\mathbb{1}_{\{V_T>V_t\}}V_T^{\gamma}|\mathcal{F}_t]$. Using (4.50), (4.49), and because $E[\mathbb{1}_{\{V_T\leq V_t\}}V_t^{\gamma}|\mathcal{F}_t]\leq V_t^{\gamma}$, we get

$$\begin{aligned}\liminf_{T\to\infty}\frac{1}{T}\frac{1}{\gamma}\ln X_T &\leq \Big[\liminf_{T\to\infty}\frac{1}{T}\frac{1}{\gamma}\ln E[V_T^{\gamma}|\mathcal{F}_t]\Big]^{+}\leq \liminf_{T\to\infty}\frac{-\rho_t^{\gamma}([\ln\frac{V_T}{V_t}]^{+})}{T}\\ &=\liminf_{T\to\infty}\frac{1}{T}\frac{1}{\gamma}\ln E[\max{(V_T,V_t)}^{\gamma}|\mathcal{F}_t]\leq \liminf_{T\to\infty}\frac{1}{T}\frac{1}{\gamma}\ln(X_T+V_t^{\gamma}).\end{aligned}\tag{4.52}$$

Due to (4.51), and the fact that $\gamma>0$, we have $(X_T+V_t^{\gamma})\overset{T\to\infty}{\longrightarrow}\infty$, and consequently $X_T\overset{T\to\infty}{\longrightarrow}\infty$. Thus,

$$|\ln(X_T+V_t^{\gamma})-\ln(X_T)|\to 0,\quad T\to\infty.$$

Using (4.52) we conclude the proof.

3) ($\Rightarrow$) For $\gamma=-1$ it is enough to consider a simple example

$$\widehat{V}_T(\omega)=\begin{cases} e^{-T} & \omega\in[0,e^{-T}],\\ e^{T} & \omega\in[e^{-T},1].\end{cases}$$

This example could be easily modified for any $\gamma<0$. For $\gamma=0$ it is enough to consider

$$\widehat{V}'_T(\omega)=\begin{cases} e^{-T^2} & \omega\in[0,\frac{1}{T}],\\ e^{T} & \omega\in[\frac{1}{T},1].\end{cases}$$

4) This is a direct result of the analogous property for negative of the dynamic entropic risk measure. See Proposition 4.1.3.

5) ($\Leftarrow$) Let $s\geq t\geq 0\in\mathbb{T}$, $V\in\mathbb{V}^1_{\ln}$, and $m_s\in\bar{L}^0_s$. It is enough to prove that

$$e^{\varphi_s^{\gamma}(V)}\geq e^{m_s}\Rightarrow e^{\varphi_t^{\gamma}(V)}\geq e^{E[m_s|\mathcal{F}_t]}.\tag{4.53}$$

It is easy to note, that

$$\begin{aligned}e^{\varphi_s^{\gamma}(V)}&=e^{\liminf_{T\to\infty}\frac{1}{T}\frac{1}{\gamma}\ln E[V_T^{\gamma}|\mathcal{F}_s]}=e^{\liminf_{T\to\infty}\ln\left[E[V_T^{\gamma}|\mathcal{F}_s]^{\frac{1}{\gamma T}}\right]}\\ &=\liminf_{T\to\infty}e^{\ln\left[E[V_T^{\gamma}|\mathcal{F}_s]^{\frac{1}{\gamma T}}\right]}=\liminf_{T\to\infty}E[V_T^{\gamma}|\mathcal{F}_s]^{\frac{1}{\gamma T}}.\end{aligned}$$

Using this, we conclude that (4.53) is equivalent to the following

$$\liminf_{T\to\infty}E[V_T^{\gamma}|\mathcal{F}_s]^{\frac{1}{\gamma T}}\geq e^{m_s}\Rightarrow \liminf_{T\to\infty}E[V_T^{\gamma}|\mathcal{F}_t]^{\frac{1}{\gamma T}}\geq e^{E[m_s|\mathcal{F}_t]}.\tag{4.54}$$

Assume that $\liminf_{T\to\infty}E[V_T^{\gamma}|\mathcal{F}_s]^{\frac{1}{\gamma T}}\geq e^{m_s}$. Due to the tower property we have

$$\liminf_{T\to\infty}E[V_T^{\gamma}|\mathcal{F}_t]^{\frac{1}{\gamma T}}=\liminf_{T\to\infty}E\big[E[V_T^{\gamma}|\mathcal{F}_s]|\mathcal{F}_t\big]^{\frac{1}{\gamma T}}.$$

Since, $0<\frac{1}{\gamma T}<1$, for T large enough, we get that the function $f(x)=x^{\frac{1}{\gamma T}},\ x>0$, is concave. Consequently, by Jensen's inequality, we continue

$$\liminf_{T\to\infty}E\big[E[V_T^{\gamma}|\mathcal{F}_s]|\mathcal{F}_t\big]^{\frac{1}{\gamma T}}\geq \liminf_{T\to\infty}E\big[E[V_T^{\gamma}|\mathcal{F}_s]^{\frac{1}{\gamma T}}|\mathcal{F}_t\big].$$

Since, $E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}}$ is non-negative for every $T \in \mathbb{T}$, by Fatou lemma, we conclude

$$\liminf_{T\to\infty} E\big[E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}}|\mathcal{F}_t\big] \geq E\big[\liminf_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}}|\mathcal{F}_t\big].$$

Finally, using the fact that $\liminf_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}} \geq e^{m_s}$, and by Jensen's inequality for $f(x) = e^x$, we get

$$E\big[\liminf_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}}|\mathcal{F}_t\big] \geq E[e^{m_s}|\mathcal{F}_t] \geq e^{E[m_s|\mathcal{F}_t]},$$

which completes the proof.

5) ($\Rightarrow$) Let $\gamma = 1$, and let $\{\widehat{V}_T\}_{T\in\mathbb{N}}$ be defined by

$$\widehat{V}_T(\omega) = \begin{cases} \frac{1}{T} & \omega \in [0, \frac{1}{T}], \\ e^T & \omega \in [\frac{1}{T}, 1]. \end{cases} \tag{4.55}$$

For $\omega \neq 0$, we have

$$\varphi_1^{-1}(\widehat{V}_T)(\omega) = \liminf_{T\to\infty} \frac{-1}{T} \ln \frac{1}{\widehat{V}_T(\omega)} = \liminf_{T\to\infty} [(-\frac{\ln T}{T}) \cdot \mathbb{1}_{[0,\frac{1}{T}]}(\omega) + 1 \cdot \mathbb{1}_{[\frac{1}{T},1]}(\omega)] = 1.$$

On the other hand

$$\varphi_0^{-1}(\widehat{V}_T) = \liminf_{T\to\infty} \frac{-1}{T} \ln E(\frac{1}{\widehat{V}_T}) = \liminf_{T\to\infty} \frac{-1}{T} \ln(1 + \frac{T-1}{T} e^{-T}) \leq \liminf_{T\to\infty} \frac{-\ln 1}{T} = 0.$$

Thus, with $m_1 = 1$, we get

$$\varphi_1^{-1}(\widehat{V}) \geq m_1 \not\Rightarrow \varphi_0^{-1}(\widehat{V}) \geq E[m_1|\mathcal{F}_0],$$

which contradicts acceptance consistency. This counterexample can be easily adjusted for any $\gamma < 0$.

Similarly, for $\gamma = 0$, we consider

$$\widehat{V}'_T(\omega) := \begin{cases} e^{-T^2} & \omega \in [0, \frac{1}{T}], \\ e^T & \omega \in [\frac{1}{T}, 1]. \end{cases}$$

6) ($\Leftarrow$) Let $t \in \mathbb{T}$, $V \in \mathbb{V}^1_{\ln}$ and $\gamma < 0$. We want to prove that for $s \in \mathbb{T}$, $s > t$, and $m_s \in \bar{L}^0_s$, we have

$$\varphi_s^\gamma(V) \leq m_s \Rightarrow \varphi_t^\gamma(V) \leq E[m_s|\mathcal{F}_t]. \tag{4.56}$$

Doing similar operations as in 5), we deduce that (4.56) is equivalent to

$$\liminf_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}} \leq e^{m_s} \Rightarrow \liminf_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{\gamma T}} \leq e^{E[m_s|\mathcal{F}_t]}. \tag{4.57}$$

Since for $\gamma < 0$ and nonnegative x the function $f(x) = x^\gamma$ is decreasing, we have that (4.57) is equivalent to

$$\Big[\liminf_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}}\Big]^\gamma \geq e^{\gamma m_s} \Rightarrow \Big[\liminf_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{\gamma T}}\Big]^\gamma \geq e^{\gamma E[m_s|\mathcal{F}_t]}$$

which is consequently equivalent to

$$\limsup_{T\to\infty}\Big[E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{\gamma T}}\Big]^\gamma \geq e^{\gamma m_s} \Rightarrow \limsup_{T\to\infty}\Big[E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{\gamma T}}\Big]^\gamma \geq e^{\gamma E[m_s|\mathcal{F}_t]},$$

From here, we conclude that (4.53) is equivalent to

$$\limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}} \geq e^{\gamma m_s} \Rightarrow \limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{T}} \geq e^{\gamma E[m_s|\mathcal{F}_t]}, \tag{4.58}$$

and thus we will verify this implication.

To give a better intuition of the proof of (4.58), first we will consider $t=0$, i.e we will show that that for any $m_s \in \bar{L}_s^0$, we have that

$$\limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}} \geq e^{\gamma m_s} \Rightarrow \limsup_{T\to\infty} E[V_T^\gamma]^{\frac{1}{T}} \geq e^{\gamma E[m_s]}. \tag{4.59}$$

Assume that $s>0$, $m_s \in \bar{L}_s^0$, and such that

$$\limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}} \geq e^{\gamma m_s}.$$

Note that, there exists a set $C \in \mathcal{F}_s$, such that $\mathbb{P}[C]>0$ and $\mathbb{1}_C e^{\gamma m_s} \geq \mathbb{1}_C E[e^{\gamma m_s}]$. Hence,

$$\mathbb{1}_C \limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}} \geq \mathbb{1}_C E[e^{\gamma m_s}].$$

By Jensen's inequality, we continue

$$\mathbb{1}_C \limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}} \geq \mathbb{1}_C e^{\gamma E[m_s]}. \tag{4.60}$$

Let $\epsilon>0$, and put $B_T^\epsilon := \{\omega\in\Omega : E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}}(\omega) \geq e^{\gamma E[m_s]-\epsilon}\}$. Notice that

$$C \subset \limsup_{T\to\infty} B_T^\epsilon,$$

which consequently implies that

$$P\Big[\limsup_{T\to\infty} B_T^\epsilon\Big] > 0. \tag{4.61}$$

From here, by Borel-Cantelli Lemma, we get that $\sum_{T=1}^\infty \mathbb{P}[B_T^\epsilon] = \infty$. Since the last series is divergent, there exists a subsequence $\{T_k^\epsilon\}_{(k=1,2,\ldots)}$ such that

$$\mathbb{P}[B_{T_k^\epsilon}^\epsilon] \geq \frac{1}{(T_k^\epsilon)^2}.$$

Using this, we have the following chain of inequalities

$$\begin{aligned}
\limsup_{T\to\infty} E[V_T^\gamma]^{\frac{1}{T}} &= \limsup_{T\to\infty} E[E[V_T^\gamma|\mathcal{F}_s]]^{\frac{1}{T}} \geq \limsup_{T\to\infty} E[\mathbb{1}_{B_T^\epsilon} E[V_T^\gamma|\mathcal{F}_s]]^{\frac{1}{T}} \\
&\geq \limsup_{T\to\infty} E[\mathbb{1}_{B_T^\epsilon} e^{(\gamma E[m_s]-\epsilon)T}]^{\frac{1}{T}} \geq e^{\gamma E[m_s]-\epsilon} \limsup_{T\to\infty} \mathbb{P}[B_T^\epsilon]^{\frac{1}{T}} \\
&\geq e^{\gamma E[m_s]-\epsilon} \limsup_{T_k^\epsilon\to\infty} \mathbb{P}[B_{T_k^\epsilon}^\epsilon]^{\frac{1}{T_k^\epsilon}} \geq e^{\gamma E[m_s]-\epsilon} \limsup_{T_k^\epsilon\to\infty} \Big[\frac{1}{(T_k^\epsilon)^2}\Big]^{\frac{1}{T_k^\epsilon}} \\
&= e^{\gamma E[m_s]-\epsilon}.
\end{aligned}$$

Hence, taking into account that $\epsilon > 0$ was arbitrary chosen, implication (4.59) follows immediately.

The proof for $t > 0$ follows similar line of ideas as for $t = 0$, although it is a bit more technical. For sake of completeness we will present the proof here too. The proof is done by contradiction: assume that (4.56) is not true for some $s \in \mathbb{T}$, $s > t$. Then, since (4.56) is equivalent to (4.58), there exists $V \in \tilde{\mathbb{V}}$, $m_s \in \bar{L}^0_s$ and $A \in \mathcal{F}_t$, $\mathbb{P}[A] > 0$ such that for

$$\limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}} \geq e^{\gamma m_s} \quad \text{and} \quad \limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{T}} < e^{\gamma E[m_s|\mathcal{F}_t]}. \tag{4.62}$$

almost surely on A. Note that there exists $\epsilon > 0$ and $A_2 \in \mathcal{F}_t$, $A_2 \subset A$, $\mathbb{P}[A_2] > 0$, such that

$$1_{A_2} \limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{T}} \leq 1_{A_2} e^{\gamma E[m_s|\mathcal{F}_t]-2\epsilon}. \tag{4.63}$$

Let us consider the following sets

$$\begin{aligned} B_T^\epsilon &:= \{\omega \in A_2 : E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}}(\omega) \geq e^{\gamma E[m_s|\mathcal{F}_t](\omega)-\epsilon}\}, \\ D_\alpha &:= \{\omega \in A_2 : \sum_{T=1}^{\infty} E[1_{B_T^\epsilon}|\mathcal{F}_t] < \alpha\}, \qquad \alpha \in \mathbb{N} \cup \{+\infty\}. \end{aligned}$$

Note that $D_n \in \mathcal{F}_t$ for any $n \in \mathbb{N}$, $D_n \subset D_m$ for $n \leq m$, and $D_\infty = \cup_{n\in\mathbb{N}} D_n \in \mathcal{F}_t$. Next we consider two cases: a) $\mathbb{P}[D_\infty] > 0$ and b) $\mathbb{P}[D_\infty] = 0$.

Case a) Since $\mathbb{P}[D_\infty] = \mathbb{P}[\lim_{n\to\infty} D_n] = \lim_{n\to\infty} \mathbb{P}[D_n] > 0$, there exists $n_0 > 0$ such that $\mathbb{P}[D_{n_0}] > 0$. Consequently,

$$\sum_{T=1}^{\infty} \mathbb{P}[B_T^\epsilon \cap D_{n_0}] < n_0.$$

From here, by Borel-Cantelli Lemma, we get

$$P\Big[\limsup_{T\to\infty} [B_T^\epsilon \cap D_{n_0}]\Big] = 0,$$

which implies that

$$1_{D_{n_0}} \limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_s]^{\frac{1}{T}} \leq 1_{D_{n_0}} e^{\gamma E[m_s|\mathcal{F}_t]-\epsilon}.$$

that contradicts (4.62) on some set of positive measure.

Case b) Let $\mathbb{P}[D_\infty] = 0$. First note that,

$$\begin{aligned} \limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{T}} &= \limsup_{T\to\infty} E[E[V_T^\gamma|\mathcal{F}_s]|\mathcal{F}_t]^{\frac{1}{T}} \geq \limsup_{T\to\infty} E[\mathbb{1}_{B_T^\epsilon} E[V_T^\gamma|\mathcal{F}_s]|\mathcal{F}_t]^{\frac{1}{T}} \\ &\geq \limsup_{T\to\infty} E[\mathbb{1}_{B_T^\epsilon} e^{(\gamma E[m_s|\mathcal{F}_t]-\epsilon)T}|\mathcal{F}_t]^{\frac{1}{T}} \geq e^{\gamma E[m_s|\mathcal{F}_t]-\epsilon} \limsup_{T\to\infty} E[1_{B_T^\epsilon}|\mathcal{F}_t]^{\frac{1}{T}}. \end{aligned} \tag{4.64}$$

Since $D_\infty \subset A_2$, and $\mathbb{P}[D_\infty] = 0$, we have that for (almost) every $\omega \in A_2$ there exists a subsequence $\{T_k^{\epsilon,\omega}\}_{k\in\mathbb{N}}$ such that

$$E\big[1_{B^\epsilon_{T_k^{\epsilon,\omega}}}|\mathcal{F}_t\big](\omega) \geq \frac{1}{(T_k^{\epsilon,\omega})^2}.$$

Using this, and (4.64), we conclude that for (almost) every $\omega \in A_2$

$$\limsup_{T\to\infty} E[1_{B_T^\epsilon}|\mathcal{F}_t]^{\frac{1}{T}}(\omega) \geq \limsup_{T_k^{\epsilon,\omega}\to\infty} E[1_{B^\epsilon_{T_k^{\epsilon,\omega}}}|\mathcal{F}_t]^{\frac{1}{T_k^{\epsilon,\omega}}}(\omega) \geq \limsup_{T_k^{\epsilon,\omega}\to\infty}\Big[\frac{1}{(T_k^{\epsilon,\omega})^2}\Big]^{\frac{1}{T_k^{\epsilon,\omega}}} = 1.$$

Thus, almost everywhere on A_2

$$\limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{T}} \geq e^{\gamma E[m_s|\mathcal{F}_t]-\epsilon}.$$

Combining the last inequality with (4.63), we get

$$1_{A_2}e^{\gamma E[m_s|\mathcal{F}_t]-2\epsilon} \geq 1_{A_2}\limsup_{T\to\infty} E[V_T^\gamma|\mathcal{F}_t]^{\frac{1}{T}} \geq 1_{A_2}e^{\gamma E[m_s|\mathcal{F}_t]-\epsilon},$$

which leads to contradiction, as $\mathbb{P}[A_2]>0$.
6) ($\Rightarrow$) As in the previous case we will consider only $\gamma=1$ and $\gamma=0$. For $\gamma=1$, we take $\{\widehat{V}_T\}_{T\in\mathbb{N}}$ defined by

$$\widehat{V}_T(\omega) = \begin{cases} Te^T & \omega\in[0,\frac{1}{T}], \\ 1 & \omega\in[\frac{1}{T},1]. \end{cases} \tag{4.65}$$

Then, we have

$$\varphi_1^1(\widehat{V}_T)(\omega) = \liminf_{T\to\infty}\frac{1}{T}\ln\widehat{V}_{T(\omega)} = \liminf_{T\to\infty}[(1+\frac{\ln T}{T})\cdot \mathbb{1}_{[0,\frac{1}{T}]}(\omega)+0\cdot\mathbb{1}_{[\frac{1}{T},1]}(\omega)] = 0, \quad \omega\neq 0.$$

On the other hand

$$\varphi_0^1(\widehat{V}_T) = \liminf_{T\to\infty}\frac{1}{T}\ln E(\widehat{V}_T) = \liminf_{T\to\infty}\frac{1}{T}\ln(e^T+\frac{T-1}{T}) \geq \liminf_{T\to\infty}\frac{T}{T} = 1.$$

Thus, with $m_1=0$, we get

$$\varphi_1^1(\widehat{V}) \leq m_1 \not\Rightarrow \varphi_0^1(\widehat{V}) \leq E[m_1|\mathcal{F}_0],$$

which contradicts rejection consistency.

Similarly, for $\gamma=0$, we consider

$$\widehat{V}'_T(\omega) = \begin{cases} e^{T^2} & \omega\in[0,\frac{1}{T}], \\ 1 & \omega\in[\frac{1}{T},1]. \end{cases}$$

$\square$

We conclude this section by presenting an example that is related to properties 4), 5) and 6).

Example 4.3.12. Let $([0,1],\mathcal{B}([0,1]),\{\mathcal{F}_t\}_{t\in\mathbb{N}_0},\mathbb{P})$ be a filtered probability space, where P is the standard Lebesgue measure, $\mathcal{B}(A)$ denotes the σ-algebra of Borel measurable sets of A, $\mathcal{F}_0$ is trivial and $\mathcal{F}_t=\sigma(K_t^1,\ldots,K_t^{2^t})$, where $K_t^i := [\frac{2(i-1)}{2^{t+1}},\frac{2i}{2^{t+1}}]$. Let $X(\omega)=\omega$ for $\omega\in[0,1]$, and let $\{\widehat{V}_T\}_{T\in\mathbb{N}_0}$ be defined by

$$\widehat{V}_T(\omega) = e^{TE[X|\mathcal{F}_T](\omega)}. \tag{4.66}$$

We will derive explicit formula for the dynamic risk sensitive criterion φ_t^γ. We start with the case of $\gamma = -1$. For fixed $t \in \mathbb{N}_0$, we get

$$\varphi_t^{-1}(\widehat{V}) = \liminf_{T\to\infty} \frac{-1}{T} \ln E[e^{-TE[X|\mathcal{F}_T]}|\mathcal{F}_t] = \liminf_{T\to\infty}(-1)\ln E[(e^{-E[X|\mathcal{F}_T]})^T|\mathcal{F}_t]^{1/T}.$$

Next for $\omega \in K_t^i$ and $T \in \mathbb{T}$, noting that $E[(e^{-E[X|\mathcal{F}_T]})^T|\mathcal{F}_t]^{1/T}(\omega)$ is in fact a power mean, we obtain

$$\limsup_{T\to\infty} E[(e^{-E[X|\mathcal{F}_T]})^T|\mathcal{F}_t]^{1/T}(\omega) \leq \limsup_{T\to\infty}[\operatorname*{ess\,sup}_{\omega\in K_t^i}(e^{-E[X|\mathcal{F}_T](\omega)})] \leq \operatorname*{ess\,sup}_{\omega\in K_t^i} e^{-X(\omega)} = e^{-\frac{2(i-1)}{2^{t+1}}}. \tag{4.67}$$

On the other hand using Jensen inequality, for any $T_0 \in \mathbb{T}$, such that $T_0 > t$, we get

$$\begin{aligned}\limsup_{T\to\infty} E[(e^{-E[X|\mathcal{F}_T]})^T|\mathcal{F}_t]^{1/T}(\omega) &= \limsup_{T\to\infty} E[E[e^{-TE[X|\mathcal{F}_T]}|\mathcal{F}_{T_0}]|\mathcal{F}_t]^{1/T}(\omega)\\ &\geq \limsup_{T\to\infty} E[e^{-TE[E[X|\mathcal{F}_T]|\mathcal{F}_{T_0}]}|\mathcal{F}_t]^{1/T}(\omega)\\ &= \limsup_{T\to\infty} E[(e^{-E[X|\mathcal{F}_{T_0}]})^T|\mathcal{F}_t]^{1/T}(\omega)\\ &= \operatorname*{ess\,sup}_{\omega\in K_t^i} e^{-E[X|\mathcal{F}_{T_0}]} = e^{-(\frac{2(i-1)}{2^{t+1}}+\frac{1}{2^{T_0+1}})}.\end{aligned} \tag{4.68}$$

Letting $T_0 \to \infty$, and combining (4.67) with (4.68), we conclude that for $\omega \in K_t^i$,

$$\varphi_t^{-1}(\widehat{V})(\omega) = (-1)\ln e^{-\frac{2(i-1)}{2^{t+1}}} = \frac{2(i-1)}{2^{t+1}}.$$

Using similar computations, it is easy to show that, for $\gamma \in \mathbb{R}$ and $\omega \in K_t^i$, we have

$$\varphi_t^\gamma(\widehat{V})(\omega) = \begin{cases} \frac{2(i-1)}{2^{t+1}} & \gamma < 0,\\ \frac{2(i-1)+2i}{2^{t+2}} & \gamma = 0,\\ \frac{2i}{2^{t+1}} & \gamma > 0.\end{cases}$$

Now, it clear from the above formula that $\varphi_t^\gamma(\widehat{V})$ is increasing in γ, so that property 4) is fulfilled. In addition, one can easily check that process $\varphi_t^\gamma(\widehat{V})$ is a submartingale (resp. supermartingale), with respect to the filtration $\{\mathcal{F}_t\}_{t\in\mathbb{N}_0}$, when $\gamma < 0$ (resp. $\gamma > 0$).

It is interesting to note that the values of $\varphi_t^\gamma(\widehat{V})$ are separated into three regimes: risk-seeking ($\gamma > 0$), risk-neutral ($\gamma = 0$) and risk-averse ($\gamma < 0$).

4.3.2 Dynamic Limit log-V@R

Let $\mathcal{X} = \mathbb{V}_\tau^0$.

Definition 4.3.13. A *Dynamic limit logarithmic Value at Risk* (Dynamic limit log-V@R) is a family $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ of mappings $\varphi_t^\alpha : \mathcal{X} \to \bar{L}_t^0$, indexed by $\alpha \in (0,1)$, and defined by

$$\varphi_t^\alpha(V) = \liminf_{T\to\infty} \frac{q_t^\alpha(\ln V_T)}{T} \tag{4.69}$$

where $q_t^\alpha(X)$ denotes $\mathcal{F}_t$-conditional (upper) α quantile of X.[15]

[15] i.e. $q_t^\alpha(X) = \operatorname{ess\,sup}\{Y \in \bar{L}_t^0 \mid E[\mathbb{1}_{\{X\leq Y\}}|\mathcal{F}_t] \leq \alpha\}$.

Proposition 4.3.14. *Let $\mathcal{X} = \mathbb{V}^1$ and let $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ denote a dynamic limit log-V@R. Then*

1) $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ is a dynamic acceptability index.

2) $\{\varphi_t^\alpha\}_{t\in\mathbb{T}}$ is increasing with α.

The proof of 1) is a direct result of Proposition 4.3.6. The proof of 2) is straightforward

Remark 4.3.15. We have introduced this mapping to show one interesting observation. The mapping $\rho(X) = -q_t^\alpha(X)$ corresponds to the conditional version of a standard Value at Risk (V@R). It is well known that V@R in general is not quasi-convex (QCV), see e.g. [77, Example 4.41]. Nevertheless, the corresponding dynamic limit growth index defined in (4.69) is quasi-concave (QCC) due to Proposition 4.3.6.

Chapter 5

Selected stochastic control problems

This chapter is devoted to the study of selected (financial) dynamic stochastic control problems in discrete-time. We will show, how one can use dynamic risk and performance measures both as value functions and as constraints in various optimisation problems. We have focused only on three representative problems, which could provide a general overview in this area. In particular we hope they will justify the need for further studies of time consistency property. For a general good survey about dynamic stochastic control and problems which arise in this area, please see [119]. Please also see e.g. [3, 109, 120, 136] and references therein for more problems, which involve stochastic control and (dynamic) risk and performance measures.

This chapter is organized as follows. The first example will be connected to dynamic limit growth indices, and will show how to solve infinite time horizon control problem for risk sensitive criterion, under certain ergodic assumptions on the underlying dynamics.

The second example will show how to solve a finite-time, multistep dynamic portfolio selection problem, when risk constraint is described by a strongly time consistent (coherent) dynamic risk measure.

Finally, the last example will show how to solve an optimal stopping problem for american options and explain why the least-square numerical approach to this problem is valid, even if the underlyings follow a non-Markovian dynamics and the problem is multidimensional.

5.1 Risk sensitive criterion for Markov decision processess

Many stochastic control methods are used in theoretical studies of portfolio management (cf. [120] and references therein). Among them, Risk Sensitive Criterion (RSC), introduced in Section 4.3.1 is one of the most recognised one. It has many advantages over the standard theoretical methods, which are usually based on expected utility criterion. Let us alone mention difficulties associated with the estimation of model parameters or traceable difficulties which arise when we try to compute optimal trading strategies for the realistic security market models. Moreover, following [22] we would like to stress out the fact, that risk sensitive criterion could be seen as a risk-to-reward criterion. Applying Taylor expansion around $\gamma = 0$, we get that the static version of RSC with parameter γ, denoted by φ^γ, could be presented as

$$\varphi^\gamma(V) = \liminf_{T\to\infty} \frac{1}{T}\frac{1}{\gamma} \ln E[V_T^\gamma] = \liminf_{T\to\infty} \frac{1}{T}\Big[E[\ln V_T] + \frac{\gamma}{2} Var(\ln V_T) + O(\gamma^2, t)\Big]. \tag{5.1}$$

This shows, that RSC could be seen as a measure of performance, which penalise expected growth rate with asymptotic variance multiplied by risk-averse parameter γ (for $\gamma < 0$). Of course this only appiles for problems, for which the last term (i.e. $O(\gamma^2, T)/T$) vanishes, when T goes to infinity. Nevertheless, this assumption is satisfied for a lot of standard dynamics, as explained in [22, Section 5], so (5.1) brings out the motivation, which led to this class of maps. Also, RSC is a Limit growth index, with gives its financial interpretation (see Section 4.3 for details). We refer to [22] for a further discussion about economic properties of RSC.

The study of RSC is connected to the optimal control literature, mostly to Markov controlled decision processes (see [91, 90, 58]) for infinite time horizon. The connection to portfolio optimisation was showed in [21], when RSC was applied to continuous time infinite time horizon, when a version of Merton's intertemporal capital asset pricing model was considered [111]. The analogous result for discrete-time market model was shown in [136].

There are many sophisticated methods used in the control theory, which guarantee the existence of the solution to Bellman equation associated with RSC. Let us alone mention the vanishing discount approach [89] or fixed point approach [58]. The assumptions under which the existence of the solutions is guaranteed are related to ergodic properties of the considered process [58, 101, 90, 89]. The most recent results relate to Doeblin's conditions [31] or Markov splitting techniques [59]. The theory of RSC is also strictly connected to Multiplicative Poisson equations [59] and Issacs equations for ergodic cost stochastic dynamic games (cf. [89, 72, 51] and references therein).

In this example we will focus on the infinite discrete time horizon, and follow the standard fixed point approach (also called span-contraction approach) used in [136]. We will also adapt some of the ideas used in [60, 58, 59, 61, 51, 90]. In particular we will consider problems of the form (2.20), with various (ergodic) assumptions imposed on set $\mathcal{Z}$, which describe all admissible portfolios, and RSC will be used as optimality criterion.

This Section is organized in follows. Subsection 5.1.1 will be devoted to the general setup in which we will introduce the problem and make all assumptions (e.g. on dynamics, control, etc.). Next, in Subsection 5.1.2 we will introduce the Bellman equation which naturally arise when we study problem introduces in the previous section. We will also solve the stochastic control problem stated before in general framework. Finally, in Subsection 5.1.3 we will show exemplary dynamics, commonly used in practise, that could be fit to our model.

As mentioned before, most of the methods and ideas of proofs in this section is taken from [136], and fitted to dynamic risk measurement framework.

5.1.1 General setup

In this subsection let $\mathcal{X} = \mathbb{V}^0$ and let $\mathbb{T} = \mathbb{N}$. For a fixed $-1 < \gamma < 0$, let φ^γ denote the unconditional version of dynamic risk sensitive criterion introduced in (4.3.9), i.e.

$$\varphi^\gamma(V) := \liminf_{T\to\infty} \frac{1}{T}\frac{1}{\gamma}\ln E[V_T^\gamma]. \quad V \in \mathcal{X} \tag{5.2}$$

Let μ^γ denote the negative of unconditional version of the entropic risk measure for random variables introduced in (4.1.2), i.e.

$$\mu^\gamma(X) = \frac{1}{\gamma}\ln E[\exp(\gamma X)], \quad X \in L^0. \tag{5.3}$$

Given the set $\mathcal{A}$ and dynamics of $V^H \in \mathcal{X}$ for any $H \in \mathcal{A}$, we want to solve the optimal stochastic control problem

$$\sup_{H\in\mathcal{A}} \varphi^{\gamma}(V^H). \tag{5.4}$$

We will now present the specification of the set $\mathcal{A}$ and the dynamics of V_H (for any $H \in \mathcal{A}$) which we will consider in this subsection.

We will assume that the market consist of m risky assets (e.g. stocks, bonds, derivative securities) and k economical factors (e.g. rates of inflation, short term interest rates, dividend yields). Prices of m risky assets will be denoted by $S^i = (S^i_t)_{t\in\mathbb{T}} \in \mathcal{X}$ for $(i = 1,\dots,m)$ and levels of k economical factors will be denoted by $X^j = (X^j_t)_{t\in\mathbb{T}} \in \mathcal{X}$ for $(j = 1,\dots,k)$. We will use notation $S := (S^1,\dots,S^m)$ and $X := (X^1,\dots,X^k)$. We will use $\mathcal{A}$ to denote the set of all U-valued predictable processes[1], where U is a compact subset of $\mathbb{R}^m$. Elements of $\mathcal{A}$ will correspond to all admissible portfolio strategies H, where $H = (H^1,\dots,H^m)$ and $H^i = (H^i_t)_{t\in\mathbb{T}} \in \mathcal{X}$ is a part of capital invested in i-th risky asset (for $i = 1,\dots,m$). We will use notation $V^H = (V^H_t)_{t\in\mathbb{T}} \in \mathcal{X}$ to denote the portfolio value process, corresponding to strategy H. We will make the following assumptions:

(A.1) The filtration $\{\mathcal{F}\}_{t\in\mathbb{T}}$ will be generated by a sequence of $k+m$ stochastic processes $W^i \in \mathcal{X}$ for $(i = 1,\dots,k+m)$ and $W := (W^1,\dots,W^{k+m})$ will form a sequence of i.i.d random vectors[2] with law η (we will also use this symbol to denote the corresponding measure).

(A.2) The factor process X will be of the form

$$X_{t+1} = G(X_t, W_t) := (G^1(X_t,W_t),\dots,G^k(X_t,W_t)),$$

where $G^i\colon \mathbb{R}^k\times\mathbb{R}^{k+m} \to \mathbb{R}^k$ is a W-continuous[3] Borel measurable function (for $i = 1,\dots,k$). Moreover we will assume that $\{X_t\}_{t\in\mathbb{T}}$ is a Markov chain[4], which is uniformly ergodic, i.e.

$$\sup_{A\in\mathcal{B}(\mathbb{R}^k)} \sup_{x,y\in\mathbb{R}^k} |\mathbb{P}[G(x,W_1)\in A] - \mathbb{P}[G(y,W_1)\in A]| < 1. \tag{5.5}$$

(A.3) For any $H \in \mathcal{A}$, we will assume that the portfolio dynamics will be of the form

$$V^H_0 = v_0, \qquad \ln\frac{V^H_t}{V^H_{t-1}} = F(X_t, H_t, W_t), \tag{5.6}$$

where $F\colon \mathbb{R}^k\times U\times\mathbb{R}^{k+m} \to \mathbb{R}$ is a Borel measurable function and $v_0 > 0$[5]. Moreover we will assume that F is W-continuous.[6]

[1] i.e. V_t is $\mathcal{F}_{t-1}$-measurable ($V \in \mathcal{X}$)

[2] i.e. $W_t = (W^1_t,\dots,W^{k+m}_t)$ is independent of $\mathcal{F}_{t-1}$ and $Law(W_t) = Law(W_{t-1})$, for $t-1 \in \mathbb{T}$.

[3] i.e. $G(x_n,W_t) \xrightarrow{a.s.} G(x,W_t)$ as $x_n \to x$ ($x_n, x \in \mathbb{R}^k$).

[4] Note that $Law(W_t) = Law(W_1)$ for any $t \in \mathbb{T}$, so $\{\mathbb{P}[G(x,W_1)\in A]\}_{A\in\mathcal{B}(\mathbb{R}^k)}$ generates transition probability of Markov chain $\{X_t\}_{t\in\mathbb{T}}$.

[5] Note that the portfolio value process V^H must be always non-negative. Also $t \geq 1$ in (5.6).

[6] i.e. $F(x_n,h_n,W_t) \xrightarrow{a.s.} F(x,h,W_t)$ as $x_n \to x$ and $h_n \to h$ ($x_n, x \in \mathbb{R}^k$, $h_n, h \in U$).

(A.4) We will assume that there exist sequences of i.i.d. random variables $\{K_t^+\}_{t\in\mathbb{T}}$ and $\{K_t^-\}_{t\in\mathbb{T}}$ ($K_t^\pm \in L_t^0$), such that for any $t \in \mathbb{T}$, we have

$$-\infty < \mu^{-1}(K_t^\pm) \leq \mu^\gamma(K_t^\pm) \leq \mu^1(K_t^\pm) < \infty, \qquad \mu^\gamma(K_t^\pm \ln K_t^\pm) \in \mathbb{R} \tag{5.7}$$

and for any $x \in \mathbb{R}^k$, $h \in U$, and $t \in \mathbb{T}$, we have

$$K_t^- \leq F(x, h, W_t) \leq K_t^+. \tag{5.8}$$

Remark 5.1.1. The upper and lower constraints introduced in (A.4) have a financial interpretation. They say that the utility (or risk) measured by μ^γ must be finite for any simple trade (in any state) and in fact it is jointly bounded by sequences K_t^- and K_t^+. As we consider the infinite-time horizon problem, this assumption relates to ergodicity conditions. Please note, that this assumption is rather weak, and fulfilled by standard models, which describe log-returns as processes of the form

$$F(x, h, W_t) = a(x, h, W_t) + \sum_{i=1}^{k+m} b(x, h) W_t^i,$$

where W_t is a random vector with multidimensional normal distribution and functions a and b are uniformly bounded. Then, random variables K_t^- and K_t^+ could be constructed using random variables $\min(W_t^1, \ldots, W_t^{k+m})$ and $\max(W_t^1, \ldots, W_t^{k+m})$, respectively. This is due to the fact that for any $h \in U$, we get

$$\min_{i\in\{1,\ldots,m\}} F(x, g^i, W_t) \leq F(x, h, W_t) \leq \max_{i\in\{1,\ldots,m\}} F(x, g^i, W_t),$$

where g^i are strategies such that we invest everything in the i-th asset. Intuitively speaking, our growth could be bounded by maximum and minimum of n-portfolios, for which the investment is made only in one asset. Of course when we make the investment, we don't know which of them will underperform/outperform us, so this bounds are theoretical.

Using the entropic representation of φ^γ and (5.6), for any $H \in \mathcal{A}$, we get

$$\varphi^\gamma(V^H) = \liminf_{t\to\infty} \frac{\mu^\gamma\big(\sum_{i=1}^t \ln \frac{V_i^H}{V_{i-1}^H}\big) + \ln v_0}{t} = \liminf_{t\to\infty} \frac{\mu^\gamma(\sum_{i=1}^t F(X_i, H_i, W_i))}{t}. \tag{5.9}$$

Under the above assumptions, from (5.9), it is not difficult to see, that the optimal value of the problem (5.4) will be finite, which is in fact the statement of Proposition 5.1.2.

Proposition 5.1.2. *Under assumption* (A.3) *and* (A.4), *we get*

$$-\infty < \sup_{H\in\mathcal{A}} \varphi^\gamma(V^H) < \infty.$$

Proof. Using (A.3) and (A.4), for any $H \in \mathcal{A}$ and $t \in \mathbb{T}$, we get $F(X_t, H_t, W_t) \leq K_t^+$. As the entropic risk measure μ^γ is strongly time consistent and law invariant, we know that μ^γ is additive for any two independent random variables [103]. Thus, for any $t \in \mathbb{T}$, we get

$$\mu^\gamma(\sum_{i=1}^t K_i^+) = \sum_{i=1}^t \mu^\gamma(K_i^+) = t\mu^\gamma(K_1^+).$$

Consequently, using (5.9) and monotonicity of μ^γ for any $H \in \mathcal{A}$, we get

$$\varphi^\gamma(V^H) = \big[\liminf_{t\to\infty} \frac{\mu^\gamma(\sum_{i=0}^{t-1} F(X_i, H_i, W_i))}{t}\big] \leq \liminf_{t\to\infty} \frac{\mu^\gamma(\sum_{i=1}^{t} K_i^+)}{t} = \mu^\gamma(K_1^+).$$

Consequently, by (5.7)

$$\sup_{H\in\mathcal{A}} \varphi^\gamma(V^H) \leq \mu^\gamma(K_1^+) < \infty.$$

The proof of the other inequality is analogous. □

Next, we will prove that under assumptions (A.1)-(A.4) there exists a Markov solution to problem (5.4).

5.1.2 Bellman equation

Using representation (5.9), it is not hard to see that the Bellman equation corresponding to (5.4) is of the form

$$v(x) + \lambda = \sup_{h\in U} \mu^\gamma(F(x, h, W_1) + v(G(x, W_1))), \tag{5.10}$$

where $\lambda \in \mathbb{R}$, $v \in \mathcal{C}(\mathbb{R}^k)$[7], $x \in \mathbb{R}^k$. To prove the existence of the solution to Bellman equation (5.10) we will adapt here the span contraction approach used e.g. in [136].

Remark 5.1.3. Bellman equation (5.10) is strictly connected to the Multiplicative Poisson Equation (MPE) defined for corresponding γ (cf. [59] and references therein). Sufficient general conditions for which there exists a solution to MPE in the classic case (i.e. using ergodicity conditions and span norm or vanishing discount approach) could be found e.g. in [58, 101, 90, 89]. For a more general conditions (obtained using splitting Markov techniques or Doeblin's condition) see e.g. [59, 31]. Also using robust representation (4.1.5) of the risk measure, one could notice that equation (5.10) corresponds to the Isaacs equation for ergodic cost stochastic dynamic game (cf. [89, 72] and references therein).

We will also make use of the span (semi-)norm of $f \in \mathcal{C}(\mathbb{R}^k)$, which is given by[8]

$$\|f\|_{\text{span}} := \sup_x f(x) - \inf_y f(y).$$

Moreover the operator T_γ corresponding to Bellman equation (5.10) is given by

$$T_\gamma f(x) := \sup_{h\in U} \mu^\gamma(F(x, h, W_1) + f(G(x, W_1))), \quad f \in \mathcal{C}(\mathbb{R}^k). \tag{5.11}$$

The operator defined in (5.11) plays a crucial role in the proof of the existence of the solution to Bellman equation (5.10) and it is strictly connected with the problem (5.4). We will now show some of it's properties.

Proposition 5.1.4. *Under assumptions* (A.1)–(A.4), *operator* T_γ *is Feller.*[9]

[7] $\mathcal{C}(\mathbb{R}^k)$ denotes the set of bounded and continuous functions $f : \mathbb{R}^k \to \mathbb{R}^k$.

[8] Notation of span norm is strictly connected to so called ω-norm and in particular to the Hilbert norm (cf. [133] and references therein).

[9] i.e. operator T transforms the set $\mathcal{C}(\mathbb{R}^k)$ of continuous bounded functions into itself.

Proof. Let $f \in \mathcal{C}(\mathbb{R}^k)$. Using (A.4) and the fact that μ^γ is cash-additive we know that for any $x \in \mathbb{R}^k$, we get

$$\mu^\gamma(K_1^-) + \inf_y f(y) \leq T_\gamma f(x) \leq \mu^\gamma(K_1^+) + \sup_y f(y),$$

which imply boundedness of $T_\gamma f(x)$. Next, let $\{(x_n, h_n)\}_{n\in\mathbb{N}}$ be a sequence such that $x_n \in \mathbb{R}^k$, $h_n \in U$ and $(x_n, h_n) \to (x, h)$ (where $x \in \mathbb{R}^k$, $h \in U$). By assumptions (A.2) and (A.3) we know that

$$e^{\gamma[F(x_n,h_n,W_1)+f(G(x_n,W_1))]} \xrightarrow{a.s.} e^{\gamma[F(x,h,W_1)+f(G(x,W_1))]}.$$

Moreover using (A.4) we know that

$$0 \leq e^{\gamma[F(x_n,h_n,W_1)+f(G(x_n,W_1))]} \leq e^{\gamma[K_1^- + \inf_y f(y)]} \quad \text{and} \quad e^{\gamma[K_1^- + \inf_y f(y)]} \in L^1$$

Thus, using dominated convergence theorem

$$E[e^{\gamma[F(x_n,h_n,W_1)+f(G(x_n,W_1))]}] \to E[e^{\gamma[F(x,h,W_1)+f(G(x,W_1))]}],$$

and consequently

$$\mu^\gamma(F(x_n, h_n, W_1) + f(G(x_n, W_1))) \to \mu^\gamma(F(x, h, W_1) + f(G(x, W_1))).$$

Let $h_z := \arg\max_{h\in U} \mu^\gamma(F(z, h, W_1) + f(G(z, W_1)))$, for any $z \in U$ (note that U is compact). Due to continuity of the function $(x, h) \mapsto \mu^\gamma(F(x, h, W_1) + f(G(x, W_1)))$, we also know that

$$\mu^\gamma(F(x_n, h_{x_n}, W_1) + f(G(x_n, W_1))) \to \mu^\gamma(F(x, h_x, W_1) + f(G(x, W_1))),$$

which imply continuity of T_γ. □

Proposition 5.1.5. *For any $t \in \mathbb{T}$, the value $T_\gamma^t 0(X_1)$ correspond to the optimal value of the problem (5.4) for a planning horizon of length t, i.e.*

$$\frac{T_\gamma^t 0(X_1)}{t} = \sup_{H\in\mathcal{A}} \frac{\mu^\gamma(\sum_{i=1}^t F(X_t, H_i, W_i))}{t}. \tag{5.12}$$

Proof. Before we prove (5.12) let us give some comments about operator T_γ and it's connection to problem (5.4). We know that $Law(W_1) = Law(W_t)$, for any $t \in \mathbb{N}$ and μ^γ is law-invariant, so we know that for any $f \in \mathcal{C}(\mathbb{R}^k)$ we get

$$T_\gamma f(x) := \sup_{h\in U} \mu^\gamma(F(x, h, W_t) + f(G(x, W_t))). \tag{5.13}$$

Moreover, as W_t is independent of $\mathcal{F}_{t-1}$, we could write

$$T_\gamma f(x) := \sup_{h\in U} \mu_{t-1}^\gamma(F(x, h, W_t) + f(G(x, W_t))). \tag{5.14}$$

where $\{\mu_t^\gamma\}_{t\in\mathbb{T}}$ corresponds to dUM defined in (4.5). For $t \in \mathbb{T}$, let $\mathcal{A}_t$ correspond to the set of all $\mathcal{F}_{t-1}$-measurable random variables, which take values in u (they could be interpreted as admissible strategies for time step t). Using (5.14) we could then write

$$T_\gamma f(Z) := \operatorname*{ess\,sup}_{H_t\in\mathcal{A}_t} \mu_{t-1}^\gamma(F(Z, H_t, W_t) + f(G(Z, W_t))). \tag{5.15}$$

for any $\mathcal{F}_{t-1}$-measurable random vector Z, which takes values in $\mathbb{R}^k$.

Thus, for any fixed $t \in \mathbb{T}$, using (5.14), (5.15), strong time consistency of dynamic entropic risk measure and Bellman principle of optimality, we get

$$\begin{aligned}
T^1_\gamma 0(X_t) &= \sup_{H_t \in \mathcal{A}_t} \mu^\gamma_{t-1}(F(X_t, H_t, W_t)) \\
&= \sup_{H \in \mathcal{A}} \mu^\gamma_{t-1}(F(X_t, H_t, W_t)), \\
T^2_\gamma 0(X_{t-1}) &= \sup_{H_{t-1} \in \mathcal{A}_{t-1}} \mu^\gamma_{t-2}(F(X_{t-1}, H_{t-1}, W_{t-1}) + \sup_{H_t \in \mathcal{A}_t} \mu^\gamma_{t-1}(F(G(X_{t-1}, W_{t-1}), H_t, W_t))) \\
&= \sup_{H_{t-1} \in \mathcal{A}_{t-1}} \mu^\gamma_{t-2}(\sup_{H_t \in \mathcal{A}_t} \mu^\gamma_{t-1}(F(X_{t-1}, H_{t-1}, W_{t-1}) + F(X_t, H_t, W_t))) \\
&= \sup_{H_{t-1} \in \mathcal{A}_{t-1}} \sup_{H_t \in \mathcal{A}_t} \mu^\gamma_{t-2}(\mu^\gamma_{t-1}(F(X_{t-1}, H_{t-1}, W_{t-1}) + F(X_t, H_t, W_t))). \\
&= \sup_{H \in \mathcal{A}} \mu^\gamma_{t-2}(F(X_{t-1}, H_{t-1}, W_{t-1}) + F(X_t, H_t, W_t)), \\
T^3_\gamma 0(X_{t-2}) &= \sup_{H \in \mathcal{A}} \mu^\gamma_{t-3}(F(X_{t-2}, H_{t-2}, W_{t-2}) + F(X_{t-1}, H_{t-1}, W_{t-1}) + F(X_t, H_t, W_t)), \\
\ldots &= \ldots \\
T^t_\gamma 0(X_1) &= \sup_{H \in \mathcal{A}} \mu^\gamma_0(\sum_{i=1}^t F(X_i, H_i, W_i)),
\end{aligned}$$

which completes the proof of (5.12). $\square$

Proposition 5.1.6. *Under assumptions (A.1)–(A.4), operator T_γ is a local contraction under $\|\cdot\|_{span}$, i.e. there exists $L : \mathbb{R}^+ \to (0, 1)$, such that*

$$\|T_\gamma f_1 - T_\gamma f_2\|_{span} \leq L(M)\|f_1 - f_2\|_{span},$$

for $M > 0$, $f_1, f_2 \in \mathcal{C}(\mathbb{R}^k)$, such that $\|f_1\|_{span} \leq M$ and $\|f_2\|_{span} \leq M$. Moreover the function L is independent of γ.

Proof. Let $M > 0$ and let $\gamma \in (-1, 0)$. Let $(\Omega, \mathcal{F}_1, \eta)$, be a probability space which corresponds to random variable W_1. For any $f \in \mathcal{C}(\mathbb{R}^k)$, such that $\|f\|_{\text{span}} \leq M$, $x \in \mathbb{R}^k$ and $h \in U$ we will use the following notation

$$h_{(x,f)} := \arg\max_{h \in U} \mu^\gamma(F(x, h, W_1) + f(G(x, W_1))), \tag{5.16}$$

$$Q_{(x,f,h)} := \arg\min_{\mathbb{Q} \in \mathcal{M}_1(\eta)} \Big[E_{\mathbb{Q}}[F(x, h, W_1) + f(G(x, W_1)] - \frac{1}{\gamma} H[Q\|\eta]\Big] \tag{5.17}$$

The measure $Q_{(x,f,h)}$ corresponds to the minimizing scenario in the robust representation (4.1.5) of μ^γ. To have a unique representation of $Q_{(x,f,h)}$ we will define it through *Esscher transformation* [85] introduced in (4.8) and given by

$$Q_{(x,f,h)}(dw) = \frac{e^{\gamma[F(x,h,w)+f(G(x,w))]}\eta(dw)}{E[e^{\gamma[F(x,h,W_1)+f(G(x,W_1))]}]}. \tag{5.18}$$

Using Proposition 4.1.7 and noticing that

$$[F(x,h,W_1)+f(G(x,W_1))]e^{F(x,h,W_1)+f(G(x,W_1))]}\in L^1,$$

due to assumption (A.4), we know that indeed (5.18) is the minimizer of (5.17). Moreover, we will define the measure $\bar{Q}_{(x,f,h)}$ on $\mathbb{R}^k$, by

$$\bar{Q}_{(x,f,h)}(A)=\frac{E\big[\mathbb{1}_{\{G(x,W_1)\in A\}}e^{\gamma[F(x,h,W_1)+f(G(x,W_1))]}\big]}{E\big[e^{\gamma[F(x,h,W_1)+f(G(x,W_1))]}\big]},\quad A\in\mathcal{B}(\mathbb{R}^k). \tag{5.19}$$

Step 1. Let $M>0$ and $f,g\in\mathcal{C}(\mathbb{R}^k)$ be such that $\|f\|_{\text{span}}\leq M$ and $\|g\|_{\text{span}}\leq M$. Using (5.16) we get

$$\begin{aligned}
T_\gamma f(x)&=\sup_{h\in U}\mu^\gamma(F(x,h,W_1)+f(G(x,W_1)))\\
&=\mu^\gamma(F(x,h_{(x,f)},W_1)+f(G(x,W_1)))\\
&=\inf_{\mathbb{Q}\in\mathcal{M}_1(\eta)}\Big[E_{\mathbb{Q}}[F(x,h_{(x,f)},W_1)+f(G(x,W_1)]-\frac{1}{\gamma}H[Q\|\eta]\Big]\\
&\leq E_{Q_{(x,g,h_{(x,f)})}}[F(x,h_{(x,f)},W_1)+f(G(x,W_1)]-\frac{1}{\gamma}H[Q_{(x,g,h_{(x,f)})}\|\eta]
\end{aligned} \tag{5.20}$$

Now, using (5.17) we get

$$\begin{aligned}
T_\gamma g(x)&=\sup_{h\in U}\mu^\gamma(F(x,h,W_1)+g(G(x,W_1)))\\
&\geq\mu^\gamma(F(x,h_{(x,f)},W_1)+g(G(x,W_1)))\\
&=\inf_{\mathbb{Q}\in\mathcal{M}_1(\eta)}\Big[E_{\mathbb{Q}}[F(x,h_{(x,f)},W_1)+g(G(x,W_1)]-\frac{1}{\gamma}H[Q\|\eta]\Big]\\
&=E_{Q_{(x,g,h_{(x,f)})}}[F(x,h_{(x,f)},W_1)+g(G(x,W_1)]-\frac{1}{\gamma}H[Q_{(x,g,h_{(x,f)})}\|\eta]
\end{aligned} \tag{5.21}$$

Combining (5.20) and (5.21) we get

$$\begin{aligned}
T_\gamma f(x)-T_\gamma g(x)&\leq E_{Q_{(x,g,h_{(x,f)})}}[f(G(x,W_1))-g(G(x,W_1))]\\
&\leq\int_{\mathbb{R}^k}[f(z)-g(z)]\bar{Q}_{(x,g,h_{(x,f)})}(dz).
\end{aligned} \tag{5.22}$$

Switching f with g in (5.22), for a fixed $y\in\mathbb{R}^k$, we get

$$T_\gamma g(y)-T_\gamma f(y)\leq\int_{\mathbb{R}^k}[g(z)-f(z)]\bar{Q}_{(y,f,h_{(y,g)})}(dz) \tag{5.23}$$

Combining (5.22) and (5.23) we get

$$\begin{aligned}
T_\gamma f(x)-T_\gamma g(x)-(T_\gamma f(y)-T_\gamma g(y))&\leq\int_{\mathbb{R}^k}\big[f(z)-g(z)\big][\bar{Q}_{(x,g,h_{(x,f)})}-\bar{Q}_{(y,f,h_{(y,g)})}](dz)\\
&\leq\frac{1}{2}\|f-g\|_{\text{span}}\|\bar{Q}_{(x,g,h_{(x,f)})}-\bar{Q}_{(y,f,h_{(y,g)})}\|_{\text{var}}.
\end{aligned} \tag{5.24}$$

where $\|\cdot\|_{\mathrm{var}}$ is a variation norm of a measure[10]. The last inequality is a result of the fact, that we can decompose the finite signed measure $\bar{Q}_{(x,g,h_{(x,f)})} - \bar{Q}_{(y,f,h_{(y,g)})}$ into positive and negative part (e.g. using Hahn-Jordan decomposition).

Step 2. We will now show that for any $\gamma \in (-1,0)$, there exists function $L^\gamma : \mathbb{R}^+ \to (0,1)$, such that for any $x, y \in \mathbb{R}^k$ and $f, g \in \mathcal{C}(\mathbb{R}^k)$ such that $\|f\|_{\mathrm{span}} \leq M$ and $\|g\|_{\mathrm{span}} \leq M$, we get

$$\|\bar{Q}_{(x,g,h_{(x,f)})} - \bar{Q}_{(y,f,h_{(y,g)})}\|_{\mathrm{var}} < 2L^\gamma(M). \tag{5.25}$$

On the contrary, let us assume that (5.25) is false. There exists $\gamma \in (-1,0)$ and a sequence

$$(x_n, y_n, f_n, g_n, A_n)_{n\in\mathbb{N}},$$

where $x_n, y_n \in \mathbb{R}^k$, $f_n, g_n \in \mathcal{C}(\mathbb{R}^k)$ ($\|f_n\|_{\mathrm{span}} \leq M$, $\|g_n\|_{\mathrm{span}} \leq M$) and $A_n \in \mathcal{B}(\mathbb{R}^k)$, such that

$$\bar{Q}_{(x_n,g_n,h_{(x_n,f_n)})}(A_n) - \bar{Q}_{(y_n,f_n,h_{(y_n,g_n)})}(A_n) \to 1. \tag{5.26}$$

Due to (5.26) we know that

$$\bar{Q}_{(x_n,g_n,h_{(x_n,f_n)})}(A_n^c) \to 0 \quad \text{and} \quad \bar{Q}_{(y_n,f_n,h_{(y_n,g_n)})}(A_n) \to 0 \tag{5.27}$$

Next, for any $x \in \mathbb{R}^k$, $h \in U$, $f \in \mathcal{C}(\mathbb{R}^k)$ ($\|f\|_{\mathrm{span}} \leq M$) and $A \in \mathcal{B}(\mathbb{R}^k)$, using Schwarz inequality, we get

$$\begin{aligned}
\bar{Q}_{(x,f,h)}(A) &= \frac{E\big[\mathbb{1}_{\{G(x,W_1)\in A\}} e^{\gamma[F(x,h,W_1)+f(G(x,W_1))]}\big]}{E\big[e^{\gamma[F(x,h,W_1)+f(G(x,W_1))]}\big]} \\
&\geq \frac{e^{\gamma\|f\|_{\mathrm{span}}} E\big[\mathbb{1}_{\{G(x,W_1)\in A\}} e^{\gamma[F(x,h,W_1)]}\big]}{e^{-\gamma\|f\|_{\mathrm{span}}} E\big[e^{\gamma[F(x,h,W_1)]}\big]} \frac{E\big[e^{-\gamma[F(x,h,W_1)]}\big]}{E\big[e^{-\gamma[F(x,h,W_1)]}\big]} \\
&\geq e^{2\gamma\|f\|_{\mathrm{span}}} \frac{E\big[\mathbb{1}_{\{G(x,W_1)\in A\}} e^{\gamma[F(x,h,W_1)]} e^{-\gamma[F(x,h,W_1)]}\big]}{E\big[e^{\gamma[F(x,h,W_1)]}\big] E\big[e^{-\gamma[F(x,h,W_1)]}\big]} \\
&= e^{2\gamma\|f\|_{\mathrm{span}}} \frac{E\big[\mathbb{1}_{\{G(x,W_1)\in A\}}\big]}{E\big[e^{\gamma[F(x,h,W_1)]}\big] E\big[e^{-\gamma[F(x,h,W_1)]}\big]}.
\end{aligned} \tag{5.28}$$

Next, using assumption (A.4), monotonicity of μ^γ and the fact that μ^γ is increasing with γ, we get

$$e^{2\gamma\|f\|_{\mathrm{span}}} \frac{E\big[\mathbb{1}_{\{G(x,W_1)\in A\}}\big]}{E\big[e^{\gamma[F(x,h,W_1)]}\big] E\big[e^{-\gamma[F(x,h,W_1)]}\big]} \geq e^{2\gamma\|f\|_{\mathrm{span}}} \frac{E\big[\mathbb{1}_{\{G(x,W_1)\in A\}}\big]}{e^{\gamma\mu^\gamma(K_1^+)} e^{\gamma\mu^\gamma(-K_1^-)}} \geq \alpha E\big[\mathbb{1}_{\{G(x,W_1)\in A\}}\big], \tag{5.29}$$

where $\alpha := e^{-2\|f\|_{\mathrm{span}}} \cdot \max\{e^{2|\mu^{-1}(K_1^+)|}, e^{2|\mu^0(K_1^+)|}, e^{2|\mu^{-1}(-K_1^-)|}, e^{2|\mu^0(-K_1^-)|}\}^{-1}$. Note that due to assumption (A.4), the value α is finite and in fact independent of γ. Combining (5.28) and (5.29) with (5.27) we get that

$$E\big[\mathbb{1}_{\{G(x_n,W_1)\in A_n^c\}}\big] \to 0 \quad \text{and} \quad E\big[\mathbb{1}_{\{G(y_n,W_1)\in A_n\}}\big] \to 0.$$

[10] For the general definition of total variation norm see e.g. [91]. In our framework, for two probability measures μ and ν on $\mathbb{R}^k$ with Borel sigma algebra, we get $\|\mu-\nu\|_{\mathrm{var}} = 2\sup_{A\in\mathcal{B}(\mathbb{R}^k)} |\mu(A)-\nu(A)|$ (again, see [91] for details).

Because of that

$$\mathbb{P}[G(x_n, W_1) \in A_n] - \mathbb{P}[G(y_n, W_1) \in A_n] \to 1,$$

which contradicts assumption (A.2), i.e. uniform ergodicity of $\{X_t\}_{t\in\mathbb{T}}$. Moreover, as the value α in (5.29) is independent of γ, and due to uniform ergodicity of $\{X_t\}_{t\in\mathbb{T}}$, we know that there will exists $L : \mathbb{R}^+ \to (0, 1)$, independent of γ, such that for any $\gamma \in (-1, 0)$, we get

$$\|\bar{Q}_{(x,g,h_{(x,f)})} - \bar{Q}_{(y,f,h_{(y,g)})}\|_{\text{var}} < L(M),$$

for $M > 0$, $x, y \in \mathbb{R}^k$ and $f, g \in \mathcal{C}(\mathbb{R}^k)$ ($\|f\|_{\text{span}} \leq M$, $\|g\|_{\text{span}} \leq M$), which concludes the proof. □

Propositions 5.1.4 and 5.1.6 allow us to show that there exists a solution to Bellman equation (5.10), while Proposition 5.1.5 tell us that the solution could be used to generate the Markov strategy, which is optimal for (5.4) as well as give the explicit value of (5.4).

Proposition 5.1.7. *Under assumptions (A.1)–(A.4) for sufficiently big $\gamma \in (-1, 0)$, there exist a unique (up to an additive constant) $v_\gamma \in \mathcal{C}(\mathbb{R}^k)$ and $\lambda_\gamma \in \mathbb{R}$, such that v_γ and λ_γ are solutions to Bellman equation (5.10).*

Proof. Using Propositions 5.1.4 and 5.1.6 it is easy to note that for any fixed $\gamma \in (-1, 0)$ the operator

$$\widetilde{T}_\gamma f(x) := \gamma T_\gamma \frac{f(x)}{\gamma},$$

is local contraction and Feller continuous. Indeed, for $f_1, f_2 \in \mathcal{C}(\mathbb{R}^k)$ ($\|f_1\|_{\text{span}} \leq M$, $\|f_2\|_{\text{span}} \leq M$), we get

$$\|\widetilde{T}_\gamma f_1 - \widetilde{T}_\gamma f_2\|_{\text{span}} = |\gamma| \cdot \|T_\gamma \frac{f_1}{\gamma} - T_\gamma \frac{f_2}{\gamma}\|_{\text{span}} \leq |\gamma| L(\frac{1}{|\gamma|} M) \|\frac{f_1 - f_2}{\gamma}\|_{\text{span}} = L(\frac{1}{|\gamma|} M) \|f_1 - f_2\|_{\text{span}}.$$

Moreover, repeating the proof of Proposition 5.1.6 (see also [136, 137] for the exact proof), we get that

$$\|\widetilde{T}_\gamma f_1 - \widetilde{T}_\gamma f_2\|_{\text{span}} \leq \widetilde{L}(M) \|f_1 - f_2\|_{\text{span}},$$

where $\widetilde{L} \colon \mathbb{R}^+ \to (0, 1)$ is independent of γ.

As $\widetilde{T}_\gamma$ is a contraction, using Banach's fixed point theorem (see e.g. [90, Appendix A]), we know that there exists at most one fixed point of operator $\widetilde{T}_\gamma$ in $\mathcal{C}(\mathbb{R}^k)$ endowed with the span norm. The same applies to operator T.

Let $\gamma \in (-1, 0)$ be such that $|\gamma| < 1 - \widetilde{L}(M)$[11], where $M := \mu^0(K_t^+) - \mu^{-1}(K_t^-)$. Using (A.4), due to monotonicity of μ^γ, it is easy to notice that

$$\|\widetilde{T}_\gamma 0\|_{\text{span}} \leq |\gamma| (\mu^\gamma(K_t^+) - \mu^\gamma(K_t^-)) \leq |\gamma| (\mu^0(K_t^+) - \mu^{-1}(K_t^-)) = |\gamma| M \leq M, \tag{5.30}$$

Moreover, as $|\gamma| < 1 - L(M)$, for any $n \in \mathbb{N}$, we get the following two inequalities:

$$\begin{aligned} \|\widetilde{T}_\gamma^n 0\|_{\text{span}} &\leq \Big[\sum_{k=0}^{n-1} L(M)^k\Big] \|\widetilde{T}_\gamma 0\|_{\text{span}} \leq \frac{|\gamma| M}{1 - L(M)} < M, \\ \|\widetilde{T}_\gamma^{n+1} 0 - \widetilde{T}_\gamma^n 0\|_{\text{span}} &\leq L(M)^n \|\widetilde{T}_\gamma 0\|_{\text{span}}. \end{aligned} \tag{5.31}$$

[11] Note that the set $\{\gamma \in (-1, 0) : |\gamma| < 1 - \widetilde{L}(M)\}$ is nonempty, as $\widetilde{L}(M) < 1$

Let us proof this statement using induction. For $n = 1$, using (5.30) we get

$$\|\widetilde{T}_\gamma^2 0\|_{\text{span}} \leq \|\widetilde{T}_\gamma^2 0 - \widetilde{T}_\gamma 0\|_{\text{span}} + \|\widetilde{T}_\gamma 0\|_{\text{span}} \leq \|\widetilde{T}_\gamma 0\|_{\text{span}}(1 + L(M)) \leq \frac{|\gamma| M}{1 - L(M)} < M,$$

which implies

$$\begin{aligned} \|\widetilde{T}_\gamma 0\|_{\text{span}} &\leq L(M)^0 \|\widetilde{T}_\gamma 0\|_{\text{span}} < M \\ \|\widetilde{T}_\gamma^2 0 - \widetilde{T}_\gamma 0\|_{\text{span}} &\leq L(M)\|\widetilde{T}_\gamma 0\|_{\text{span}}. \end{aligned}$$

Next, let us assume that (5.31) holds for a fixed $n \in \mathbb{N}$. Then, we get

$$\|\widetilde{T}_\gamma^{n+1} 0\|_{\text{span}} \leq \|\widetilde{T}_\gamma^{n+1} 0 - \widetilde{T}_\gamma^n 0\|_{\text{span}} + \|\widetilde{T}_\gamma^n 0\|_{\text{span}} \leq \Big[\sum_{k=0}^{n} L(M)^k\Big] \|\widetilde{T}_\gamma 0\|_{\text{span}} < M,$$

$$\|\widetilde{T}_\gamma^{n+2} 0 - \widetilde{T}_\gamma^{n+1} 0\|_{\text{span}} \leq L(M)\|\widetilde{T}_\gamma^{n+1} 0 - \widetilde{T}_\gamma^n\|_{\text{span}} \leq L(M)^{n+1}\|\widetilde{T}_\gamma 0\|_{\text{span}}.$$

which concludes the proof of (5.31).

Thus, by (5.31) we know that there exists a unique $\widetilde{v}_\gamma \in \mathbb{C}(\mathbb{R}^k)$ (up to an additive constant), such that

$$\|T\widetilde{v}_\gamma - \widetilde{v}_\gamma\|_{\text{span}} = 0.$$

Moreover, $v_\gamma(x) := \frac{\widetilde{v}_\gamma(x)}{\gamma}$ is a span norm fixed point of operator T_γ, as

$$\|T_\gamma v_\gamma - v_\gamma\|_{\text{span}} = \|\frac{1}{\gamma}\widetilde{T}_\gamma \widetilde{v}_\gamma - \frac{\widetilde{v}_\gamma}{\gamma}\|_{\text{span}} = 0. \tag{5.32}$$

Thus, there exists $\lambda_\gamma \in \mathbb{R}$[12], such that $v_\gamma(x)$ and λ_γ are solutions to Bellman equation (5.10) for a fixed $\gamma \in (-1, 1 - L(m))$.

□

Proposition 5.1.8. *Under assumptions (A.1)–(A.4) for sufficiently big $\gamma \in (-1, 0)$, there exist (Markov) solution to problem (5.4). Moreover, the optimal value is equal to λ_γ, and the optimal strategy is defined through v_γ.*

Proof. The proof of Proposition 5.1.8 for $\gamma \in (-1, 1 - L(M))$ is in fact a direct implication of Proposition 5.1.7 and classical results (i.e. verification theorems) from the theory of Risk-sensitive control (see e.g. [89, Theorem 2.2]). Nevertheless, we will show the proof based on properties of dUM to show how the dynamic version of risk measure is connected to the optimal solution, and also give some comment on the relation between infinite and finite time horizon optimal control problems.

Let us fix $\gamma \in (-1, 1 - L(M))$[13]. Let $v_\gamma \in \mathcal{C}(\mathbb{R}^k)$ and $\lambda_\gamma \in \mathbb{R}$ be a solution to Bellman equation (5.10). Following the proof of Proposition 5.1.7, it is easy to note that $\|v_\gamma\|_{\text{span}} \leq \frac{M}{|\gamma|}$ and $\|T_\gamma^t 0\|_{\text{span}} \leq \frac{M}{|\gamma|}$ for any $t \in \mathbb{T}$, where $M := \mu^0(K_t^+) - \mu^{-1}(K_t^-)$. Indeed for any $t \in \mathbb{T}$, we get

$$\gamma T_\gamma^t 0 = \gamma T_\gamma(T_\gamma^{t-1} 0) = \gamma \frac{1}{\gamma}\widetilde{T}_\gamma(\gamma T_\gamma^{t-1} 0) = \widetilde{T}_\gamma^2(\gamma T_\gamma^{t-2} 0) = \ldots = \widetilde{T}_\gamma^t(\gamma 0) = \widetilde{T}_\gamma^t(0),$$

[12] i.e. $\lambda_\gamma := T_\gamma v_\gamma(0) - v_\gamma(0)$

[13] this bound correspond to the one in Proposition 5.1.7

and we have shown in Proposition 5.1.7, that $\|\widetilde{T}^t_\gamma(0)\|_{\text{span}} \leq M$. Thus, we get

$$\begin{aligned}\|T^t_\gamma 0 - T^t_\gamma v_\gamma\|_{\text{span}} &\leq L(|\gamma|M)\|T^{t-1}_\gamma 0 - T^{t-1}_\gamma v_\gamma + \lambda_\gamma\|_{\text{span}} = L(|\gamma|M)\|T^{t-1}_\gamma 0 - T^{t-1}_\gamma v_\gamma\|_{\text{span}} \\ &\leq L(|\gamma|M)^2\|T^{t-2}_\gamma 0 - T^{t-2}_\gamma v_\gamma + \lambda_\gamma\|_{\text{span}} = L(|\gamma|M)^2\|T^{t-2}_\gamma 0 - T^{t-2}_\gamma v_\gamma\|_{\text{span}} \\ &\leq L(|\gamma|M)^t\|v_\gamma\|_{\text{span}}. \end{aligned} \tag{5.33}$$

Combining (5.12) and (5.33) we get

$$\begin{aligned}\sup_{H\in\mathcal{A}} \varphi^\gamma(V^H) &= \sup_{H\in\mathcal{A}} \liminf_{t\to\infty} \frac{\mu^\gamma(V^H)}{t} \leq \liminf_{t\to\infty} \sup_{H\in\mathcal{A}} \frac{\mu^\gamma(V^H)}{t} = \liminf_{t\to\infty} \frac{T^t_\gamma 0(X_1)}{t} \\ &= \liminf_{t\to\infty} \frac{T^t_\gamma v_\gamma(X_1)}{t} \leq \liminf_{t\to\infty} \frac{\lambda_\gamma t + \|v_\gamma\|_{\text{span}}}{t} = \lambda_\gamma\end{aligned}$$

On the other hand, the value λ_γ is obtained for the optimal Markov strategy H^* given by

$$H^*_t(x) = \arg\max_{h\in U} \mu^\gamma(F(x,h,W_1) + v_\gamma(G(x,W_1)))^{14},$$

which implies that λ_γ is optimal value for the problem (5.4) and optimal strategy is defined in terms of v_γ. □

5.1.3 Examplery dynamics

In this subsection let us present examples of dynamics for which assumptions (A.1)-(A.2) are fulfilled.

Example 5.1.9. In this example we will assume that time $\mathbb{T} = \mathbb{R}_+$ is continuous, but we can only reshape our portfolio in discrete time moments $n \in \mathbb{N}$. For $n \in \mathbb{N}$ and $(z = 1,\ldots,k+m)$, let us assume that W^z_n denotes the trajectory of $w_z(t) - w_z(n)$ $(n \leq t \leq n+1)$, where $\{w_z(t)\}_{z=1}^{k+m}$ are independent Brownian motions (which generates the filtration). Let us assume that the dynamics of the risky assets and factors is given by

$$\begin{aligned} X^j_n &= b_j(X_{n-1}) + \sum_{z=1}^{k+m} \delta_{jz}[w_z(n) - w_z(n-1)], && n\in\mathbb{N} \\ \frac{dS^i_t}{S^i_t} &= a_i(X_n)\,\mathrm{d}t + \sum_{z=1}^{k+m} \sigma_{iz}\,\mathrm{d}w_z(t), && t\in[n,n+1)\end{aligned}$$

where for $(i = 1,\ldots,m)$, $(j = 1,\ldots,k)$ and $(z = 1,\ldots,k+m)$: $a_i, b_i : \mathbb{R}^k \to \mathbb{R}$ are measurable and bounded functions, b_i is continuous, $\delta_{jz} \in \mathbb{R}$, $\sigma_{iz} \in \mathbb{R}$ and $\text{rank}((\sigma_{iz})_{z=1,\ldots,k+m}) = k$. Let $h_i(t)$ denote the part of the capital invested at time t in the i-th risky asset and let[15]

$$U = \{(h_1,\ldots,h_m) \in [0,1]^m : \sum_{i=1}^m h_i = 1\}.$$

[14] The exact strategy is given by $H^*_t = \arg\max_{H_t\in\mathcal{A}_t} \mu^\gamma_{t-1}(F(X_t,H_t,W_t) + v_\gamma(X_{t+1}))$.

[15] Note that we do not allow short selling, nor short borrowing.

Moreover, let $H_n^i = h_i(n)$. Using Ito's Lemma (see [136] for details) we get function F of the form

$$\begin{aligned}F(X_n, H_n, W_n) = &\sum_{i=1}^{m} \int_n^{n+1} a_i(X_n)h_i(s)\,\mathrm{d}s - \frac{1}{2}\sum_{z=1}^{k+m}\int_n^{n+1}\Big(\sum_{i=1}^{m} h_i(s)\sigma_{iz}\Big)^2\mathrm{d}s\\ &+\int_n^{n+1}\sum_{i=1}^{m} h_i(s)\sum_{z=1}^{k+m}\sigma_{iz}\,\mathrm{d}w_z(s).\end{aligned}$$

One can check that assumptions (A.1)-(A.4) will hold in this framework. See [136], where in fact equivalents of all Propositions from Section 5.1.2 are directly proved. For clarity, let us show the existence of the upper bound in (A.4). Following similar arguments as in Remark 5.1.1, we get

$$\begin{aligned}F(X_n, H_n, W_n) &= \ln\frac{V_{n+1}}{V_n} = \ln\sum_{i=1}^{m} H_n^i\frac{S_{n+1}^i}{S_n^i} = \ln\sum_{i=1}^{m} H_n^i e^{a_i(X_n)+\sum_{z=1}^{k+m}\sigma_{iz}[w_z(n+1)-w_z(n)]}\\ &\leq \sum_{i=1}^{m} H_n^i\big(a_i(X_n)+\sum_{z=1}^{k+m}\sigma_{iz}[w_z(n+1)-w_z(n)]\big)\\ &\leq \|a\|_{sup} + \max_{1\leq z\leq k+m}\sigma_{iz}[w_z(n+1)-w_z(n)]).\end{aligned}$$

Thus, $K_n^+ := \|a\|_{sup} + \max_{1\leq z\leq k+m}\sigma_{iz}[w_z(n+1)-w_z(n)])$ will satisfy (5.8). Moreover, it is easy to check that K_n^+ will satisfy (5.7), as for a Gaussian X, we get $e^{|X|}\in L^1$.

Example 5.1.10. Let us assume that assumptions (A.1) and (A.2) hold and the dynamics of risky assets is given by

$$\frac{S_{t+1}^i}{S_t^i} = \xi_i(X_t, W_t),$$

for $t\in\mathbb{N}$, where for $(i = 1,\ldots,m)$, ξ_t is a measurable vector function. Moreover the set u will be of the form $\{(h_1,\ldots,h_m)\in[0,1]^m : \sum_{i=1}^{m} h_i \leq 1\}$. Then we could define F explicitly, as

$$F(X_n, H_n, W_n) = \sum_{i=1}^{m} H_n^i\xi_i(X_n, W_n) + (1-\sum_{i=1}^{m} H_n^i).$$

To get assumptions (A.3) and (A.4) we need to impose additional assumptions on W and ξ_i. In particular we can consider the discretized version of Example 5.1.9 by setting $W_n^i = w_i(n+1)-w_i(n)$ and

$$\xi^i(X_n, W_n) = \exp\Big\{a_i(X_n) - \frac{1}{2}\sum_{z=1}^{k+m}\sigma_{iz}^2 + \sum_{z=1}^{k+m}\sigma_{iz}W_n^j\Big\}. \tag{5.34}$$

See [137] for details in general case and [60] for the case when (5.34) holds.

5.2 Portfolio optimisation with WV@R constrains

The stochastic control problems related to portfolio optimisation have a long history and have been studied intensively over the last 60 years. A major contribution, which could be regarded as the birth of modern portfolio analysis was done by H. Markowitz in 1952 [108]. Markowitz introduced the mean-variance optimisation problem

$$\begin{cases} E[X] \to max \\ Var(X) \leq c \\ X \in \mathcal{A} \end{cases} \tag{5.35}$$

where the random variable X is related to the (final) return of some financial portfolio, $Var(X)$ is it's variance, $c \in \mathbb{R}_+$ denotes the risk constraint and the set $\mathcal{A}$ describes all admissible values of X. One could also reformulate this problem, using portfolio Profits and losses instead of returns. The problem (5.35) was reformulated in many ways [79, 3, 120], mostly due to the fact that variance is not a good risk measure, as it penalizes profits in exactly the same way as losses.

We will focus on the problem, which substitutes the classical variance, with different dUM, namely a strongly time consistent coherent dynamic risk measure (on finite time horizon). We will use a family of dynamic WV@R as building blocks, due to their high analytical traceability. See [44] for a more detailed description for this class of maps.

Quasi-convex and coherent risk measures are used commonly in portfolio selection problems [3, 109], as they allow the use of many methods from the convex analysis and dynamic stochastic control. While the methods for static (one step) case are well studied, the dynamic selection model often cause a lot of problems. Bellman's principle of optimality is a crucial property, when we consider problems in a dynamic framework. Thus, a strong time consistency is a desired property, as explained in [7]. Moreover, the assumption about coherence of risk measure is very convenient as it often allows to reshape the problem in such a way, that we only need to maximize one objective function, instead of dealing with risk constraints (in a standard framework, one might say that Sharpe Ratio [66] could be used to solve problem 5.35).

This section is organized as follows. Subection 5.2.1 will be devoted to the general setup in which we will introduce the problem and make all assumptions (e.g. on dynamics, control, etc.). Next, in Subection 5.2.2 we will use the dynamic programming approach to completely solve problem introduces before.

This section is based on [44], written by A. Cherny. While we tried to shed some new light into problem presented in this paper, all results in this Section could be considered as counterparts of results from [44].

5.2.1 General setup

In this Section we will assume that the time horizon is finite, i.e. $\mathbb{T} = \{0, 1, \ldots, T\}$ for a fixed $T \in \mathbb{T}$ and $\mathcal{X} = \mathbb{V}^0$. Let ν be a fixed probability measure on $(0, 1]$ and let ρ^ν denote a corresponding static WV@R (for random variables on L^0) defined in (4.18), i.e.

$$\rho^\nu(X) := \int_0^1 \rho^\alpha(X)\nu(\mathrm{d}\alpha) = -\inf_{\mathbb{Q}\in\mathcal{D}^\nu} E_{\mathbb{Q}}[X], \quad X \in L^0, \tag{5.36}$$

where ρ^α is TV@R defined in (4.10) and D^ν is defined in (4.19). Moreover let $\{\rho^\nu_t\}_{t\in\mathbb{T}}$ denote the $\mathcal{D}^\nu$-composite dRM for stochastic processes defined in (5.37), where $\mathcal{D}^\nu$ denote the family of measures defining dWV@R in (4.21), i.e. for $(t = T-1, \dots, 1)$ we define recursively

$$\begin{aligned} \rho^{\nu}_T(V) &:= -V_T \\ \rho^{\nu}_t(V) &:= -\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{D}^{\nu}_t} E_{\mathbb{Q}}[V_t - \varphi^{\nu}_{t+1}(V)|\mathcal{F}_t]. \end{aligned} \tag{5.37}$$

Given the set $\mathcal{A}$ and dynamics of $V^H \in \mathcal{X}$ for any $H \in \mathcal{A}$, we want to solve the optimal stochastic control problem

$$\begin{cases} E\Big[\sum_{i=1}^T V_i^H\Big] \to max \\ \rho^{\nu}_0(V^H) \leq c \\ H \in \mathcal{A} \end{cases} \tag{5.38}$$

for a fixed $c \in \mathbb{R}_+$. Process V^H will correspond to the stream of dividend payoffs related to portfolio strategy, so the cumulated value at time t is given by $\sum_{i=1}^t V_i^H$. As we will work in concave framework, we will use φ^{ν} to denote the negative of ρ^{ν} and $\{\varphi^{\nu}_t\}_{t\in\mathbb{T}}$ to denote the negative of $\{\rho^{\nu}_t\}_{t\in\mathbb{T}}$, and deal with constraints $\varphi^{\nu}_0(V^H) \geq -c$, rather than $\rho^{\nu}_0(V^H) \leq c$.

We will now present the specification of the set $\mathcal{A}$ and the dynamics of V_H (for any $H \in \mathcal{A}$) which we will consider in this Section.

Similarly as in Subsection 5.1.2, prices of d risky assets will be denoted by $S^i = (S^i_t)_{t\in\mathbb{T}} \in \mathcal{X}$ for $(i = 1, \dots, d)$ and we will use notation $S := (S^1, \dots, S^d)$. We will use $\mathcal{A}$ to denote the set of all $\mathbb{R}$-valued, d-dimensional predictable processes. Elements of $\mathcal{A}$ will correspond to all admissible portfolio strategies H, where $H = (H^1, \dots, H^d)$ and $H^i = (H^i_t)_{t\in\mathbb{T}} \in \mathcal{X}$ is a part of capital invested in i-th risky asset (for $i = 1, \dots, d$). We will use notation $V^H = (V^H_t)_{t\in\mathbb{T}} \in \mathcal{X}$ to denote the portfolio cash-flows, corresponding to strategy H. We will make the following assumptions:

(A.1) The filtration $\{\mathcal{F}_t\}_{t\in\mathbb{T}}$ will be generated by a sequence of d stochastic processes $W^i \in \mathcal{X}$ for $(i = 1, \dots, d)$ and $W := (W^1, \dots, W^d)$ will form a sequence of i.i.d integrable random vectors with continuous distribution. Moreover, we will assume that $E[W_1] \geq 0$.[16]

(A.2) The price process $S = \{S_t\}_{t=0,\dots,T}$ will be of the form

$$S_0 \in \mathbb{R}^d_+, \qquad S_t = S_0 + \sum_{i=1}^t \sigma_i W_i, \tag{5.39}$$

where σ_t is a non-degenerate $\mathcal{F}_{t-1}$-measurable $d \times d$ matrix for $t \in \mathbb{T}$. Moreover, we will assume that $E[X_t] \neq 0$[17], where $X_t := (S_t - S_{t-1})$ denotes the stream of cash-flows associated with S.

(A.3) For any $H \in \mathcal{A}$, we will assume that the portfolio dynamics will be of the form

$$V^H_0 = 0, \quad V^H_t = \langle H_t, X_t\rangle, \tag{5.40}$$

where $\langle \cdot, \cdot \rangle$ denotes a standard scalar product and $t \in \mathbb{T}$ $(t > 0)$.

[16] We will use the notation $E[W_1] = (E[W^1_1], \dots, E[W^d_1])$.
[17] i.e. $(E[X^1_t], \dots, E[X^d_t]) \neq (0, 0, \dots, 0)$.

(A.4) We will assume that for any $h \in \mathbb{R}^d \setminus \{0\}$ and $t \in \mathbb{T}$ we get

$$-\infty < \varphi^{\nu}(\langle h, X_t \rangle) < 0, \tag{5.41}$$

where φ^{ν} is negative of a map given in (5.36).

Remark 5.2.1. In (A.2) the assumption $E[X_t] \neq 0$ ensures that there exists a trade with strictly positive reward. On the other hand in (A.3), the inequality (5.41) tell us that any simple trade has finite, strictly positive risk (negative utility). This assumption relates to no-arbitrage property (see e.g. [42, 44] for so called *no good deal* bounds) and is strictly related to the space $L^1_s(\mathcal{D}^{\nu})$ defined in (3.39), as for the family of dWV@Rs, we get $L^1_s(\mathcal{D}^{\nu}) = L^1_w(\mathcal{D}^{\nu})$ (see [42, Proposition 2.6]). Please also note that (5.41) implies that the probability of $\langle H_t, X_t \rangle$ being negative is always positive (for all $t \in \mathbb{T}$ and $H_t \neq 0$).

Remark 5.2.2. It is worth mentioning that the dynamics introduced in (5.39) cover the class of multidimensional GARCH models, which are a common tool used by practitioners to model prices (or log-returns) of assets [11].

5.2.2 Dynamic programming equations

In this Subsection we will consider the simplified problem (5.38). We will assume that $c = 1$ and $\sigma_t = \mathbb{I}_d$ for any $t \in \mathbb{T}$, where $\mathbb{I}_d$ denotes the d-dimensional unit matrix, i.e. we will consider the problem

$$\begin{cases} E\Big[\sum_{i=1}^{T} \langle H_i, W_i \rangle\Big] \to max \\ \varphi_0^{\nu}(\langle H, W \rangle) \geq -1 \\ H \in \mathcal{A} \end{cases} \tag{5.42}$$

Let us note that the problem (5.42) is scalable, in the sense, that if we consider cH instead of H (for $c \in \mathbb{R}_+$) in (5.42), then both the risk and the reward will be rescaled linearly. Because of that we can use normalized notation, which will be convenient, when we will define Bellman equations.

Remark 5.2.3. It is worth mentioning that because (5.42) is scalable, we might consider the related problem[18]

$$\sup_{H \in \mathcal{A}} \frac{E\big[\sum_{i=1}^{T} \langle H_i, W_i \rangle\big]}{-\varphi_0^{\nu}(\langle H, W \rangle)}, \tag{5.43}$$

for which the optimal value will be the same as in problem (5.42). The map introduced in (5.43) is in fact a RAROC PM, which we have introduced in Section 4.2.2. For more information about general risk-reward ratios, see [39].

As we know $\{\varphi_t^{\nu}\}_{t \in \mathbb{T}}$ is a strongly time consistent dUM. Because of that we might want to use Bellman's principle of optimality to define the corresponding Bellman equation and then decompose problem (5.38) into a series of conditional problems. We will use here the 'forward' approach. Let us define recursively the sequence of real numbers (for $t \in \mathbb{N}$)

$$\begin{aligned} U_0 &:= 0, \\ U_t &:= \sup_{h \in \mathcal{H}} E[\langle h, W_1 \rangle + U_{t-1}(\langle h, W_1 \rangle - a(h))^+], \end{aligned} \tag{5.44}$$

[18] Note that $\varphi_0^{\nu}(\langle H, X \rangle) \leq 0$, due to assumption (A.4), as will be proved later.

where

$$a(h) := \inf\{x \in \mathbb{R} : \ \varphi^{\nu}(\langle h, W_1\rangle \wedge x) \geq -1\}, \tag{5.45}$$

$$\mathcal{H} := \{h \in \mathbb{R}^d : \ \varphi^{\nu}(\langle h, W_1, \rangle) \geq -1\}. \tag{5.46}$$

The value U_t will correspond to the optimal value in problem (5.42) for time horizon of length t. Let us now present some heuristics which will give insight on the financial and mathematical interpretation of (5.44).

Remark 5.2.4. In the proof of Proposition 5.2.5 we will exploit the fact, that $\{\varphi_t^{\nu}\}_{t\in\mathbb{T}}$ is strongly time consistent dUMs.[19] Note that the conditional expectation also admits tower property, which is a form of strong time consistency. Indeed, the inequality (5.47) is a direct result of the Bellman principle of optimality together with scale-invariance nature of the problem. Those properties allow us to consider the conditional problem at time 1 and bound it by the optimal solution computed at time 0, for time horizon $T-1$ (which is expressed by U_{t-1}). From strong time consistency we also get that this bound is attained, and the optimal solution exists, as will be proved later.

Proposition 5.2.5. *The optimal value for the problem* (5.42) *with time horizon T is equal to U_T.*

Proof. For clarity, we will present only the general overview of the proof of Proposition 5.2.5, which will explain, why the optimal value is expressed through (5.44). For more details see [44, Theorem 3.1], where the detailed proof is provided for the general problem (5.38).[20]

We know that for 0-step problem, the optimal value U_0 equals to 0 , as $V_0^H = 0$ for any H. We will use induction to prove the thesis. Let us assume that the optimal value of the problem (5.42) for $(t-1)$-time horizon equals to U_{t-1} and consider problem (5.42) for time horizon of length t.

Using Bellman's principle of optimality and looking at the problem (5.42) at time 1 (i.e. for a fixed first step), it is easy to notice, that the remaining strategy (and thus the optimal value) should be the same (up to a non-negative constant[21]) as the strategy for the problem (5.42) with time horizon $t-1$. In other words, as the problem is scalable, assuming U_{t-1} is optimal value for $(t-1)$-step problem, and knowing that $\varphi_1^{\nu}(\langle H, X\rangle) - \langle H_1, W_1\rangle \leq 0$ (due to assumption (A.4), see [44, Lemma 3.3] for the proof), we should have the property

$$\frac{E\Big[\sum_{i=1}^{T}\langle H_i, W_i\rangle | \mathcal{F}_1\Big] - \langle H_1, W_1\rangle}{-[\varphi_1^{\nu}(\langle H, W\rangle) - \langle H_1, W_1\rangle]} \leq U_{t-1}, \tag{5.47}$$

where the interpretation of risk-reward ratio in (5.47) could be found in Remark 5.2.3. The detailed proof of (5.47) could be found in [44, Lemma 3.5]. Next, inequality (5.47) could be rewritten as

$$E\Big[\sum_{i=1}^{T}\langle H_i, W_i\rangle | \mathcal{F}_1\Big] \leq \langle H_1, W_1\rangle + U_{t-1}(\langle H_1, W_1\rangle - \varphi_1^{\nu}(\langle H, W\rangle)).$$

---

[19] On the space of stochastic processes described by all admissible strategies, defined in (5.40).

[20] The definition of the $\mathcal{D}^{\nu}$-composite dUM in [44] is slightly different, i.e. the map $\{\widetilde{\varphi}_t^{\nu}\}_{t\in\mathbb{T}}$ of the form $\widetilde{\varphi}_t^{\nu}(V) = \varphi_t^{\nu}(V) - V_t$ is considered. Nevertheless, all the results could be translated directly.

[21] This non-negative constant will be responsible for the risk control (amount of risk we can take), as the problem is scalable. We will show that it is indeed equal to $(\langle h, W_1\rangle - a(h))^+$ later.

Taking expectation on both sides we get

$$E\Big[\sum_{i=1}^{T}\langle H_i, W_i\rangle\Big] \le E\Big[\langle H_1, W_1\rangle + U_{t-1}(\langle H_1, W_1\rangle) - \varphi_1^{\nu}(\langle H, W\rangle)\Big]. \tag{5.48}$$

Thus, the good candidate for the optimal value U_t (with fixed control H_1) could be computed solving the problem

$$\begin{cases} E[Z + U_{t-1}(Z - Y(H))] \to max \\ Z - Y(H) \ge 0 \\ \varphi^{\nu}(Y(H)) \ge -1 \\ Y(H) \in L^0 \end{cases} \tag{5.49}$$

where $Z = \langle H_1, W_1\rangle$ is fixed and $Y(H) = \varphi_1^{\nu}(\langle H, W\rangle)$ is a function of H.[22] The first constraint in problem (5.49) is the result of assumption (A.4), while the second follows from the fact that $\{\varphi_t^{\nu}\}_{t\in\mathbb{T}}$ is constructed recursively as in (5.37) and due to the initial risk constraints, i.e. we get

$$\varphi^{\nu}(\varphi_1^{\nu}(\langle H, X\rangle) = \sup_{\mathbb{Q}\in D_0^{\nu}} E_{\mathbb{Q}}\Big[0 + \varphi_1^{\nu}(\langle H, W\rangle)\Big] = \varphi_0^{\nu}(\langle H, W\rangle) \ge -1.$$

Moreover, please note that we are only interested in Zs, such that $\varphi^{\nu}(Z) \ge -1$ (otherwise, the first constraint will not be fulfilled, due to monotonicity of φ^{ν}). For such random variables, problem (5.49) could be solved explicitly using Proposition 5.2.6.

Proposition 5.2.6. *Let $Z \in L^0$ be such that $\varphi^{\nu}(Z) \ge -1$. Then the optimal value in problem (5.49) is attained for $Y(H) = Z \wedge a$, where $a = \inf\{x \in \mathbb{R} : \ \varphi^{\nu}(Z \wedge x) \ge -1\}$.*

The proof of Proposition 5.2.6 could be found in [44, Lemma 3.2]. Using this Proposition for any $Z = \langle H_1, W_1\rangle$ such that $\varphi^{\nu}(Z) \ge -1$,[23] and taking the supremum, we get the formula for the global optimal value U_t (formally it is only the global upper bound, the proof that it is indeed attained could be found in [41, Lemma 3.7]), which will take the form

$$\begin{aligned} U_t &= \sup_{h\in\mathcal{H}} E[\langle h, W_1\rangle + U_{t-1}(\langle h, W_1\rangle - \langle h, W_1\rangle \wedge a(h))] \\ &= \sup_{h\in\mathcal{H}} E[\langle h, W_1\rangle + U_{t-1}(\langle h, W_1\rangle - a(h))^{+}], \end{aligned}$$

which is precisely the optimal value U_t defined in (5.44). Moreover, due to the fact that the set $\mathcal{H}$ is convex compact, which easily follows from (A.4), we know that the supremum in (5.44) is attained and can be substituted by maximum. □

Using similar arguments as before, one could also find the optimal control for problem (5.42). Indeed, It has been proven in [41, Theorem 3.1] that the optimal strategy H^* for problem (5.42) will be given by

$$H_t^* = C_{t-1}h^*(U_{T-t+1}), \tag{5.50}$$

[22] This problem is related to the maximisation of the right side in inequality (5.48). We assume that H_1 in H is fixed as well.

[23] The space of all such Zs coincide with the space $\mathcal{H}$, defined in (5.46).

where $h^*\colon \mathbb{R}_+ \to \mathbb{R}^d$ is defined as

$$h^*(x) := \arg\max_{h\in\mathcal{H}} E[\langle h, W_1\rangle + x(\langle h, W_1\rangle - a(h))^+],$$

and the adapted process $\{C_n\}_{t\in\mathbb{T}}$ is defined recursively, setting $C_0 = 1$ and

$$C_t = C_{t-1}(\langle h^*(U_{T-t+1}), W_n\rangle - a(h^*(U_{T-t+1})))^+.$$

One could interpret the value C_t as the amount of risk (in particular scenario) we can take at time $t \in \mathbb{T}$ for the problem considered over the period $(t+1, T)$.

Let us now go back to the original problem (5.38). We know that this problem could be easily transformed to problem (5.42) using simple change of variables. Thus, let us now provide the formulas for optimal value and optimal control for general problem (5.42).

Proposition 5.2.7. *The optimal value for the problem (5.38) with time horizon T equals cU_T and the optimal strategy H^* is given by*

$$H_t^* = cC_{t-1}\Big[(\sigma_t^*)^{-1} \cdot h^*(U_{T-t+1})\Big], \tag{5.51}$$

for $t \in \mathbb{T}$, where σ_t^* is the transpose of σ_t.

The exact proof of Proposition 5.2.7 could be found in [41, Theorem 3.1]. It also follows easily from Proposition 5.2.5. As for any t, the matrix σ_t is non-degenerate (and thus invertible). The set $\mathcal{A}$ consists of all predictable processes, so we know that by setting $\widetilde{H} = c[(\sigma_t^*)^{-1}H]$ for any $H \in \mathcal{A}$, we can transform problem (5.38) into problem (5.42). The optimal strategy in (5.51) is exactly such transformation of optimal strategy introduced in (5.50).

5.3 American option pricing – the least square approach

For over a decade several variants of the so called least-squares method of American option pricing have been widely used by financial practitioners and at the same time studied by researchers. The origins of the method can be found in the work of Carriere [30], Tsitsiklis, Van Roy [139] (see also [138]), Longstaff, Schwartz [106] and Clément, Lamberton, Protter [48]. Basically the method seeks a way of approximating conditional expectations needed in the valuation process either directly as in [106] and [48], or indirectly through the value function as in [139].

While the problem of optimal stopping for american options has been completely solved from the mathematical point of view with the help of Snell envelopes [104], the practical implementations still cause a lot of problems. The main reason is essentially the fact that the information space for conditional expectation, or in other words its range, is in many interesting cases infinite dimensional. Inevitably, in these cases any approximation of conditional expectations, or value functions depending on conditional expectations, has to involve significantly restrictive extrinsic assumptions to make practical computations possible. Due to this fact, one has to look for various algorithms, which approximate the optimal value.

In this Section we will extend the methods proposed by Clément, Lamberton and Protter [48], so that they cover the case of American style options with path dependent pay-offs, with a non-Markovian multidimensional underlying and with a very general approach to regression.

For possible practical applications of the results proposed in this Section, please see [99], where three computational examples are provided. First example concerns the pricing of a one year Eurodollar American put and call options, which (under the standard risk-neutral measure) are based on non-Markovian dynamics. The second example focuses on multidimensional options. Finally, the last example shows how to numerically price american options, when the underlying instruments follow the Heston-Nandi GARCH(1,1) model (see [99] for more details).

This Section is organized as follows. In the end of the introduction we provide more detailed discussion about least-square method, to connect our result with the existing literature. Subsection 5.3.1 is devoted to the short review of consequences of the classic Dobrushin-Minlos theorem, which can lead to viable numerical approximations of conditional expectations. Next, in Subsection 5.3.2 we give a short overview of Snell envelopes and comment on the relation to American-style option pricing. Finally, in Subsection 5.3.3 we extend the methods proposed by Clément, Lamberton and Protter [48] for American style options.

This Section will be based on [99].

Additional remarks

Let us now provide some more detailed insight about the least square approach to american option pricing. Several papers on this subject have been published — we will mention just a few of them.

A modification of the algorithm from [106] was studied in [48] from the point of view of the convergence of the method.

Glasserman and Yu [87] investigated in 2004 the convergence of the least-squares like methods, where — basically — the necessary conditional expectations are approximated by finite linear combinations of approximating functions. More specifically they look into the problem of accuracy of estimations when the number of approximating functions and the number of simulated trajectories increase. They assume that the underlying is a multidimensional Markov process. The rather pessimistic outcome, from the practical point of view, is that for polynomials as the approximating functions and for conventional (resp. geometric) Brownian motion as the underlying, the number of required paths may grow exponentially in the degree (resp. the square of the degree) of the polynomials. Glasserman and Yu remark that similar property may hold also for more general approximating functions (with the number of approximating functions replacing the maximal degree).

Also in 2004 Stentoft [135] analyzed and extended the convergence results presented in [48]. In particular he has considered the problem of choosing the optimal number of regressors in relation to the number of simulated trajectories.

In 2005 Egloff [64] proposed an extension to the original Longstaff-Schwartz [106] as well as Tsitsiklis – Van Roy ([138], [139]) algorithms by treating the optimal stopping problem for multidimensional discrete time Markov processes as a generalized statistical learning problem. His results also improve those from [48]. Egloff comments that despite very good performance of least-squares algorithms in some practical calculations, precise estimates of the statistical quantities involved in these procedures may be difficult, leading to some less impressive performance in other cases.

Zanger [144] proposed in 2009 another extension to the least-squares method by considering fairly arbitrary subsets of information spaces as the approximating sets. He has also produced some new and interesting convergence results showing in particular that sometimes the exponential dependence on the number of time steps can be avoided.

Two features seem to be common to the articles mentioned above. Firstly, the underlying is assumed to be Markovian. Secondly, the convergence rates of the method, in all its incarnations, are not encouraging from the computational point of view. In the present paper, we extend the Clément, Lamberton, Protter approach [48] to show that the method converges even if the underlying is not a Markov process and if the pay-offs are path-dependent, with a fairly general setting for the regression approximating conditional expectations. Obviously by giving up the Markov property and aiming at better approximation of conditional expectation, the potential computational complexity increases considerably. However, the main advantage of relaxation of the assumptions is the increase in freedom to customize the method. Moreover, we would like to argue that the least-squares methods should be seen as a general framework leading to a variety of specific implementations.

The main reason is essentially the fact that the information space for conditional expectation, or in other words its range, is in many interesting cases infinite dimensional. Inevitably, in these cases any approximation of conditional expectations, or value functions depending on conditional expectations, has to involve significantly restrictive extrinsic assumptions to make practical computations possible. While general convergence results are necessary to motivate the overall approach and some computational complexity may be addressed along the lines of [128], it is most likely that the future developments will evolve closer to simplified time-series models. It is quite conceivable that an alternative source of realism and numerical efficiency could exploit the advances in both time-series analysis and frame theory (see e.g. [98]). The empirical basis for such speculations comes from the fact, that in many real problems even taking only a few non-linear regressors, and sometimes ignoring lack of the Markov property, often leads to satisfactory results from the practical point of view. There seem to be much anecdotal evidence coming from the financial industry supporting the last statement.

It should be mentioned that the least squares approach can be also seen as part of the stochastic mesh framework proposed by Broadie and Glasserman ([27], [28]; see also [105] and [86]).

5.3.1 Approximation of conditional expectation

In the L^2 framework, dealing only with random variables of finite variance, we can rely on the Hilbert space geometry in addressing the issues of interest (see [134]). A closed subspace $S \subset L^2$ is said to be *probabilistic* if it contains constants and is closed with respect to taking the maximum of two of its elements, i.e. if $X, Y \in S$, then $X \vee Y \in S$. For any non-empty set $\mathbf{X} \subset L^2$, its *lattice envelope* $\mathrm{Latt}(\mathbf{X})$ is defined as the smallest probabilistic subspace of L^2 containing $\mathbf{X}$. Obviously, even if $\mathbf{X}$ consists of just one random variable, $\mathrm{Latt}(\mathbf{X})$ can be infinite-dimensional. Moreover, if $\mathbf{X} = \{X_1, \ldots, X_n\}$ and $\mathcal{B}_n$ denotes the σ-algebra of Borel sets in $\mathbb{R}^n$, then it is not difficult to prove that

$$\mathrm{Latt}(\mathbf{X}) = L^2(\sigma(\mathbf{X})) = L^2((X_1, \ldots, X_n)^{-1}(\mathcal{B}_n)).$$

The latter will be referred to as the *information space generated by* $X_1, \ldots, X_n$. Since this is also the range of the orthogonal projection $\mathrm{E}[\cdot \,|\, X_1, \ldots, X_n]$, it would be desirable from the numerical standpoint to be able to approximate such projections, with projections onto smaller finite-dimensional vector spaces using available least-squares algorithms.

To this end one could use the following theorem, which is a slight reformulation of a result of Dobrushin and Minlos [62].

Theorem 5.3.1. *Let $\alpha > 0$. Let $\mathcal{P}_n$ denote the space of all polynomials of n real variables. If $X_1, \ldots, X_n$ are random variables such that $e^{|X_j|} \in L^\alpha$ for $j = 1, \ldots, n$, then:*

(a) *$P(X_1, \ldots, X_n) \in L^p$ for any polynomial $P \in \mathcal{P}_n$ and $p \in [1, \infty)$;*

(b) *the vector space $\{P(X_1, \ldots, X_n) : P \in \mathcal{P}_n\}$ is dense in L^p for every $p \in [1, \infty)$.*[24]

It should be noted that the converse to part (a) is false as shown in the following example.

Example 5.3.2. Let $n = 1$ and let $\mathcal{F} = \sigma(X_1)$. Define an atomic probability measure $\mathbb{P}$ on the real axis via its probability mass function

$$f(x) = \mathbb{P}[X_1 = m] = \frac{\sum_{m=1}^{\infty} \frac{\delta(x-m)}{m^{\ln m}}}{\sum_{m=1}^{\infty} \frac{1}{m^{\ln m}}},$$

where $\delta(z) = \mathbb{1}_{\{z=0\}}$. If $q \geq 1$, then $\sum_{m=1}^{\infty} \frac{m^q}{m^{\ln m}} < \infty$. On the other hand, for any $\alpha > 0$, we get $\sum_{m=1}^{\infty} \frac{e^{\alpha m}}{m^{\ln m}} = \infty$.

If the probability measure $\mathbb{P}$ has a bounded support, in $\mathbb{R}^n$, then the assumption of the Dobrushin-Minlos theorem 5.3.1 is trivially satisfied. In fact, in this special case the conclusion of the theorem follows directly from the Stone-Weierstrass Theorem. It is also easy to see that if X is Gaussian, then $e^{|X|} \in L^1$. However, if X is lognormal, then its moment generating function does not exist in the interval $(0, \infty)$ and hence $e^{\alpha|X|} \notin L^\alpha$ for all $\alpha > 0$.

In concrete applications, the condition $e^{|X|} \in L^\alpha$ can sometimes be achieved by changing the probability distribution of "very large" values of $|X|$. For instance, this can be accomplished by truncation of probability distribution or some direct attenuation of the random variable X. Another possibility is the use of suitable weight functions. In this context the Dubrushin-Minlos theorem 5.3.1 can be used to justify the density part in the construction of several classic polynomial bases in spaces of square integrable functions, associated with the names of Jacobi, Gagenbauer, Legendre, Chebyshev, Laguerre and Hermite (see e.g. [47]).

Let V be an information space generated by random variables $X_1, \ldots, X_n$. Suppose that one can furnish a sequence of Borel functions $q_m : \mathbb{R}^n \longrightarrow \mathbb{R}$, with $m \in \mathbb{N}$, such that the set $\{q_m(X_1, \ldots, X_n) : m \in \mathbb{N}\}$ is linearly dense in V (e.g. with the help of the Dobrushin-Minlos theorem). Then the conditional expectation operator $\mathrm{E}[\cdot \mid X_1, \ldots, X_n]$ is the pointwise limit of the sequence of projections onto linear spaces $V^m = \{q_k(X_1, \ldots, X_n) : 1 \leq k \leq m\}$ as $m \nearrow \infty$. This observation leads to an auxiliary concept of admissible projection systems.

Given a discrete time filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_T \subset \mathcal{F}$ in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we define an *admissible projection system* as a family of orthogonal projections

$$\left\{ P_t^m : L^2(\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}) \right\}_{\substack{t = 1, \ldots, T \\ m \in \mathbb{N}}} \tag{5.52}$$

with ranges $V_t^m = P_t^m \left(L^2(\Omega, \mathcal{F}, \mathbb{P}) \right)$, such that for all $t = 1, \ldots, T$ and $m \in \mathbb{N}$ we have

$$V_t^m \subset V_t^{m+1}$$

[24] Note that we assume $\mathcal{F} = \sigma(X_1, \ldots, X_n)$ in $L^p = L^p(\Omega, \mathcal{F}, \mathbb{P})$.

and

$$\overline{\bigcup_{k\in\mathbb{N}} V_t^k} = L^2(\Omega, \mathcal{F}_t, \mathbb{P}).$$

Note that for any such system and for any fixed t, we get pointwise convergence of the projections P_t^m to $\mathrm{E}[\cdot \mid \mathcal{F}_t]$. However, this is not a norm convergence unless the underlying sequence of subspaces becomes constant after finitely many steps.

5.3.2 Approximation of Snell envelope

We consider the optimal stopping problem (2.22) for a fixed $Z = (Z_t)_{t=1}^T \in \mathbb{V}^1$, i.e.

$$\sup_{\nu \in C_0^T} \varphi(Z_\nu) \tag{5.53}$$

where C_0^T denote the set of all stopping times with values in $\mathbb{T} = \{0, 1, \dots, T\}$. As we have mentioned in the introduction (see Theorem 2.3.2), the dynamic programming principle for the problem (5.53) could be also rewritten in terms of the series of stopping times (τ_t), defined by recursion

$$\begin{aligned}
\tau_T &= T, \\
\tau_t &= t\mathbf{1}_{\{Z_t \geq E[Z_{\tau_{t+1}} \mid \mathcal{F}_t]\}} + \tau_{t+1}\mathbf{1}_{\{Z_t < E[Z_{\tau_{t+1}} \mid \mathcal{F}_t]\}}, \qquad t = 1, \dots, T-1.
\end{aligned}$$

In particular, we get $U_t = E[Z_{\tau_t} | \mathcal{F}_t]$ and consequently, τ_0 is optimal for (Z_t).

The key element in any numerical implementation of Snell envelopes is the ability to approximate the conditional expectation operator. Except for the finite case, one has to deal with infinite-dimensional spaces of random variables. Some elucidation seems to be in order here.

Given an admissible projection system (P_t^m), as in (5.52), for a fixed $m \in \mathbb{N}$ we define the stopping times $\tau_t^{[m]}$ by recursion:

$$\begin{aligned}
\tau_T^{[m]} &= T, \\
\tau_t^{[m]} &= t\mathbf{1}_{\{Z_t \geq P_t^m(Z_{\tau_{t+1}^{[m]}})\}} + \tau_{t+1}^{[m]}\mathbf{1}_{\{Z_t < P_t^m(Z_{\tau_{t+1}^{[m]}})\}}, \qquad t = 1, \dots, T-1.
\end{aligned} \tag{5.54}$$

The following theorem generalizes a result due to Clément, Lamberton and Protter (see Theorem 3.1 in [48]):

Theorem 5.3.3. *If (P_t^m) is an admissible projection system, then*

$$\lim_{m\to\infty} \mathrm{E}\left[Z_{\tau_t^{[m]}} \mid \mathcal{F}_t\right] = \mathrm{E}[Z_{\tau_t} \mid \mathcal{F}_t]$$

for $t = 1, \dots, T$, where the convergence is in L^2. In particular

$$\lim_{m\to\infty} \mathrm{E}\left[Z_{\tau_t^{[m]}}\right] = \mathrm{E}[Z_{\tau_t}]$$

in L^2.

Proof. Despite a much more general setting we have adopted here and slightly different notation, we can proceed as in [48]. Since the case $t = T$ is obvious, we can use induction on t. Assume that the formula is true for $t+1$. Let $\mathrm{E}_t[\,\cdot\,] = \mathrm{E}[\,\cdot \mid \mathcal{F}_t]$. Define five subsets of Ω as collections of points satisfying the following inequalities:

$$
\begin{aligned}
C_1 &= \{Z_t \geq P_t^m(Z_{\tau_{t+1}^{[m]}})\} \quad , \quad C_2 = \Omega \setminus C_1, \\
C_3 &= \{\ Z_t \geq \mathrm{E}_t[Z_{\tau_{t+1}}]\ \} \quad , \quad C_4 = \Omega \setminus C_3, \\
C_5 &= \left\{ \left| Z_t - \mathrm{E}_t[Z_{\tau_{t+1}}] \right| \ \leq \ \left| \mathrm{E}_t[Z_{\tau_{t+1}}] - P_t^m(Z_{\tau_{t+1}^{[m]}}) \right| \right\}.
\end{aligned}
$$

Obviously, for $t < T$ we have the formulas

$$
\begin{aligned}
\tau_t^{[m]} &= t\mathbf{1}_{C_1} + \tau_{t+1}^{[m]}\mathbf{1}_{C_2}, \\
\tau_t &= t\mathbf{1}_{C_3} + \tau_{t+1}\mathbf{1}_{C_4}.
\end{aligned}
$$

Observe that

$$
\begin{aligned}
\mathrm{E}_t\left[Z_{\tau_t^{[m]}} - Z_{\tau_t}\right] &= \mathrm{E}_t\left[Z_t\mathbf{1}_{C_1} + Z_{\tau_{t+1}^{[m]}}\mathbf{1}_{C_2} - Z_t\mathbf{1}_{C_3} - Z_{\tau_{t+1}}\mathbf{1}_{C_4}\right] \\
&= Z_t(\mathbf{1}_{C_1} - \mathbf{1}_{C_3}) + \mathrm{E}_t\left[Z_{\tau_{t+1}^{[m]}}\right]\mathbf{1}_{C_2} - \mathrm{E}_t\left[Z_{\tau_{t+1}}\right]\mathbf{1}_{C_4} \\
&= Z_t(\mathbf{1}_{C_1} - \mathbf{1}_{C_3}) + \mathrm{E}_t\left[Z_{\tau_{t+1}^{[m]}}\right]\mathbf{1}_{C_2} - \mathrm{E}_t[Z_{\tau_{t+1}}](\mathbf{1}_{C_1} + \mathbf{1}_{C_2} - \mathbf{1}_{C_3}) \\
&= \mathrm{E}_t\left[Z_{\tau_{t+1}^{[m]}} - Z_{\tau_{t+1}}\right]\mathbf{1}_{C_2} + \left(Z_t - \mathrm{E}_t\left[Z_{\tau_{t+1}}\right]\right)(\mathbf{1}_{C_1} - \mathbf{1}_{C_3}) \\
&= \mathrm{E}_t\left[Z_{\tau_{t+1}^{[m]}} - Z_{\tau_{t+1}}\right]\mathbf{1}_{C_2} + L_t^m.
\end{aligned}
$$

The first component of the sum in the last equality goes to zero by the induction hypothesis and the fact that $\mathrm{E}_t\mathrm{E}_{t+1} = \mathrm{E}_t$. We need to estimate the last term. To this end note that

$$
|\mathbf{1}_{C_1} - \mathbf{1}_{C_3}| \leq |\mathbf{1}_{C_1 \cap C_4} - \mathbf{1}_{C_2 \cap C_3}| \leq \mathbf{1}_{C_5},
$$

because $(C_1 \cap C_4) \cup (C_2 \cap C_3) \subset C_5$. Hence

$$
\begin{aligned}
L_t^m &\leq \left| Z_t - \mathrm{E}_t\left[Z_{\tau_{t+1}}\right] \right| \mathbf{1}_{C_5} \\
&\leq \left| \mathrm{E}_t[Z_{\tau_{t+1}}] - P_t^m(Z_{\tau_{t+1}^{[m]}}) \right|, \text{ by the definition of } C_5, \\
&\leq \left| \mathrm{E}_t[Z_{\tau_{t+1}}] - P_t^m\left(\mathrm{E}_t[Z_{\tau_{t+1}}]\right) \right| + \left| P_t^m\left(\mathrm{E}_t[Z_{\tau_{t+1}}]\right) - P_t^m(Z_{\tau_{t+1}^{[m]}}) \right| \\
&= \left| \mathrm{E}_t[Z_{\tau_{t+1}}] - P_t^m\left(\mathrm{E}_t[Z_{\tau_{t+1}}]\right) \right| + \left| P_t^m\left(\mathrm{E}_t[Z_{\tau_{t+1}}]\right) - P_t^m(\mathrm{E}_t[Z_{\tau_{t+1}^{[m]}}]) \right| \\
&\leq \left| \mathrm{E}_t[Z_{\tau_{t+1}}] - P_t^m\left(\mathrm{E}_t[Z_{\tau_{t+1}}]\right) \right| + \left| \mathrm{E}_t\left[Z_{\tau_{t+1}} - Z_{\tau_{t+1}^{[m]}}\right] \right|,
\end{aligned}
$$

because of the tower property of projections and the fact that the norm of a projection is at most one. The last term goes to zero by the induction hypothesis. The second last one because of the L^2 density of the union of ranges of the projection forming the admissible projection system. $\square$

It is straightforward to check, that Theorem 5.3.3 is also true for vector valued stochastic processes.

5.3.3 The general case of the least squares method of option pricing

In what follows we will denote the set of all real $(m \times n)$-matrices by $\mathbb{R}^{m\times n}$ with the convention that $\mathbb{R}^m = \mathbb{R}^{1\times m}$. Throughout the section we will use notation and methods similar to those introduced in [48] but adapted to our less restrictive assumptions.

Suppose that $(X_t)_{t=0}^T$ is a discrete time d-dimensional stochastic process, with X_0 being a constant. This process is meant to represent the prices of the underlying assets for an American style option we wish to valuate.

Let

$$X = (X_1, \dots, X_T) : \Omega \longrightarrow \mathbb{R}^{d\times T}$$

and let $\mathcal{F}_t = \sigma(X_0, \dots, X_t) = \sigma(X_1, \dots, X_t)$ for $t = 1, \dots, T$. Given a family of Borel functions

$$f_t : \mathbb{R}^{d\times(t+1)} \longrightarrow \mathbb{R}_+, \qquad t = 0, \dots, T,$$

we define

$$Z_t = f_t(X_0, \dots, X_t), \qquad t = 0, \dots, T.$$

This sequence represents suitably discounted intrinsic prices of the option we want to consider. Such a general choice of functions f_t, expands the potential applicability well beyond American put options.

Our goal is to to calculate U_0, where U_t is the Snell envelope of Z_t and since $U_0 = \max(Z_0, \mathrm{E}[Z_{\tau_1}])$, we basically want to approximate numerically $\mathrm{E}[Z_{\tau_1}]$. Let us now give some comments on the algorithm which does that and present the corresponding notation.

Step 1 - The setup

First of all, to use Theorem 5.3.3 we need to chose an admissible projection system for the filtration associated with X. This is equivalent to choosing for each $t \in \{1, \dots, T\}$ a suitable sequence of Borel functions

$$q_t^k : \mathbb{R}^{d\times T} \longrightarrow \mathbb{R}, \qquad k \in \mathbb{N},$$

which depend only on the first t column variables, and are such that the sequence $\{q_t^k(X)\}_{k\in\mathbb{N}}$ is linearly dense and linearly independent in the space $L^2(\Omega, \sigma(X_1, \dots, X_t), \mathbb{P})$. Then, we can select an increasing sequence of integers $(k_m)_{m\in\mathbb{N}}$, such that the spaces

$$V_t^m = \mathrm{Lin}\{q_t^k(X) \,:\, k = 1, \dots, k_m\},$$

and the orthogonal projections $P_t^m : L^2(\Omega, \sigma(X), \mathbb{P}) \longrightarrow V_t^m$ have all the right properties. The symbol "Lin" denotes the linear envelope of the given set of vectors.

If the stopping times $\tau^{[m]}$ are defined as in (5.54), then for some $\alpha_t^m \in \mathbb{R}^{k_m \times 1}$ we have

$$P_t^m \left(Z_{\tau_{t+1}^{[m]}} \right) = e_t^m(X)\, \alpha_t^m,$$

where the mapping e_t^m is given by the formula

$$e_t^m : \mathbb{R}^{d\times T} \ni x \mapsto \left(q_t^1(x), \ldots, q_t^{k_m}(x) \right) \in \mathbb{R}^{k_m}.$$

In view of our assumptions, the Gram matrix of the components of $e_t^m(X)$ (with respect to the inner product $(Y_1, Y_2) \mapsto \mathrm{E}[Y_1 Y_2]$), that is the matrix

$$A_t^m = \left[\mathrm{E}\left[q_t^i(X) q_t^j(X) \right] \right]_{1\le i,j\le k_m} \in \mathbb{R}^{k_m \times k_m},$$

is invertible and hence

$$\alpha_t^m = (A_t^m)^{-1} \begin{bmatrix} \mathrm{E}\left[Z_{\tau_{t+1}^{[m]}}\, q_t^1(X) \right] \\ \vdots \\ \mathrm{E}\left[Z_{\tau_{t+1}^{[m]}}\, q_t^{k_m}(X) \right] \end{bmatrix}.$$

Next, we want to use Monte-Carlo simulation, to approximate α_t^m and $\tau_t^{[m]}$. To do so we need to introduce the notation for the sample handling (i.e. independent trajectories) and to all the corresponding estimates of the functions.

Step 2 - Monte-Carlo simulation

Given a number N, let

$$X^{(n)} = \left(X_1^{(n)}, \ldots, X_T^{(n)} \right) \in \mathbb{R}^{d\times T}$$

denote independent trajectories of the process X, for $n = 1, 2, \ldots, N$. Each simulation has the fixed starting point $X_0^{(n)} = X_0 \in \mathbb{R}^{d\times 1}$. Define

$$Z_t^{(n)} := f_t \left(X_0^{(n)}, \ldots, X_t^{(n)} \right)$$

and let

$$\widehat{Z}_t = \begin{bmatrix} Z_t^{(1)} \\ \vdots \\ Z_t^{(N)} \end{bmatrix} \in \mathbb{R}^{N\times 1}.$$

This column vector consists simply of the values at time t of all simulated trajectories of the process Z. Define also

$$V_t^{(m,N)} = \mathrm{Lin} \left\{ \begin{bmatrix} q_t^k(X^{(1)}) \\ \vdots \\ q_t^k(X^{(N)}) \end{bmatrix} : k = 1, \ldots, k_m \right\} \subset \mathbb{R}^{N\times 1}$$

and

$$P_t^{(m,N)} = \mathrm{Proj}_{V_t^{(m,N)}} : \mathbb{R}^{N\times 1} \longrightarrow \mathbb{R}^{N\times 1}$$

with respect to the inner product $\frac{\langle x,y\rangle}{N}$, where $\langle x,y\rangle$ denotes the standard scalar product. Note that

$$\begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix} \in \mathbb{R}^{N\times k_m}$$

and

$$V_t^{(m,N)} = \mathrm{Lin}\left\{\text{the columns of } \begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix}\right\} \subset \mathbb{R}^{N\times 1}.$$

Next, we want to approximate stopping times $\tau_t^{[m]}$ defined similarly as in (5.54). To do so, we define the approximative stopping times, $\tau_t^{n,m,N}$ by the formula

$$\begin{aligned} \tau_T^{n,m,N} &= T, \\ \tau_t^{n,m,N} &= t\mathbf{1}_{\left\{Z_t^{(n)} \geq \prod_n \left[P_t^{(m,N)}(\widehat{Z}_{\tau_{t+1}^{n,m,N}})\right]\right\}} + \tau_{t+1}^{n,m,N}\mathbf{1}_{\left\{Z_t^{(n)} < \prod_n \left[P_t^{(m,N)}(\widehat{Z}_{\tau_{t+1}^{n,m,N}})\right]\right\}}, \\ & \quad \text{for } t = 1, \ldots, T-1, \end{aligned}$$

where

$$\prod_n : \mathbb{R}^{N\times 1} \ni \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \mapsto x_n \in \mathbb{R}.$$

Then, for some $\alpha_t^{(m,N)} \in \mathbb{R}^{k_m\times 1}$ we have

$$P_t^{(m,N)}\left(\begin{bmatrix} Z^{(1)}_{\tau_{t+1}^{1,m,N}} \\ \vdots \\ Z^{(N)}_{\tau_{t+1}^{N,m,N}} \end{bmatrix}\right) = \begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix} \alpha_t^{(m,N)}.$$

Let $A_t^{(m,N)}$ be the $(k_m \times k_m)$-Gram matrix associated with the columns of the matrix

$$\begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix},$$

that is

$$A_t^{(m,N)} = \frac{1}{N}\begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix}^* \begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix}.$$

This is simply the Gram matrix estimator for the given sample.

We know that $\alpha_t^{(m,N)}$ is a solution of the equation

$$A_t^{(m,N)}\alpha_t^{(m,N)} = \frac{1}{N}\begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix}^* \begin{bmatrix} Z^{(1)}_{\tau_{t+1}^{1,m,N}} \\ \vdots \\ Z^{(N)}_{\tau_{t+1}^{N,m,N}} \end{bmatrix}.$$

By the Law of Large Numbers $A_t^{(m,N)} \xrightarrow{a.s.} A_t^m$ as $N \to \infty$, and hence for sufficiently large N the matrix $A_t^{(m,N)}$ is invertible (almost surely). In this case

$$\alpha_t^{(m,N)} = \frac{1}{N}\left(A_t^{(m,N)}\right)^{-1}\begin{bmatrix} e_t^m(X^{(1)}) \\ \vdots \\ e_t^m(X^{(N)}) \end{bmatrix}^* \begin{bmatrix} Z^{(1)}_{\tau_{t+1}^{1,m,N}} \\ \vdots \\ Z^{(N)}_{\tau_{t+1}^{N,m,N}} \end{bmatrix}.$$

For convenience we shall write

$$\alpha^m = \left(\alpha_1^m, \ldots, \alpha_{T-1}^m\right), \qquad \alpha^{(m,N)} = \left(\alpha_1^{(m,N)}, \ldots, \alpha_{T-1}^{(m,N)}\right).$$

Both objects are $k_m \times (T-1)$-matrices.

Step 3 - Showing the convergence

Before showing that $\alpha_t^{(m,N)}$ converges to α_t^m and consequently, $\frac{1}{N}\sum_{n=1}^N Z^{(n)}_{\tau_t^{n,m,N}}$ converges to $\mathrm{E}\left[Z_{\tau_t^{[m]}}\right]$, we will introduce some additional notation. Let

$$B_t := \{(a^m, z, x) \;:\; z_t < e_t^m(x)a_t^m\} \subset \mathbb{R}^{k_m\times(T-1)} \times \mathbb{R}^T \times \mathbb{R}^{d\times T}$$

for $t = 1, \ldots, T-1$, where $a^m = (a_1^m, \ldots, a_{T-1}^m)$, $z = (z_1, \ldots, z_T)$, and $x = (x_1, \ldots, x_T)$. By B_t^c we will denote the complement of B_t. We define an auxiliary function

$$F_t : \mathbb{R}^{k_m\times(T-1)} \times \mathbb{R}^T \times \mathbb{R}^{d\times T} \longrightarrow \mathbb{R},$$

by recursion:

$$\begin{aligned} F_T(a^m, z, x) &= z_T, \\ F_t(a^m, z, x) &= z_t \mathbf{1}_{B_t^c} + F_{t+1}(a^m, z, x)\mathbf{1}_{B_t}, \qquad t = 1, \ldots, T-1. \end{aligned}$$

Since $\mathbf{1}_{C\cap D} = \mathbf{1}_C \mathbf{1}_D$ for any two sets C and D, it is easy to see that

$$F_t(a^m, z, x) = z_t \mathbf{1}_{B_t^c} + \sum_{s=t+1}^{T-1} z_s \mathbf{1}_{B_t\cap\ldots\cap B_{s-1}\cap B_s^c} + z_T \mathbf{1}_{B_t\cap\ldots\cap B_{T-1}}$$

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

$$\begin{aligned} F_t(a^m, z, x) \quad &\text{is} \quad \text{independent of } a_1^m, \ldots, a_{t-1}^m; \\ F_t(\alpha^m, Z, X) &= Z_{\tau_t^{[m]}}; \\ F_t(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}) &= Z^{(n)}_{\tau_t^{n,m,N}}. \end{aligned}$$

For $t = 2, \ldots, T$, we define also three other auxiliary functions:

$$\begin{aligned}
G_t(a^m, z, x) &= F_t(a^m, z, x) e^m_{t-1}(x); \\
\phi_t(a^m) &= \mathrm{E}[F_t(a^m, Z, X)]; \\
\psi_t(a^m) &= \mathrm{E}[G_t(a^m, Z, X)].
\end{aligned}$$

Using this notation one can see that for $t = 1, \ldots, T-1$:

$$\alpha^m_t = (A^m_t)^{-1} \psi_{t+1}(\alpha^m); \tag{5.55}$$

$$\alpha^{(m,N)}_t = (A^{(m,N)}_t)^{-1} \frac{1}{N} \sum_{n=1}^{N} G_{t+1}(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}). \tag{5.56}$$

The following estimate is a higher-dimensional counterpart of Lemma 3.1 in [48].

Lemma 5.3.4. *With the above notation, we get*

$$|F_t(a, z, x) - F_t(\tilde{a}, z, x)| \leq \sum_{s=t}^{T} |z_s| \left[\sum_{s=t}^{T-1} \mathbf{1}_{\{|z_s - e^m_s(x)\tilde{a}_s| \leq |e^m_s(x)| \|\tilde{a}_s - a_s\|\}} \right], \tag{5.57}$$

where $1 \leq t \leq T-1$, $a = (a_1, \ldots, a_{T-1}) \in \mathbb{R}^{k_m \times (T-1)}$, $\tilde{a} = (\tilde{a}_1, \ldots, \tilde{a}_{T-1}) \in \mathbb{R}^{k_m \times (T-1)}$, $z \in \mathbb{R}^T$ *and* $x \in \mathbb{R}^{d \times T}$.

Proof. Let $\tilde{B}_t = \{z_t < e^m_t(x)\tilde{a}_t\}$. Note first that

$$\begin{aligned}
|\mathbf{1}_{B_t} - \mathbf{1}_{\tilde{B}_t}| &= \mathbf{1}_{B^c_t \cap \tilde{B}_t} + \mathbf{1}_{B_t \cap \tilde{B}^c_t} \\
&\leq \mathbf{1}_{\{|z_t - e^m_t(x)\tilde{a}_t| \leq |e^m_t(x)| \|\tilde{a}_t - a_t\|\}}.
\end{aligned}$$

Moreover

$$\begin{aligned}
|\mathbf{1}_{A_1 \cap A_2} - \mathbf{1}_{C_1 \cap C_2}| &= |\mathbf{1}_{A_1}\mathbf{1}_{A_2} - \mathbf{1}_{C_1}\mathbf{1}_{C_2}| \\
&= |\mathbf{1}_{A_1}\mathbf{1}_{A_2} - \mathbf{1}_{A_1}\mathbf{1}_{C_2} + \mathbf{1}_{A_1}\mathbf{1}_{C_2} - \mathbf{1}_{C_1}\mathbf{1}_{C_2}| \\
&\leq \mathbf{1}_{A_1}|\mathbf{1}_{A_2} - \mathbf{1}_{C_2}| + \mathbf{1}_{C_2}|\mathbf{1}_{A_1} - \mathbf{1}_{C_1}| \\
&\leq |\mathbf{1}_{A_1} - \mathbf{1}_{C_1}| + |\mathbf{1}_{A_2} - \mathbf{1}_{C_2}|,
\end{aligned}$$

for any A_1, A_2, C_1, C_2. Consequently

$$\begin{aligned}
|\mathbf{1}_{B_t \cap \ldots \cap B_{s-1} \cap B^c_s} - \mathbf{1}_{\tilde{B}_t \cap \ldots \cap \tilde{B}_{s-1} \cap \tilde{B}^c_s}| &\leq \sum_{u=t}^{s-1} |\mathbf{1}_{B_u} - \mathbf{1}_{\tilde{B}_u}| + |\mathbf{1}_{B^c_s} - \mathbf{1}_{\tilde{B}^c_s}| \\
&= \sum_{u=t}^{s} |\mathbf{1}_{B_u} - \mathbf{1}_{\tilde{B}_u}|,
\end{aligned}$$

because $|\mathbf{1}_{B^c_s} - \mathbf{1}_{\tilde{B}^c_s}| = \mathbf{1}_{B^c_s \Delta \tilde{B}^c_s} = \mathbf{1}_{B_s \Delta \tilde{B}_s} = |\mathbf{1}_{B_u} - \mathbf{1}_{\tilde{B}_u}|$, where Δ denotes the symmetric difference of sets. Similarly

$$|\mathbf{1}_{B_t \cap \ldots \cap B_{T-1}} - \mathbf{1}_{\tilde{B}_t \cap \ldots \cap \tilde{B}_{T-1}}| \leq \sum_{u=t}^{T-1} |\mathbf{1}_{B_u} - \mathbf{1}_{\tilde{B}_u}|,$$

Therefore

$$\begin{aligned}|F_t(a,z,x)-F_t(\tilde{a},z,x)| &= \Big| z_t(\mathbf{1}_{B_t}-\mathbf{1}_{\tilde{B}_t}) \\ &\quad +\sum_{s=t+1}^{T-1} z_s(\mathbf{1}_{B_t\cap\ldots\cap B_{s-1}\cap B_s^c}-\mathbf{1}_{\tilde{B}_t\cap\ldots\cap\tilde{B}_{s-1}\cap\tilde{B}_s^c}) \\ &\quad +z_T(\mathbf{1}_{B_t\cap\ldots\cap B_{T-1}}-\mathbf{1}_{\tilde{B}_t\cap\ldots\cap\tilde{B}_{T-1}})\Big| \\ &\leq \left(\sum_{s=t}^{T}|z_s|\right)\left(\sum_{s=t}^{T-1}|\mathbf{1}_{B_s}-\mathbf{1}_{\tilde{B}_s}|\right),\end{aligned}$$

as needed. □

The next lemma is a direct extension of Lemma 3.2 from [48].

Lemma 5.3.5. *Let* $\mathbb{P}[e_t^m(X)\alpha_t^m = Z_t] = 0$*, for* $t \in \{1,\ldots,T-1\}$*. With the above notations and assumptions*

$$\lim_{N\to\infty} \alpha_t^{(m,N)} \overset{a.s.}{=} \alpha_t^m,$$

for $t \in \{1,\ldots,T-1\}$*.*

Proof. We know that $A_t^{(m,N)} \xrightarrow{a.s.} A_t^{(m)}$ because of the Law of Large Numbers. Hence, in view of (5.55) and (5.56), we need to prove that:

$$\frac{1}{N}\sum_{n=1}^{N} G_t(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}) \xrightarrow{a.s.} \psi_t(\alpha^{(m)}).$$

We use induction on t starting at $T-1$. For $t=T-1$, we have $G_{t+1}(a^m,z,x) = z_T e_{T-1}^m(x)$, so the statement is true as the Law of Large Numbers implies that

$$\frac{1}{N}\sum_{n=1}^{N} Z_T^{(n)} e_t^m(X^{(n)}) \xrightarrow{a.s.} \mathrm{E}[Z_T e_T^m(X)],$$

which is what we need. Assume that the statement is true for t. The Law of Large Numbers implies that

$$\frac{1}{N}\sum_{n=1}^{N} G_t(\alpha^m, Z^{(n)}, X^{(n)}) \xrightarrow{a.s.} \psi_t(\alpha^m),$$

so it suffices to prove that $\lim_{N\to\infty} G_N = 0$, where

$$G_N = \frac{1}{N}\sum_{n=1}^{N}\left(G_t(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}) - G_t(\alpha^m, Z^{(n)}, X^{(n)})\right).$$

We have

$$\begin{aligned}|G_N| &\leq \frac{1}{N}\sum_{n=1}^{N}|e_{t-1}^m(X^{(n)})||F_t(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}) - F_t(\alpha^m, Z^{(n)}, X^{(n)})| \\ &\leq \frac{1}{N}\sum_{n=1}^{N}|e_{t-1}^m(X^{(n)})|\left(\sum_{s=t}^{T}|Z_s^{(n)}|\right)\left(\sum_{s=t}^{T-1}\mathbf{1}_{W_I(s,N)}\right),\end{aligned}$$

where

$$W_I(s,N) = \{|Z_s^{(n)} - \alpha_s^m e_s^m(X^{(n)})| \le |\alpha_s^{(m,N)} - \alpha_s^m||e_s^m(X^{(n)})|\}.$$

For $s = t, \dots, T-1$

$$\alpha_s^{(m,N)} \xrightarrow{a.s.} \alpha_s^m, \qquad N \to \infty.$$

Let

$$\begin{aligned} W_{II}(s,\epsilon) &= \{\{|Z_s^{(n)} - \alpha_s^m e_s^m(X^{(n)})| \le \epsilon |e_s^m(X^{(n)})|\}\}, \\ W_{III}(s,\epsilon) &= \{|Z_s - \alpha_s^m e_s^m(X)| \le \epsilon |e_s^m(X)|\}. \end{aligned}$$

So $\forall\ \epsilon > 0$

$$\begin{aligned} \limsup |G_N| &\overset{a.s.}{\le} \limsup \frac{1}{N}\sum_{n=1}^{N}\left[|e_{t-1}^m(X^{(n)})|\left(\sum_{s=t}^{T}|Z_s^{(n)}|\right)\left(\sum_{s=t}^{T-1}\mathbf{1}_{W_{II}(s,\epsilon)}\right)\right] \\ &\overset{a.s.}{=} \mathrm{E}\left[|e_{t-1}^m(X)|\left(\sum_{s=t}^{T}|Z_s^{(n)}|\right)\left(\sum_{s=t}^{T-1}\mathbf{1}_{W_{III}(s,\epsilon)}\right)\right], \end{aligned}$$

The last equality follows from the Law of Large Numbers. If $\epsilon \to 0$, we get convergence to zero, because of the probability assumption: if $A, B, Y \ge 0$ and $\mathbb{P}(A = 0) = 0$, then as $\epsilon \searrow 0$

$$\int_{\{A \le \epsilon B\}} Y d\mathbb{P} \searrow \int_{\bigcap_{\epsilon>0}\{A \le \epsilon B\}} Y d\mathbb{P} = \int_{\{A=0\}} Y d\mathbb{P} = 0.$$

□

Finally, we are ready to present Theorem 5.3.6, which is a direct extension of Theorem 3.2 from [48].

Theorem 5.3.6. *Let $\mathbb{P}[e_t^m(X)\alpha_t^m = Z_t] = 0$, for $t \in \{1, \dots, T-1\}$. With the above notations and assumptions*

$$\frac{1}{N}\sum_{n=1}^{N} Z^{(n)}_{\tau_t^{n,m,N}} \xrightarrow{a.s.} \mathrm{E}\left[Z_{\tau_t^{[m]}}\right], \text{ as } N \to \infty,$$

for $t = 1, \dots, T$, provided that

$$\mathbb{P}(e_t^m(X)\alpha_t^m = Z_t) = 0.$$

Proof. The thesis is equivalent to the statement

$$\frac{1}{N}\sum_{n=1}^{N} F_t(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}) \xrightarrow{a.s.} \phi_t(\alpha^{(m)}).$$

As before we use induction on t starting at $T-1$. For $t = T-1$, we have $F_{t+1}(a^m, z, x) = z_T$, so the statement is true as the the Law of Large Numbers implies that

$$\frac{1}{N}\sum_{n=1}^{N} Z_T^{(n)} \xrightarrow{a.s.} \mathrm{E}[Z_T],$$

which is what we need. Assume that the statement is true for t. The Law of Large Numbers implies that

$$\frac{1}{N}\sum_{n=1}^{N} F_t(\alpha^m, Z^{(n)}, X^{(n)}) \xrightarrow{a.s.} \phi_t(\alpha^m),$$

so it suffices to prove that $\lim_{N\to\infty} F_N = 0$, where

$$F_N = \frac{1}{N}\sum_{n=1}^{N}\left(F_t(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}) - F_t(\alpha^m, Z^{(n)}, X^{(n)})\right).$$

We have

$$\begin{aligned} |F_N| &\leq \frac{1}{N}\sum_{n=1}^{N} |F_t(\alpha^{(m,N)}, Z^{(n)}, X^{(n)}) - F_t(\alpha^m, Z^{(n)}, X^{(n)})| \\ &\leq \frac{1}{N}\sum_{n=1}^{N}\left(\sum_{s=t}^{T}|Z_s^{(n)}|\right)\left(\sum_{s=t}^{T-1}\mathbf{1}_{W_I(s,N)}\right). \end{aligned}$$

Now for any $\epsilon > 0$

$$\begin{aligned} \limsup |F_N| &\overset{a.s.}{\leq} \limsup \frac{1}{N}\sum_{n=1}^{N}\left[\left(\sum_{s=t}^{T}|Z_s^{(n)}|\right)\left(\sum_{s=t}^{T-1}\mathbf{1}_{W_{II}(s,\epsilon)}\right)\right] \\ &\overset{a.s.}{=} \mathrm{E}\left[\left(\sum_{s=t}^{T}|Z_s^{(n)}|\right)\left(\sum_{s=t}^{T-1}\mathbf{1}_{W_{III}(s,\epsilon)}\right)\right], \end{aligned}$$

The last equality follows from the Law of Large Numbers. If $\epsilon \to 0$, we get convergence to zero, which is precisely what the conclusion of the theorem asserts.

$\square$

Theorems 5.3.3 and 5.3.6 provide a recipe for approximation of $\mathrm{E}[Z_{\tau_1}]$ and hence also

$$U_0 = \max\left(Z_0, \mathrm{E}[Z_{\tau_1}]\right),$$

as required.

Appendix A

Appendix

A.1 Proofs deferred to the Appendix

Proof of Proposition 2.1.3. Let $s, t \in \mathbb{T}$, $s > t$. Using the convention $0 \cdot \pm\infty = 0$ and by Beppo-Levi monotone convergence theorem for $X, Y \in \bar{L}^0$ such that $X, Y \geq 0$ and $\lambda \in L_t^0$ such that $\lambda \geq 0$ we get

$$
\begin{aligned}
E[\lambda X|\mathcal{F}_t] &= \lambda E[X|\mathcal{F}_t]; && \text{(A.1)}\\
E[X|\mathcal{F}_t] &= E[E[X|\mathcal{F}_s]|\mathcal{F}_t]; && \text{(A.2)}\\
E[X|\mathcal{F}_t] + E[Y|\mathcal{F}_t] &= E[X + Y|\mathcal{F}_t]. && \text{(A.3)}
\end{aligned}
$$

Moreover we know that for any $A \in \mathcal{F}_t$ and $X \in \bar{L}^0$ we get

$$
E[X|\mathcal{F}_t] = 1_A E[X|\mathcal{F}_t] + 1_{A^c} E[X|\mathcal{F}_t] \quad \text{and} \quad E[-X|\mathcal{F}_t] \leq -E[X|\mathcal{F}_t]. \tag{A.4}
$$

The last inequality is the result of the convention $\infty - \infty = -\infty$.[1] Now let $X, Y \in \bar{L}^0$.

1) Let $\lambda \in L_t^0$. If $\lambda \geq 0$ then using (A.1) we get

$$
\begin{aligned}
E[\lambda X|\mathcal{F}_t] &= E[(\lambda X)^+|\mathcal{F}_t] - E[(\lambda X)^-|\mathcal{F}_t] = E[\lambda X^+|\mathcal{F}_t] - E[\lambda X^-|\mathcal{F}_t] = \\
&= \lambda E[X^+|\mathcal{F}_t] - \lambda E[X^-|\mathcal{F}_t] = \lambda E[X|\mathcal{F}_t].
\end{aligned}
$$

Now for general $\lambda \in L_t^0$ using (A.4) we get

$$
\begin{aligned}
E[\lambda X|\mathcal{F}_t] &= E[1_{\{\lambda\geq 0\}}\lambda X + 1_{\{\lambda<0\}}\lambda X|\mathcal{F}_t] = 1_{\{\lambda\geq 0\}}\lambda E[X|\mathcal{F}_t] + 1_{\{\lambda<0\}}(-\lambda)E[-X|\mathcal{F}_t] \leq \\
&\leq 1_{\{\lambda\geq 0\}}\lambda E[X|\mathcal{F}_t] + 1_{\{\lambda<0\}}\lambda E[X|\mathcal{F}_t] = \lambda E[X|\mathcal{F}_t].
\end{aligned}
$$

3) On the set $\{E[X|\mathcal{F}_t] = -\infty\} \cup \{E[Y|\mathcal{F}_t] = -\infty\}$ the inequality is trivial due to the convention $\infty - \infty = -\infty$. On the other hand the set $\{E[X|\mathcal{F}_t] > -\infty\} \cap \{E[Y|\mathcal{F}_t] > -\infty\}$ could be represented as the union of the sets $\{E[X|\mathcal{F}_t] > n\} \cap \{E[Y|\mathcal{F}_t] > n\}$ for $n \in \mathbb{Z}$ on which the inequality becomes the equality.

[1] We know that $\infty \leq -(-\infty)$ and on the set $\{E[-X|\mathcal{F}_t] \neq \infty\}$ inequality in (A.4) is trivial.

2) Using 1) and 3) we get

$$\begin{aligned}E[E[X|\mathcal{F}_s]|\mathcal{F}_t] &= E[E[X^+|\mathcal{F}_s] - E[X^-|\mathcal{F}_s]|\mathcal{F}_t]\\ &\geq E[E[X^+|\mathcal{F}_s]|\mathcal{F}_t] + E[-E[X^-|\mathcal{F}_s]|\mathcal{F}_t]\\ &= E[X^+|\mathcal{F}_t] - E[X^-|\mathcal{F}_t]\\ &= E[X|\mathcal{F}_t].\end{aligned}$$

The proof for $X \in L^0$, could also be found in [44, Lemma 3.4]. □

Proof of Proposition 2.2.4. Let $\mathcal{X} = L^\infty$ and $f\colon \mathcal{X} \to \bar{L}^0$.
1) Let $X, Y \in \mathcal{X}$, $\lambda \in \mathcal{X}_t$ $(0 \leq \lambda \leq 1)$. We get

$$f(\lambda X + (1-\lambda)Y) \geq f(\lambda X) + f((1-\lambda)Y) = \lambda f(X) + (1-\lambda)f(Y).$$

2) Let $\lambda = \frac{1}{2}$. We get

$$f(X+Y) = 2\frac{1}{2}f(X+Y) = 2f(\frac{X}{2} + \frac{Y}{2}) \leq f(X) + f(Y).$$

3) From (MI), (tCA) and (tA), for any $X \in L^\infty$, we get $f(X) \in \mathcal{X}_t$, so

$$\begin{aligned}f(\lambda X + (1-\lambda)Y) &= \lambda f(X) + (1-\lambda)f(Y) + f(\lambda(X - f(X)) + (1-\lambda)(Y - f(Y))\\ &\leq \lambda f(X) + (1-\lambda)f(Y) + [f(X - f(X)) \vee f(Y - f(Y))]\\ &= \lambda f(X) + (1-\lambda)f(Y).\end{aligned}$$

4) Let $A \in \mathcal{F}_t$. We get

$$\mathbb{1}_A f(\mathbb{1}_A X) = \mathbb{1}_A f(\mathbb{1}_A X + \mathbb{1}_{A^c} \operatorname{ess\,inf} X) \leq \mathbb{1}_A f(X) \leq \mathbb{1}_A f(\mathbb{1}_A X + \mathbb{1}_{A^c} \operatorname{ess\,sup} X) = \mathbb{1}_A f(\mathbb{1}_A X).$$

5) Let $A \in \mathcal{F}_t$ and $X \in \mathcal{X}$. We get

$$\begin{aligned}&\mathbb{1}_A f(X) = \mathbb{1}_A(\mathbb{1}_A f(X) + \mathbb{1}_{A^c} f(0)) \geq \mathbb{1}_A f(\mathbb{1}_A X + \mathbb{1}_{A^c} 0) = \mathbb{1}_A f(\mathbb{1}_A X)\\ &\mathbb{1}_A f(X) = \mathbb{1}_A f(\mathbb{1}_A(\mathbb{1}_A X) + \mathbb{1}_{A^c} X) \leq \mathbb{1}_A f(\mathbb{1}_A X) + \mathbb{1}_{A \cap A^c} f(X) = \mathbb{1}_A f(\mathbb{1}_A X).\end{aligned}$$

□

Proof of Proposition 2.2.13. We will show the proof for φ^+. Let $t \in \mathbb{T}$.
(Adaptivity) It is easy to note that for any $X \in \bar{L}^0$ and $A \in \mathcal{F}_t$ we get

$$\Big[\mathbb{1}_A \operatorname*{ess\,inf}_{Y \in \mathcal{Y}_A^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c}(\infty)\Big] \in \bar{L}^0_t. \tag{A.5}$$

Indeed, for any $X \in \bar{L}^0$, ess inf of the set of $\mathcal{F}_t$-measurable random variables $\{\varphi_t(Y)\}_{Y \in \mathcal{Y}_A^+(X)}$ is $\mathcal{F}_t$-measurable (see [97, Appendix A]), which implies (A.5) for any $A \in \mathcal{F}_t$. Thus, $\varphi_t^+(X) \in \bar{L}^0_t$.
(Monotonicity) If $X \geq X'$ then for any $A \in \mathcal{F}_t$ we get $\mathcal{Y}_A^+(X) \subseteq \mathcal{Y}_A^+(X')$. Thus, for any $A \in \mathcal{F}_t$ we get

$$\mathbb{1}_A \operatorname*{ess\,inf}_{Y \in \mathcal{Y}_A^+(X)} \varphi_t(Y) \geq \mathbb{1}_A \operatorname*{ess\,inf}_{Y \in \mathcal{Y}_A^+(X')} \varphi_t(Y),$$

which implies $\varphi_t^+(X) \geq \varphi_t^+(X')$.
(Locality) Let $B \in \mathcal{F}_t$ and $X \in \bar{L}^0$. In (2.10) it is enough to consider $A \in \mathcal{F}_t$, such that $\mathcal{Y}_A^+(X) \neq \emptyset$, as otherwise we get $\varphi_t^+(X) \equiv \infty$. For any such $A \in \mathcal{F}_t$, we get

$$\mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_A^+(X)} \varphi_t(Y) = \mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(X)} \varphi_t(Y). \tag{A.6}$$

Indeed, let us assume that $\mathcal{Y}_A^+(X) \neq \emptyset$. As $\mathcal{Y}_A^+(X) \subseteq \mathcal{Y}_{A\cap B}^+(X)$, we get

$$\mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_A^+(X)} \varphi_t(Y) \geq \mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(X)} \varphi_t(Y).$$

On the other hand for any $Y \in \mathcal{Y}_{A\cap B}^+(X)$ and any fixed $Z \in \mathcal{Y}_A^+(X)$ (note that $\mathcal{Y}_A^+(X) \neq \emptyset$), we get

$$\mathbb{1}_B Y + \mathbb{1}_{B^c} Z \in \mathcal{Y}_A^+(X).$$

Thus, using locality of φ_t, we get

$$\mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(X)} \varphi_t(Y) = \mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(X)} \mathbb{1}_B\varphi_t(\mathbb{1}_B Y + \mathbb{1}_{B^c} Z) \geq \mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_A^+(X)} \varphi_t(Y),$$

which proves (A.6). Next, it is easy to see that $\mathcal{Y}_{A\cap B}^+(X) = \mathcal{Y}_{A\cap B}^+(\mathbb{1}_B X)$ and thus

$$\mathbb{1}_A \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(X)} \varphi_t(Y) = \mathbb{1}_A \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(\mathbb{1}_B X)} \varphi_t(Y). \tag{A.7}$$

Combining (A.6), (A.7) and the fact that $\mathcal{Y}_A^+(X) \neq \emptyset$ implies $\mathcal{Y}_A^+(\mathbb{1}_B X) \neq \emptyset$, we get

$$\begin{aligned}
\mathbb{1}_B\varphi_t^+(X) &= \mathbb{1}_B \operatorname*{ess\,inf}_{A\in\mathcal{F}_t} \Big[\mathbb{1}_A \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_A^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c}(\infty)\Big] \\
&= \mathbb{1}_B \operatorname*{ess\,inf}_{A\in\mathcal{F}_t} \Big[\mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_A^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c\cap B}(\infty)\Big] \\
&= \mathbb{1}_B \operatorname*{ess\,inf}_{A\in\mathcal{F}_t} \Big[\mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c\cap B}(\infty)\Big] \\
&= \mathbb{1}_B \operatorname*{ess\,inf}_{A\in\mathcal{F}_t} \Big[\mathbb{1}_{A\cap B} \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_{A\cap B}^+(\mathbb{1}_B X)} \varphi_t(Y) + \mathbb{1}_{A^c\cap B}(\infty)\Big] \\
&= \mathbb{1}_B \operatorname*{ess\,inf}_{A\in\mathcal{F}_t} \Big[\mathbb{1}_A \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_A^+(\mathbb{1}_B X)} \varphi_t(Y) + \mathbb{1}_{A^c}(\infty)\Big] \\
&= \mathbb{1}_B\varphi_t^+(\mathbb{1}_B X).
\end{aligned}$$

(Extension) If $X \in \mathcal{X}$, then for any $A \in \mathcal{F}_t$, we get $X \in \mathcal{Y}_A^+(X)$. Thus,

$$\varphi_t^+(X) = \operatorname*{ess\,inf}_{A\in\mathcal{F}_t} \Big[\mathbb{1}_A \operatorname*{ess\,inf}_{Y\in\mathcal{Y}_A^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c}(\infty)\Big] = \operatorname*{ess\,inf}_{A\in\mathcal{F}_t} \Big[\mathbb{1}_A\varphi_t(X) + \mathbb{1}_{A^c}(\infty)\Big] = \varphi_t(X).$$

As above results are true for any $t \in \mathbb{T}$, we have proved that φ^+ is an extension of φ. The proof for φ^- is analogous. Let us now show (2.12) for φ^+.

Let $\widehat{\varphi}$ be an extension of φ. Let $X \in \bar{L}^0$ and $t \in \mathbb{T}$. Due to monotonicity and locality of $\widehat{\varphi}_t$, for any $A \in \mathcal{F}_t$ and $Y \in \mathcal{Y}_A^+(X)$ we get $\mathbb{1}_A \widehat{\varphi}_t(X) \leq \mathbb{1}_A \widehat{\varphi}_t(Y)$. Thus, recalling that $\operatorname{ess\,inf} \emptyset = \infty$, we get

$$\widehat{\varphi}_t(X) \leq \mathbb{1}_A \operatorname*{ess\,inf}_{Y \in \mathcal{Y}_A^+(X)} \widehat{\varphi}_t(Y) + \mathbb{1}_{A^c}(\infty) = \mathbb{1}_A \operatorname*{ess\,inf}_{Y \in \mathcal{Y}_A^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c}(\infty). \tag{A.8}$$

As (A.8) is true for any $A \in \mathcal{F}_t$, we get

$$\widehat{\varphi}_t(X) \leq \operatorname*{ess\,inf}_{A \in \mathcal{F}_t} \Big[\mathbb{1}_A \operatorname*{ess\,inf}_{Y \in \mathcal{Y}_A^+(X)} \varphi_t(Y) + \mathbb{1}_{A^c}(\infty) \Big] = \varphi_t^+(X).$$

The proof of the second inequality is analogous. □

Proof of Proposition 2.2.15. Let $\varphi_t : L^0 \to \bar{L}_t^0$ be tUM. We will show 1) and 2) just for $\widehat{\varphi}_t$. The proof for $\bar{\varphi}_t$ is analogous, knowing that $\widehat{\varphi}_t$ satisfies 1) and 2).

1) For any $X \in \widehat{\mathcal{X}}$ and $n \in \mathbb{Z}$, we get $X \vee n \in \mathcal{X}$, so the map is properly defined. The properties (tTI) and (tIP) are always satisfied, as $\mathcal{X} \subseteq \bar{L}^0$. (tA) and (MI) follows immediately. Let us prove (tL). Let $A \in \mathcal{F}_t$. Without loss of generality, we could assume that $n < 0$. We get

$$\begin{aligned}
\mathbb{1}_A \widehat{\varphi}_t(\mathbb{1}_A X) &= \mathbb{1}_A \lim_{n \to -\infty} \varphi_t\Big((\mathbb{1}_A X) \vee n\Big) = \mathbb{1}_A \lim_{n \to \infty} \varphi_t\Big(\mathbb{1}_A (X \vee n)\Big) \\
&= \lim_{n \to -\infty} \mathbb{1}_A \varphi_t\Big(\mathbb{1}_A (X \vee n)\Big) = \lim_{n \to -\infty} \mathbb{1}_A \varphi_t\Big(X \vee n\Big) = \mathbb{1}_A \lim_{n \to -\infty} \varphi_t\Big(X \vee n\Big) \\
&= \mathbb{1}_A \widehat{\varphi}_t(X),
\end{aligned}$$

where we use appropriately the convention (2.4), if needed.

2) Assume that φ_t is (tCA) and let $X \in \widehat{\mathcal{X}}$. First, we will prove cash additivity of $\widehat{\varphi}_t$ for $m \in \mathcal{X}_t$. Without loss of generality, we could assume that $n < 0$. We know that

$$\begin{aligned}
\widehat{\varphi}_t(X + m) &= \lim_{n \to -\infty} \varphi_t\Big((X + m) \vee n\Big) = \lim_{n \to -\infty} \varphi_t\Big(X \vee (n - m) + m\Big) \\
&= \lim_{n \to -\infty} \varphi_t\Big(X \vee (n - m)\Big) + m.
\end{aligned}$$

Thus, it is enough to show that

$$\widehat{\varphi}_t(X) = \lim_{n \to -\infty} \varphi_t\Big(X \vee (n - m)\Big). \tag{A.9}$$

For any $k \in \mathbb{N}$, we have that

$$\mathbb{1}_{\{-k<m<k\}} \Big[X \vee (n - k)\Big] \leq \mathbb{1}_{\{-k<m<k\}} \Big[X \vee (n - m)\Big] \leq \mathbb{1}_{\{-k<m<k\}} \Big[X \vee (n + k)\Big].$$

Thus, using the fact that φ_t is (tL), we get

$$\mathbb{1}_{\{-k<m<k\}} \widehat{\varphi}_t(X) = \mathbb{1}_{\{-k<m<k\}} \lim_{n \to \infty} \varphi_t\Big(X \vee (n - m)\Big).$$

Since $m \in \mathcal{X}_t$ and $\mathcal{X} \subseteq L^0$, we have that $\mathbb{P}[\{-k < m < k\}] \to 1$ as $k \to \infty$ which proves the equality (A.9).

Now, let $m \in \widehat{\mathcal{X}}$. Using the above result, the fact that $\widehat{\varphi}_t$ is (tL) and because for any $k \in \mathbb{Z}$ we get $\mathbb{1}_{\{m>k\}}m \in \mathcal{X}_t$, we deduce that

$$\mathbb{1}_{\{m>-\infty\}}\widehat{\varphi}_t(X+m) = \mathbb{1}_{\{m>-\infty\}}(\widehat{\varphi}_t(X)+m).$$

On the other hand

$$\begin{aligned}\mathbb{1}_{\{m=-\infty\}}\widehat{\varphi}_t(X+m) &= \mathbb{1}_{\{m=-\infty\}}\lim_{n\to-\infty}\varphi_t((-\infty)\vee n) = \mathbb{1}_{\{m=-\infty\}}\lim_{n\to-\infty}(\varphi_t(0)+n)\\ &= \mathbb{1}_{\{m=-\infty\}}(-\infty) = \mathbb{1}_{\{m=-\infty\}}(\widehat{\varphi}_t(X)+m).\end{aligned}$$

Combining those above two equalities, (tCA) of $\widehat{\varphi}_t$ follows immediately.
3) We know that $L^\infty \subseteq \mathcal{X}$. If $X \in L^\infty$, then there exists $n, m \in \mathbb{Z}$ such that $m\wedge(X\vee n) = m\wedge X = X$ which concludes the proof both for $\widehat{\varphi}_t$ and $\bar{\varphi}_t$. Now, let $X \in \mathcal{X}$ and let us assume that φ_t satisfies (FP). Put $X_n := X \vee n$ for $n \in \mathbb{Z}$. The sequence $\{X\}_{n\in\mathbb{Z}}$ is $\mathcal{X}$-dominated by X. Moreover $X_n \xrightarrow{a.s.} X$. Hence, we have that

$$\widehat{\varphi}_t(X) = \lim_{n\to-\infty}\varphi_t(X_n) \leq \limsup_{n\to-\infty}\varphi_t(X_n) \leq \varphi_t(X) \leq \lim_{n\to-\infty}\varphi_t(X_n) = \widehat{\varphi}_t(X),$$

where the last inequality is the consequence of the fact that for any $n \in \mathbb{N}$ we have $X \leq X_n$, which implies $\varphi_t(X) \leq \varphi_t(X_n)$.

Now, let us assume that φ_t satisfies (LP). We know that (LP) implies (FP), so for $X \in \mathcal{X}$ we could write

$$\bar{\varphi}_t(X) = \lim_{m\to\infty}\varphi_t(X\wedge m).$$

Put $X_m := X \wedge m$, for $m \in \mathbb{N}$. The sequence $\{X_m\}_{m\in\mathbb{N}}$ is $\mathcal{X}$-dominated by X and $X_m \xrightarrow{a.s.} X$, which implies $\bar{\varphi}_t(X) = \varphi_t(X)$.

□

A.2 Classical definition of time-consistency

The unifying approach of time consistency for dynamic monetary risk measures, based on so called benchmark set, was suggested in [140] and used e.g. in [2, 7, 36, 57]. The reformulation for processes can be found e.g. in [1, 36, 17]. For simplicity, in this subsection (if not stated otherwise) we will assume that $\mathcal{X} = L^\infty$. Before introducing time-consistency, we need to recall the definition of benchmark set, which will define a specific subset of financial positions (test set) to which we could compare our position (see [2] for more details).

Definition A.2.1. We will call $\mathcal{Y} = \{\mathcal{Y}_t\}_{t\in\mathbb{T}}$ ($\mathcal{Y}_t \subseteq \mathcal{X}$) a *benchmark set* if for any $t \in \mathbb{T}$ we get

$$0 \in \mathcal{Y}_t \quad \text{and} \quad \mathcal{Y}_t + \mathbb{R} = \mathcal{Y}_t.$$

We are now ready to present the definition of (benchmark) time consistency.

Definition A.2.2. Let φ be dUM and let $\mathcal{Y}$ be a benchmark set. We will say that φ is *acceptance (resp. rejection) time consistent with respect to the benchmark set* $\mathcal{Y}$, if

$$\varphi_s(X) \geq \varphi_s(Y) \quad (\textit{resp.} \leq) \quad \Longrightarrow \quad \varphi_t(X) \geq \varphi_t(Y) \quad (\textit{resp.} \leq), \tag{A.10}$$

for all $X \in \mathcal{X}$ and $Y \in \mathcal{Y}_s$.

Similar definition could be obtained for the space of adapted stochastic processes with additional assumptions that values of X and Y must coincide up to time t (see [1] for details).

Intuitively, the more elements in the family $\{\mathcal{Y}_s\}_{s\in\mathbb{T}}$ the stronger the degree of time consistency of φ. Let us now introduce various types of time-consistency, which were studied in the literature:

Definition A.2.3. Let φ be dUM. We say that φ is:

- *Weakly acceptance (resp. rejection) benchmark time consistent*, if it it acceptance (resp. rejection) time consistent with respect to $\mathcal{Y} = \{\mathcal{Y}_t\}_{t\in\mathbb{T}}$, where $\mathcal{Y}_t = \mathbb{R}$.
- *Middle acceptance (resp. rejection) benchmark time consistent*, if it it acceptance (resp. rejection) time consistent with respect to $\mathcal{Y} = \{\mathcal{Y}_t\}_{t\in\mathbb{T}}$, where $\mathcal{Y}_t = \mathcal{X}_t$.
- *Strongly benchmark time consistent*, if it is acceptance (or rejection) time consistent with respect to $\mathcal{Y} = \{\mathcal{Y}_t\}_{t\in\mathbb{T}}$, where $\mathcal{Y}_t = \mathcal{X}$.[2]

In L^∞ framework, there are many equivalent reformulations for time consistency. Let us mention those, which are connected to the dynamic programming.

Proposition A.2.4. *Let φ be a monetary dRM and let $\mathcal{X} = L^\infty$. Then φ is*

- *Weakly acceptance (resp. rejection) benchmark time consistent, if and only if for $s, t \in \mathbb{T}$, $s > t$, $X \in \mathcal{X}$,*
$$\varphi_s(X) \geq 0 \quad (\textit{resp.} \leq) \quad \Longrightarrow \quad \varphi_t(X) \geq 0 \quad (\textit{resp.} \leq);$$
- *Middle acceptance (resp. rejection) benchmark time consistent, if and only if for $s, t \in \mathbb{T}$, $s > t$, $X \in \mathcal{X}$.*
$$\varphi_t(-\varphi_s(X)) \geq \varphi_t(X) \quad (\textit{resp.} \leq);$$
- *Strongly benchmark time consistent, if and only if for $s, t \in \mathbb{T}$, $s > t$, $X \in \mathcal{X}$.*
$$\varphi_t(-\varphi_s(X)) = \varphi_t(X);$$
for $s, t \in \mathbb{T}$, $s > t$, $X \in \mathcal{X}$.

A.3 Some results from convex analysis

We will present here some results from a functional (and convex) analysis. For a good general survey about L^p-spaces, definition of locally convex topological spaces, Banach spaces, Hahn-Banach theorem, etc., see [77, Appendix A.7]. Moreover, we will also present here results for a static case, i.e. we will consider maps of the form $f : \mathcal{X} \to \bar{R}$. For the $\mathcal{F}_t$-conditional equivalents of results from this section, cf. [68, 67, 82] and references therein.

Definition A.3.1. Let $\mathcal{X}$ be a topological vector space (on $\mathbb{R}$). We will call

$$\mathcal{X}^* := \{l : \mathcal{X} \to \mathbb{R} \mid l \text{ is continuous and linear}\}$$

a *(topological) dual space of $\mathcal{X}$*.

[2]Note that in this case, the inequality is symmetric, so acceptance and rejection time consistency are the same.

Remark A.3.2. In Definition A.3.1 we have assumed that $\mathcal{X}$ is embedded with a certain topology and the dual space is defined wrt. to this topology. Note that with different topologies, we could obtain different dual spaces of $\mathcal{X}$. For example L^∞ (with topology induced by $\|\cdot\|_\infty$ norm) is the dual of L^1, but the converse is generally not true. However, if we endow L^∞ with the weak* topology[3], then L^1 is the dual of L^∞ [77].

Remark A.3.3. For $p \in [1,\infty)$, the dual of L^p space embedded with $\|\cdot\|_p$ norm is the L^q space (with $\|\cdot\|_q$ norm), where q denotes the conjugate index of p (i.e. $\frac{1}{p}+\frac{1}{q}=1$).

Definition A.3.4. The *Fenchel-Legendre* transform of a function $f : \mathcal{X} \to \mathbb{R}\cup\{+\infty\}$ is the function f^* on $\mathcal{X}^*$, defined as

$$f^*(l) = \sup_{X\in\mathcal{X}} (l(X) - f(X))$$

Remark A.3.5. The function f^* is proper (P2), convex (CV) and lower semi-continuous (LSC) as the supremum of affine functions. Moreover, if f is (CV) and (P2), then we call f^* *the conjugate of* f. Note that in concave framework, for $-f$ we obtain the similar result.

Theorem A.3.6. *[Fenchel-Moreau Theorem] Let $\mathcal{X}$ be a locally convex vector space. Let f be adapted* (A)*, proper* (P2)*, convex* (CV) *and lower semi-continuous* (LSC) *wrt. weak* topology. Then $f = f^{**}$, where f^{**} denote biconjugate of f.*

Again, we refer to [77, Appendix A.7] and references therein for more classical results about dual spaces, etc. (Banach Alaoglu Theorem, James Theorem, Dunford-Pettis Theorem, etc.)

A.4 Subsidiary results

Example A.4.1 (Construction of $\widetilde{\mathcal{X}}$-extension for any tUM and dUM defined on $\mathcal{X}$)**.** Let $\mathcal{X}, \widetilde{\mathcal{X}} \subseteq \bar{L}^0$ be such that $\mathcal{X} \subset \widetilde{\mathcal{X}}$. Let $\varphi_t : \mathcal{X} \to \bar{L}^0$ be tUM. We will show, how to construct the exemplary $\bar{L}^0$-extension of φ_t.

Proof. We will show, how, for a given $\varphi_t : \mathcal{X} \to \bar{L}^0_t$ tUM, construct the exemplary $\bar{L}^0$-extension of φ_t (note, that such extension will be valid for any $\widetilde{\mathcal{X}} \subseteq \bar{L}^0$). In fact we will show how to construct two $\bar{L}^0$-extensions, which usually do not coincide. For $X \in \bar{L}^0$, let

$$\mathcal{A}_t(X) := \{A \in \mathcal{F}_t \mid \ \mathbb{1}_A X \in \mathcal{X},\ \mathbb{P}[A] > 0\}.$$

First, let us show that If $\mathcal{A}_t(X) \neq \emptyset$, then for any $X \in \mathcal{X}$ and $n \in \mathbb{N}$, there exists $A_n \in \mathcal{A}_t$, such that

$$\sup_{A\in\mathcal{A}_t(X)} \mathbb{P}[A \setminus A_n] \leq \frac{1}{n}. \tag{A.11}$$

On the contrary let us assume that (A.11) is not true for some $n \in \mathbb{N}$. Let $A^0_n \in \mathcal{A}_t$ (any element from $\mathcal{A}_t$). There exists $A \in \mathcal{A}_t(X)$ such that $\mathbb{P}[A \setminus A^0_n] > \frac{1}{n}$. We put $A^1_n = A^0_n \cup A$. It is easy to note that $A^1_n \in \mathcal{A}_t(X)$. Again, There exists $A' \in \mathcal{A}_t$ such that $\mathbb{P}[A' \setminus A^1_n] > \frac{1}{n}$. We put $A^2_n = A^1_n \cup A'$

[3] i.e. $\sigma(L^\infty, L^1)$, the initial topology with respect to the dual space $\mathcal{X}^*$ see [77] for details

and again note that $A_n^2 \in \mathcal{A}_t(X)$. Doing similar operations we could obtain the set $A_n^{n+1} \in \mathcal{A}_t(X)$. Because the family $\{A_n^i\}_{i=1}^{n+1}$ is ascending, we get

$$\mathbb{P}[A_n^{n+1}] = \mathbb{P}[A_n^{n+1} \setminus A_n^n] + \mathbb{P}[A_n^n] = \ldots = \sum_{i=1}^{n+1} \mathbb{P}[A_n^i \setminus A_n^{i-1}] + \mathbb{P}[A_n^0] > 1.$$

This contradicts the assumption that P is a probability measure. Thus, (A.11) must by true.

Now let $\{A_n\}_{n\in\mathbb{N}}$ be a family of sets $A_n \in \mathcal{A}_t(X)$ satisfying (A.11). It is easy to see that

$$\sup_{A\in\mathcal{A}_t(X)} \mathbb{P}[A \setminus \bigcup_{n\in\mathbb{N}} A_n] = 0. \tag{A.12}$$

Let $A_1' := A_1$ and $A_n' := (A_n \setminus \bigcup_{m\leq n-1} A_m)$ for $n \geq 2$. Let $K = \{n \in \mathbb{N} \mid \mathbb{P}[A_n'] > 0\}$. For any $k \in K$, we get $A_k \in \mathcal{A}_t(X)$ and for any $k_1, k_2 \in K$, such that $k_1 \neq k_2$, we get $\mathbb{P}[A_{k_1} \cap A_{k_2}] = 0$. Moreover, if $\mathcal{A}_t(X) \neq \emptyset$, then $1 \in K$ and consequently, using (A.12), we get

$$\sup_{A\in\mathcal{A}_t(X)} \mathbb{P}[A \setminus \bigcup_{k\in K} A_k'] = 0. \tag{A.13}$$

Thus, we have shown that for any $X \in \bar{L}^0$, such that $\mathcal{A}_t(X) \neq \emptyset$, there exists a family $\{A_i^X\}_{i=1}^{N_X}$, where $N_X \in \mathbb{N} \cup \{\infty\}$, such that for $i, j \in \mathbb{N}$, $i < j \leq N_X$, we get

$$A_i^X \in \mathcal{A}_t(X), \quad A_i^X \cap A_j^X = \emptyset, \quad \sup_{A\in\mathcal{A}_t(X)} \mathbb{P}[A \setminus \bigcup_{k=1}^{N_X} A_k^X] = 0. \tag{A.14}$$

Moreover, for $X \in \bar{L}^0$, such that $\mathcal{A}_t(X) = \emptyset$, we put $N_X = 1$ and $A_1^X = \emptyset$. Let

$$B_t^X := \bigcup_{i=1}^{N_X} A_i^X. \tag{A.15}$$

We know that for any $X \in \bar{L}^0$, the set B_t^X is $\mathcal{F}_t$ measurable (note that we might get $B_t^X \notin \mathcal{A}_t$). Let $\varphi_t^1 : \bar{L}^0 \to \bar{L}_t^0$ and $\varphi_t^2 : \bar{L}^0 \to \bar{L}_t^0$ be defined by

$$\varphi_t^1(X)(\omega) = \begin{cases} \sum_{i=1}^{N_X} \mathbb{1}_{A_i^X} \varphi(\mathbb{1}_{A_i^X} X)(\omega) & \text{if } \omega \in B_t^X \\ -\infty & \text{if } \omega \in B_t^{X^+} \setminus B_t^X, \\ \infty & \text{otherwise,} \end{cases} \tag{A.16}$$

$$\varphi_t^2(X)(\omega) = \begin{cases} \sum_{i=1}^{N_X} \mathbb{1}_{A_i^X} \varphi(\mathbb{1}_{A_i^X} X)(\omega) & \text{if } \omega \in B_t^X \\ \infty & \text{if } \omega \in B_t^{X^-} \setminus B_t^X, \\ -\infty & \text{otherwise,} \end{cases} \tag{A.17}$$

Let us show that φ^1 is an $\bar{L}^0$-extension of φ.

(Monotonicity) Let $t \in \mathbb{T}$ and $X, Y \in \bar{L}^0$, be such that $X \geq Y$. We will prove that

$$\varphi_t^1(X) \geq \varphi_t^1(Y) \tag{A.18}$$

First of all, it is easy to see[4], that for any $Z_1, Z_2 \in \bar{L}^0$

$$\mathcal{A}_t(Z_1) \subseteq \mathcal{A}_t(Z_2) \Rightarrow B_t^{Z_1} \subseteq B_t^{Z_2} \tag{A.19}$$

Indeed, using (A.15), if $\mathcal{A}_t(Z_1) \subseteq \mathcal{A}_t(Z_2)$, then for $i = 1, 2, \ldots, N_{Z_1}$, we get $A_i^{Z_1} \in \mathcal{A}_t(Z_2)$. Now, by (A.14) we get $\mathbb{P}[A_i^{Z_1} \setminus B_t^{Z_2}] = 0$ for any $i = 1, 2, \ldots, N_{Z_1}$, so $\mathbb{P}[B_t^{Z_1} \setminus B_t^{Z_2}] = 0$.

Due to locality of φ we get (A.18) on $B_t^X \cap B_t^Y$. Next, it is easy to see that $\mathcal{A}_t(X) \subseteq \mathcal{A}_t(X^+)$ (note that L^p is the Frechet lattice [13] for $p \in \{0, 1, \infty\}$ and $0 \in \mathcal{X}$). Thus, using (A.19), we get $B_t^X \subseteq B_t^{X^+}$. Moreover, using the same arguments, we get $B_t^{X^+} \subseteq B_t^{Y^+}$, and consequently $B_t^X \subseteq B_t^{Y^+}$. Because of that, and the fact that inequality (A.18) is trivial on the set $B_t^{Y^+} \setminus B_t^Y$, we get (A.18) on $B_t^X \setminus B_t^Y$, and consequently on B_t^X.

Next, is is easy to note, that for any $i \leq N^{X^+}$ and $j \leq N^X$, using the fact that $X \geq Y$, we get

$$\mathbb{1}_{A_i^{X^+} \setminus A_j^X} X \notin \mathcal{X} \Rightarrow \mathbb{1}_{A_i^{X^+} \setminus A_j^X} Y \notin \mathcal{X}.$$

Thus, $B_t^{X^+} \setminus B_t^X \subseteq B_t^{Y^+} \setminus B_t^Y$, and consequently (A.18) is true on $B_t^{X^+}$.

Combining it with the fact, that (A.18) is trivial on $\Omega \setminus B_t^{X^+}$, we obtain monotonicity.

(Locality) Let $t \in \mathbb{T}$, $A \in \mathcal{F}_t$ and $X \in \bar{L}^0$. It is easy to note that $\mathcal{A}_t(X) \subseteq \mathcal{A}_t(\mathbb{1}_A X)$. Using (A.19), we get $B_t^X \subseteq B_t^{\mathbb{1}_A X}$. On the other hand if only $\mathbb{P}[A \cap A_i^{\mathbb{1}_A X}] > 0$, then $A \cap A_i^{\mathbb{1}_A X} \in \mathcal{A}_t(X)$, so we could state that $A \cap A_i^{\mathbb{1}_A X} \subseteq B_t(X)$. Combining those observations, we get

$$A \cap B_t^X = \bigcup_{i=1}^{N_X} A \cap A_i^X = \bigcup_{j=1}^{N_{(\mathbb{1}_A X)}} \bigcup_{i=1}^{N_X} A \cap A_i^X \cap A_j^{\mathbb{1}_A X} = \bigcup_{j=1}^{N_{(\mathbb{1}_A X)}} A \cap A_j^{\mathbb{1}_A X} = A \cap B_t^{\mathbb{1}_A X} \tag{A.20}$$

Thus, using the fact that φ is local and (A.20), we get

$$\begin{aligned}
\mathbb{1}_{A \cap B_t^X} \varphi_t^1(X) &= \sum_{i=1}^{N_X} \mathbb{1}_{A \cap A_i^X} \varphi(\mathbb{1}_{A_i^X} X) = \sum_{j=1}^{N_{(\mathbb{1}_A X)}} \sum_{i=1}^{N_X} \mathbb{1}_{A \cap A_i^X \cap A_j^{(\mathbb{1}_A X)}} \varphi(\mathbb{1}_{A_i^X} X) \\
&= \sum_{j=1}^{N_{(\mathbb{1}_A X)}} \mathbb{1}_{A \cap A_j^{(\mathbb{1}_A X)}} \varphi(\mathbb{1}_{A_j^{(\mathbb{1}_A X)}} X) = \mathbb{1}_{A \cap B_t^X} \varphi_t^1(\mathbb{1}_A X).
\end{aligned} \tag{A.21}$$

We also get

$$\begin{aligned}
\mathbb{1}_{A \cap (\Omega \setminus B_t^X)} \varphi_t^1(X) &= \mathbb{1}_{A \cap (B_t^{X^+} \setminus B_t^X)}(-\infty) + \mathbb{1}_{A \cap (\Omega \setminus B_t^{X^+})}(\infty) \\
&= \mathbb{1}_{A \cap (B_t^{(\mathbb{1}_A X)^+} \setminus B_t^{(\mathbb{1}_A X)})}(-\infty) + \mathbb{1}_{A \cap (\Omega \setminus B_t^{(\mathbb{1}_A X)^+})}(\infty) \\
&= \mathbb{1}_{A \cap (\Omega \setminus B_t^X)} \varphi_t^1(\mathbb{1}_A X).
\end{aligned} \tag{A.22}$$

Combining (A.21) and (A.22), we obtain locality.

(Extension) Of course if $X \in L^p$, then we get $B_t^X = \Omega$ and in fact it is enough to consider $A_1^X = \Omega$. Thus, we get $\varphi_t^1(X) = \varphi_t(1_\Omega X) = \varphi_t(X)$. This concludes the proof, that φ^1 is the extension of φ. The proof for φ^2 is analogous.

$\square$

[4] We shall write $A \subseteq B$, if $\mathbb{P}[A \setminus B] = 0$.

Bibliography

[1] B. Acciaio, H. Föllmer, and I. Penner, Risk assessment for uncertain cash flows: model ambiguity, discounting ambiguity, and the role of bubbles, Finance and Stochastics **16** (2012), no. 4, 669–709 (English).

[2] B. Acciaio and I. Penner, Dynamic risk measures, In G.Di Nunno and B. Öksendal (Eds.) Advanced Mathematical Methods for Finance (2011), 1–34.

[3] B. Acciaio and V. Goldammer, Optimal portfolio selection via conditional convex risk measures on l p, Decisions in Economics and Finance **36** (2013), no. 1, 1–21.

[4] C. Acerbi, Spectral measures of risk: a coherent representation of subjective risk aversion, Journal of Banking & Finance **26** (2002), no. 7, 1505–1518.

[5] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, and S. I. Marcus, Discrete-time controlled markov processes with average cost criterion: a survey, SIAM Journal on Control and Optimization **31** (1993), no. 2, 282–344.

[6] P. Artzner, F. Delbaen, J. Eber, and D. Heath, Thinking coherently, Risk **10** (1997), 68–71.

[7] P. Artzner, F. Delbaen, J. Eber, D. Heath, and H. Ku, Coherent multiperiod risk adjusted values and Bellman's principle, Annals of Operations Research **152** (2007), 5–22.

[8] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, Coherent measures of risk, Math. Finance **9** (1999), no. 3, 203–228.

[9] C. R. Bacon, Practical portfolio performance measurement and attribution, vol. 568, John Wiley & Sons, 2011.

[10] E. N. Barron, P. Cardaliaguet, and R. Jensen, Conditional essential suprema with applications, Appl. Math. Optim. **48** (2003), no. 3, 229–253.

[11] L. Bauwens, S. Laurent, and J. V. Rombouts, Multivariate garch models: a survey, Journal of applied econometrics **21** (2006), no. 1, 79–109.

[12] D. P. Bertsekas and S. E. Shreve, Stochastic optimal control: The discrete time case, vol. 139, Academic Press New York, 1978.

[13] S. Biagini and M. Frittelli, On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures, Optimality and Risk - Modern Trends in Mathematical Finance, Springer Berlin Heidelberg, 2010, pp. 1–28 (English).

[14] S. Biagini and J. Bion-Nadal, Dynamic quasi-concave performance measures, arXiv preprint arXiv:1212.3958 (2012).

[15] ______, Dynamic quasi concave performance measures, Journal of Mathematical Economics (2014).

[16] T. R. Bielecki, I. Cialenco, and T. Chen, Dynamic conic finance via backward stochastic difference equations, Working Paper (2014).

[17] T. R. Bielecki, I. Cialenco, S. Drapeau, and M. Karliczek, Dynamic Assessment Indices, ArXiv e-prints (2013).

[18] T. R. Bielecki, I. Cialenco, and M. Pitera, Dynamic limit growth indices in discrete time, arXiv preprint arXiv:1312.1006 (2013).

[19] ______, A unified approach to time consistency of dynamic risk measures and dynamic performance measures in discrete time, Working Paper (2014).

[20] T. R. Bielecki, I. Cialenco, and Z. Zhang, Dynamic coherent acceptability indices and their applications to finance, Mathematical Finance (2012).

[21] T. R. Bielecki and S. R. Pliska, Risk-sensitive dynamic asset management, Appl. Math. Optim. **39** (1999), no. 3, 337–360.

[22] ______, Economic properties of the risk sensitive criterion for portfolio management, Review of Accounting and Finance **2** (2003), 3–17.

[23] T. 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** (2013), no. 01, 1350002.

[24] J. Bion-Nadal, Conditional risk measure and robust representation of convex conditional risk measures, CMAP preprint # 557 (2004).

[25] ______, Time consistent dynamic risk processes, Stochastic Process. Appl. **119** (2009), no. 2, 633–654.

[26] T. Bj́ork and A. Murgoci, A theory of markovian time-inconsistent stochastic control in discrete time, Finance and Stochastics **18** (2014), no. 3, 545–592.

[27] M. Broadie and P. Glasserman, Pricing american-style securities using simulation, Journal of Economic Dynamics and Control **21** (1997), 1323–1352.

[28] ______, A stochastic mesh method for pricing high-dimensional american options, Tech. report, Columbia Business School, New York, 2002, A revised version of a working paper from 1997.

[29] J. Y. Campbell, The econometrics of financial markets, princeton University press, 1997.

[30] J. F. Carriere, Valuation of early-exercise price for options using simulations and nonparametric regression, Insurance: Mathematics and Economics **19** (1996), 19–30.

[31] R. Cavazos-Cadena and D. Hernández-Hernández, A characterization of the optimal risk-sensitive average cost in finite controlled markov chains, The Annals of Applied Probability **15** (2005), no. 1A, 175–212.

[32] Z. Chen, T. Chen, and M. Davison, Choquet expectation and peng's g-expectation, The Annals of Probability **33** (2005), no. 3, 1179–1199.

[33] Z. Chen, K. He, and R. Kulperger, Risk measures and nonlinear expectations., Journal of Mathematical Finance **3** (2013), no. 3.

[34] Z. Chen, R. Kulperger, and L. Jiang, Jensen's inequality for g-expectation: part 1, Comptes Rendus Mathematique **337** (2003), no. 11, 725–730.

[35] P. Cheridito, F. Delbaen, and M. Kupper, Coherent and convex monetary risk measures for bounded càdlàg processes, Stochastic Process. Appl. **112** (2004), no. 1, 1–22.

[36] ______, Dynamic monetary risk measures for bounded discrete-time processes, Electron. J. Probab. **11** (2006), no. 3, 57–106.

[37] P. Cheridito and T. Li, Risk measures on Orlicz hearts, Math. Finance **19** (2009), no. 2, 189–214.

[38] P. Cheridito, F. Delbaen, and M. Kupper, Coherent and convex monetary risk measures for unbounded cadlag processes, Finance and Stochastics **9** (2005), no. 3, 369–387.

[39] P. Cheridito and E. Kromer, Reward-risk ratios, Journal of Investment Strategies (2012).

[40] P. Cheridito and M. Kupper, Composition of time-consistent dynamic monetary risk measures in discrete time, International Journal of Theoretical and Applied Finance **14** (2011), no. 01, 137–162.

[41] A. S. Cherny, Weighted V@R and its properties, Finance Stoch. **10** (2006), no. 3, 367–393.

[42] ______, Pricing with coherent risk, Probability Theory and Its Applications **52** (2007), no. 3, 506–540.

[43] ______, Capital allocation and risk contribution with discrete-time coherent risk, Math. Finance **19** (2009), no. 1, 13–40. MR 2488400 (2010c:91068)

[44] ______, Risk-reward optimization with discrete-time coherent risk, Mathematical Finance **20** (2010), no. 4, 571–595.

[45] A. S. Cherny and D. B. Madan, New measures for performance evaluation, The Review of Financial Studies **22** (2009), no. 7, 2571–2606.

[46] A. S. Cherny and V. P. Maslov, On minimization and maximization of entropy in various disciplines, Theory of Probability & Its Applications **48** (2004), no. 3, 447–464.

[47] T. S. Chihara, An introduction to orthogonal polynomials, Gordon and Breach, New York, 1978.

[48] E. Clément, D. Lamberton, and P. Protter, An analysis of a least squares regression method for american option pricing, Finance and Stochastics **6** (2002), 449–471.

[49] F. Coquet, Y. Hu, J. Mémin, and S. Peng, Filtration-consistent nonlinear expectations and related g-expectations, Probability Theory and Related Fields **123** (2002), no. 1, 1–27.

[50] J. Cvitanic and I. Karatzas, On dynamic meaures of risk, Finance Stoch. **3** (1999), no. 4, 451–482.

[51] P. Dai Pra, L. Meneghini, and W. J. Runggaldier, Connections between stochastic control and dynamic games, Mathematics of Control, Signals and Systems **9** (1996), no. 4, 303–326.

[52] R.-A. Dana, A representation result for concave schur concave functions, Mathematical Finance **15** (2005), no. 4, 613–634.

[53] F. Delbaen, The structure of m-stable sets and in particular of the set of risk neutral measures, In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX, Lecture Notes in Math., vol. 1874, Springer, Berlin, 2006, pp. 215–258.

[54] F. Delbaen, S. Peng, and E. Rosazza Gianin, Representation of the penalty term of dynamic concave utilities, Finance Stoch. **14** (2010), 449–472, 10.1007/s00780-009-0119-7.

[55] F. Delbaen, Coherent risk measures, Blätter der DGVFM **24** (2000), 733–739, 10.1007/BF02809088.

[56] ______, Coherent risk measures on general probability spaces, Advances in finance and stochastics, Springer, 2002, pp. 1–37.

[57] K. Detlefsen and G. Scandolo, Conditional and dynamic convex risk measures, Finance Stoch. **9** (2005), no. 4, 539–561.

[58] G. B. Di Masi and Ł. Stettner, Risk-sensitive control of discrete-time markov processes with infinite horizon, SIAM Journal on Control and Optimization **38** (1999), no. 1, 61–78.

[59] ______, On additive and multiplicative (controlled) poisson equations, BANACH CENTER PUBLICATIONS **72** (2006), 57.

[60] ______, Remarks on risk neutral and risk sensitive portfolio optimization, From Stochastic Calculus to Mathematical Finance, Springer, 2006, pp. 211–226.

[61] ______, Infinite horizon risk sensitive control of discrete time markov processes under minorization property, SIAM Journal on Control and Optimization **46** (2007), no. 1, 231–252.

[62] R. L. Dobrushin and R. A. Minlos, Polynomials in linear random functions, Russian Math. Surveys **32:2** (1977), 71–127.

[63] S. Drapeau and M. Kupper, Risk preferences and their robust representation, Mathematics of Operations Research **38** (2013), no. 1, 28–62.

[64] D. Egloff, Monte Carlo algorithms for optimal stopping and statistical learning, Ann. Appl. Probab. **15** (2005), no. 2, 1396–1432.

[65] E. F. Fama and K. R. French, Multifactor explanations of asset pricing anomalies, The journal of finance **51** (1996), no. 1, 55–84.

[66] B. J. Feibel, Investment performance measurement, vol. 116, John Wiley & Sons, 2003.

[67] D. Filipović, M. Kupper, and N. Vogelpoth, Separation and duality in locally l0-convex modules, Journal of Functional Analysis **256** (2009), no. 12, 3996 – 4029.

[68] ______, Approaches to conditional risk, SIAM Journal on Financial Mathematics **3** (2012), no. 1, 402–432.

[69] D. Filipović and G. Svindland, The canonical model space for law-invariant convex risk measures is l1, Mathematical Finance **22** (2012), no. 3, 585–589.

[70] P. C. Fishburn, Utility theory, Wiley Online Library, 1988.

[71] W. H. Fleming and S. J. Sheu, Risk-sensitive control and an optimal investment model, Mathematical Finance **10** (2000), no. 2, 197–213.

[72] W. H. Fleming and D. Hernández-Hernández, Risk-sensitive control of finite state machines on an infinite horizon i, SIAM Journal on Control and Optimization **35** (1997), no. 5, 1790–1810.

[73] H. Föllmer and I. Penner, Convex risk measures and the dynamics of their penalty functions, Statist. Decisions **24** (2006), no. 1, 61–96.

[74] H. Föllmer, Spatial risk measures and their local specification: The locally law-invariant case, Statistics & Risk Modeling **31** (2014), no. 1, 79–101.

[75] H. Föllmer and T. Knispel, Convex risk measures: Basic facts, law-invariance and beyond, asymptotics for large portfolios, Handbook of the Fundamentals of Financial Decision Making, Eds. LC MacLean and WT Ziemba, World Scientific (2013).

[76] H. Föllmer and A. Schied, Convex measures of risk and trading constraints, Finance and stochastics **6** (2002), no. 4, 429–447.

[77] ______, Stochastic finance. an introduction in discrete time, extended ed., de Gruyter Studies in Mathematics, vol. 27, Walter de Gruyter & Co., Berlin, 2004.

[78] ______, Convex risk measures, Encyclopedia of Quantitative Finance (2010).

[79] H. Föllmer, A. Schied, and S. Weber, Robust preferences and robust portfolio choice, Handbook of Numerical Analysis **15** (2009), 29–87.

[80] M. Frittelli and E. Rosazza Gianin, Dynamic convex measures, Risk Measures in 21st Century, G. Szegö ed., J. Wiley, 2004, pp. 227–248.

[81] M. Frittelli and G. Scandolo, Risk measures and capital requirements for processes, Math. Finance **16** (2006), no. 4, 589–612.

[82] M. Frittelli and M. Maggis, Dual representation of quasi-convex conditional maps, SIAM Journal on Financial Mathematics **2** (2011), no. 1, 357–382.

[83] ______, Complete duality for quasiconvex dynamic risk measures on modules of the lp-type, Statistics & Risk Modeling **31** (2014), no. 1, 103–128.

[84] M. Frittelli and E. Rosazza Gianin, Law invariant convex risk measures, Advances in Mathematical Economics, Springer, 2005, pp. 33–46.

[85] H. U. Gerber, An introduction to mathematical risk theory, vol. 8, SS Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania Philadelphia, 1979.

[86] P. Glasserman, Monte Carlo methods in financial engineering, Springer, New York, 2010.

[87] P. Glasserman and B. Yu, Number of paths versus number of basis functions in American option pricing, Ann. Appl. Probab. **14** (2004), no. 4, 2090–2119.

[88] T. Guo, S. Zhao, and X. Zeng, On random convex analysis–the analytic foundation of the module approach to conditional risk measures, arXiv preprint arXiv:1210.1848 (2012).

[89] D. Hernández-Hernández and S. I. Marcus, Risk sensitive control of markov processes in countable state space, Systems & control letters **29** (1996), no. 3, 147–155.

[90] O. Hernández-Lerma, Adaptive markov control processes, Springer, 1989.

[91] O. Hernández-Lerma and J. B. Lasserre, Discrete-time markov control processes, Springer, 1996.

[92] A. Jobert and L. Rogers, Valuations and dynamic convex risk measures, Math. Finance **18** (2008), no. 1, 1–22.

[93] P. Jorion, Value at risk: the new benchmark for managing financial risk, vol. 2, McGraw-Hill New York, 2007.

[94] E. Jouini, W. Schachermayer, and N. Touzi, Law invariant risk measures have the fatou property, Advances in mathematical economics, Springer, 2006, pp. 49–71.

[95] M. Kaina and L. Rüschendorf, On convex risk measures on l p-spaces, Mathematical methods of operations research **69** (2009), no. 3, 475–495.

[96] O. Kallenberg, Foundations of Modern Probability, Springer, 2002.

[97] I. Karatzas and S. E. Shreve, Methods of mathematical finance, vol. 39, Springer, 1998.

[98] M. Klimek, M. Matsuura, and Y. Okabe, Stochastic flows and finite block frames, Journal of Mathematical Analysis and Applications **342** (2008), no. 2, 816–829.

[99] M. Klimek and M. Pitera, The least squares method for option pricing revisited, arXiv preprint arXiv:1404.7438 (2014).

[100] P. Koch-Medina and C. Munari, Law-invariant risk measures: extension properties and qualitative robustness, arXiv preprint arXiv:1401.3121 (2014).

[101] I. Kontoyiannis and S. P. Meyn, Spectral theory and limit theorems for geometrically ergodic markov processes, Annals of Applied Probability (2003), 304–362.

[102] D. M. Kreps, Notes on the theory of choice, vol. 2, Westview press Boulder, 1988.

[103] M. Kupper and W. Schachermayer, Representation results for law invariant time consistent functions, Mathematics and Financial Economics **2** (2009), no. 3, 189–210.

[104] D. Lamberton and B. Lapeyre, Introduction to stochastic calculus applied to finance, CRC press, 2007.

[105] G. Liu and L. J. Hong, Revisit of stochastic mesh method for pricing american options, Operations Research Letters **37** (2009), no. 6, 411–414.

[106] F. A. Longstaff and E. S. Schwartz, Valuing american options by simulation: a simple least-squares approach, Review of Financial Studies **14** (2001), 113–148.

[107] M. Maggis, On quasiconvex conditional maps. duality results and applications to finance, Ledizioni, 2011.

[108] H. Markowitz, Portfolio selection, The journal of finance **7** (1952), no. 1, 77–91.

[109] E. Mastrogiacomo and E. Rosazza Gianin, Portfolio optimization with quasiconvex risk measures, Available at SSRN 2282472 (2013).

[110] A. J. McNeil, R. Frey, and P. Embrechts, Quantitative risk management: concepts, techniques, and tools, Princeton university press, 2010.

[111] R. C. Merton, An intertemporal capital asset pricing model, Econometrica: Journal of the Econometric Society (1973), 867–887.

[112] J. Orihuela and M. R. Galán, Lebesgue property for convex risk measures on orlicz spaces, Mathematics and Financial Economics **6** (2012), no. 1, 15–35.

[113] S. Peng, Nonlinear expectations and risk measures, Proceedings of the CIME-EMS summer school, Bressanone, Italy (2003).

[114] ———, Nonlinear expectations, nonlinear evaluations and risk measures, Stochastic methods in finance, Springer, 2004, pp. 165–253.

[115] ———, G-expectation, g-brownian motion and related stochastic calculus of itô type, Stochastic analysis and applications, Springer, 2007, pp. 541–567.

[116] I. Penner, Dynamic convex risk measures: Time consistency, prudence, and sustainability, Ph.D. thesis, Humboldt Universität zu Berlin, 2007.

[117] J. Penot, Characterization of solution sets of quasiconvex programs, Journal of optimization theory and applications **117** (2003), no. 3, 627–636.

[118] G. Peskir and A. Shiryaev, Optimal stopping and free-boundary problems, Springer, 2006.

[119] H. Pham, On some recent aspects of stochastic control and their applications, Probability Surveys **2** (2005), no. 506-549, 1.

[120] J.-L. Prigent, Portfolio optimization and performance analysis, CRC Press, 2007.

[121] F. Riedel, Dynamic coherent risk measures, Stochastic Process. Appl. **112** (2004), no. 2, 185–200.

[122] R. T. Rockafellar, Convex analysis, vol. 28, Princeton university press, 1997.

[123] V. A. Rokhlin, On the fundamental ideas of measure theory, Matematicheskii Sbornik **67** (1949), no. 1, 107–150.

[124] B. Roorda, J. M. Schumacher, and J. Engwerda, Coherent acceptability measures in multiperiod models, Mathematical Finance **15** (2005), no. 4, 589–612.

[125] E. Rosazza Gianin, Risk measures via g-expectations, Insurance: Mathematics and Economics **39** (2006), no. 1, 19–34.

[126] E. Rosazza Gianin and C. Sgarra, Acceptability indexes via g-expectations: an application to liquidity risk, Mathematics and financial economics **7** (2013), no. 4, 457–475.

[127] D. J. Rudolph, Fundamentals of measurable dynamics: Ergodic theory on lebesgue spaces, Clarendon Press Oxford, 1990.

[128] J. Rust, Using randomization to break the curse of dimensionality, Econometrica **65** (1997), no. 3, 487–516.

[129] A. Ruszczyński, Risk-averse dynamic programming for markov decision processes, Mathematical programming **125** (2010), no. 2, 235–261.

[130] A. Ruszczyński and A. Shapiro, Conditional risk mappings, Mathematics of Operations Research **31** (2006), no. 3, 544–561.

[131] ———, Optimization of convex risk functions, Mathematics of operations research **31** (2006), no. 3, 433–452.

[132] A. Shapiro, On a time consistency concept in risk averse multistage stochastic programming, Operations Research Letters **37** (2009), no. 3, 143–147.

[133] Y. Shen, W. Stannat, and K. Obermayer, Risk-sensitive markov control processes, SIAM Journal on Control and Optimization **51** (2013), no. 5, 3652–3672.

[134] C. G. Small and D. L. McLeish, Hilbert space methods in probability and statistical inference, Wiley, New York, 1994.

[135] L. Stentoft, Convergence of the least squares Monte Carlo approach to american option valuation, Management Sci. **50** (2004), no. 9, 1193–1203.

[136] Ł. Stettner, Risk sensitive portfolio optimization, Mathematical Methods of Operations Research **50** (1999), no. 3, 463–474.

[137] ______, Duality and risk sensitive portfolio optimization, Contemporary Mathematics **351** (2004), 333–348.

[138] J. N. Tsitsiklis and B. Van Roy, Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives, IEEE Trans. Automat. Control **44** (1999), no. 10, 1840–1851.

[139] ______, Regression methods for pricing complex American-style options, IEEE Trans. on Neutral Networks **12** (2001), no. 4, 694–703.

[140] S. Tutsch, Update rules for convex risk measures, Quantitative Finance **8** (2008), no. 8, 833–843.

[141] S. Weber, Distribution-invariant risk measures, entropy, and large deviations, J. Appl. Probab. **44** (2007), no. 1, 16–40.

[142] P. Whittle, Risk-sensitive optimal control, Wiley New York, 1990.

[143] C. Zalinescu, Convex analysis in general vector spaces, World Scientific, 2002.

[144] D. Z. Zanger, Convergence of a least-squares Monte Carlo algorithm for bounded approximating sets, Appl. Math. Finance **16** (2009), no. 1-2, 123–150.

[145] J. M. Zapata-García, On characterization of locally l0-convex topologies induced by a family of l0-seminorms, Available at SSRN (2014).

List of symbols

(t·), t·	The prefix t denotes the $\mathcal{F}_t$-conditional version of the property, family, etc. (see e.g. page 10)
(d·), d·	The prefix d denotes the dynamic version of the property, family, etc. (see e.g. page 14)

Acronyms - properties

(A)	10	Adapted
(SBA)	10	Subadditive
(SPA)	10	Superadditive
(AD)	10	Additive
(N)	10	Normalized
(M)	10	Monotone
(MI)	10	Monotone increasing
(MD)	10	Monotone decreasing
(P)	10	Proper
(P1)	10	Proper (in concave framework)
(P2)	10	Proper (in convex framework)
(F)	10	Finite
(L)	10	Local
(IP)	10	Independent of the past
(TI)	10	Translation invariant
(CA)	10	Cash additive
(CCA)	10	Counter cash additive
(QCC)	10	Quasi-concave
(CC)	10	Concave
(QCV)	10	Quasi-convex
(CV)	10	Convex
(SI)	10	Scale invariant
(PH)	10	Positively homogeneous
(LP)	12	Lebesgue Property
(FP)	12	Fatou property
(LI)	12	Law-invariant
(USC)	10	Upper semi-continous
(LSC)	10	Lower semi-continous

Acronyms - families of maps

UM	13	Utility measure
RM	13	Risk measure
PM	13	Performance Index
AI	13	Acceptability Index
LGI	67	Limit Growth Index
CE	13	Certainty Equivalent

Mathematical notation

$\operatorname{ess\,inf}_t X$	8	$\mathcal{F}_t$-condiitonal essential infimum of X
$\operatorname{ess\,inf}_{i\in I} X_i$	8	essential upper bound (essential infimum) of family $\{X_i\}_{i\in I}$
$\operatorname{ess\,inf}_{\omega\in A} X$	8	unconditional essential infimum of X on set A
$X \vee Y$	6	$\max(X,Y)$
$X \wedge Y$	6	$\min(X,Y)$
$L^0(\Omega,\mathcal{G},\mathbb{P})$	6	the set of all $\mathcal{G}$ measurable RVs with values in $(-\infty,\infty)$
$\bar{L}^0(\Omega,\mathcal{G},\mathbb{P})$	6	the set of all $\mathcal{G}$ measurable RVs with values in $[-\infty,\infty]$
$\widehat{L}^p(\Omega,\mathcal{G},\mathbb{P})$	6	$\{X\in\bar{L}^0(\Omega,\mathcal{G},\mathbb{P}) \mid (X\wedge 0)\in L^p(\Omega,\mathcal{G},\mathbb{P})\}$
$L^p(\mathcal{G})$	6	$L^p(\Omega,\mathcal{G},\mathbb{P})$; the same applies to $\bar{L}^p(\mathcal{G})$ and $\widehat{L}^p(\mathcal{G})$
L^p_t	6	$L^p(\Omega,\mathcal{F}_t,\mathbb{P})$; the same applies to $\bar{L}^p_t$ and $\widehat{L}^p_t$
$\mathbb{V}^p$	6	$\{(V_t)_{t\in\mathbb{T}} \mid V_t\in L^p_t,\ t\in\mathbb{T}\}$
$\mathbb{V}^p_{\ln}$	6	$\{(V_t)_{t\in\mathbb{T}} \mid \ln V_t\in L^p_t,\ V_t>0,\ t\in\mathbb{T}\}$
$\mathbb{V}^p_\tau$	6	$\{(V_t)_{t\in\mathbb{T}} \mid V_t\in L^p_t,\ V_t\geq 0,\ \text{and}\ V_t=V_{t\wedge\tau^V},\ t\in\mathbb{T}\}$, $\tau^V:=\inf\{t\in\mathbb{T}\mid V_t=0\}$
$\widehat{f}(X)$	15	$\liminf_{n\to-\infty} f(X\vee n)$
$\bar{f}(X)$	16	$\liminf_{m\to\infty}\widehat{f}(X\wedge m)$
φ^+	15	upper $\bar{L}^0$-extension of φ
φ^-	15	lower $\bar{L}^0$-extension of φ
$\mathrm{Lin}\{\mathcal{Z}\}$	107	linear envelope of $\mathcal{Z}$
$\langle x,y\rangle$	109	standard scalar product
Proj_V	108	projection into V
$\mathrm{Latt}(\mathbf{X})$	103	lattice envelope of $\mathbf{X}$
$\mathcal{B}(A)$	6	σ-algebra of Borel measurable sets of A
$\mathcal{C}(A)$	87	the set of all bounded and continuous functions $f: A\to A$.
$\lVert f\rVert_{\mathrm{span}}$	87	span norm of function f
$\lVert \mu\rVert_{\mathrm{var}}$	91	total variation norm of measure μ
$H(\mu\lVert\nu)$	58	relative entropy of measures μ and ν
$H_t(\mu\lVert\nu)$	58	$\mathcal{F}_t$-cond. relative entropy of measures μ and ν
A^c	6	closure of set A
$\mathcal{M}_1$	17	$\mathcal{M}_1(\Omega,\mathcal{F})$; the set of all probability measures on $(\Omega,\mathcal{F})$.
$\mathcal{M}_1(\mathbb{P})$	17	$\mathcal{M}_1(\Omega,\mathcal{F},P)$; the set of all pr. meas. on $(\Omega,\mathcal{F})$, which are abs. cont. wrt. P
$\mathcal{M}_1^q(P)$	17	$\{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P}) \mid \frac{\mathrm{d}Q}{\mathrm{d}P}\in L^q\}$
$\mathcal{M}_{1,f}(P)$	17	$\{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P}) \mid\ Q\in ba(\mathcal{F})\}$
$ba(\mathcal{F})$	17	the set of all finitely additive signed measures on σ-algebra $\mathcal{F}$.