ON NONLINEAR EXPECTATIONS AND MARKOV CHAINS UNDER MODEL UNCERTAINTY

MAX NENDEL[1]

Abstract. The aim of this work is to give an overview on nonlinear expectation and to relate them to other concepts that describe model uncertainty or imprecision in a probabilistic framework. We discuss imprecise versions of stochastic processes with a particular interest in imprecise Markov chains. First, we focus on basic properties and representations of nonlinear expectations with additional structural assumptions such as translation invariance or convexity. In a second step, we discuss how stochastic processes under nonlinear expectations can be constructed via primal and dual representations. We illustrate the concepts by means of imprecise Markov chains with a countable state space, and show how families of Markov chains give rise to imprecise versions of Markov chains. We discuss dual representations and differential equations related to the latter.

Key words: Nonlinear expectation, imprecise probability, Choquet capacity, imprecise Markov chain, nonlinear transition probability

AMS 2010 Subject Classification: 28E05; 60G20; 60J27; 60J35

1. Introduction

Model uncertainty appears in many scientific disciplines, when, for example, due to statistical estimation methods, only a confidence interval for certain parameters of a model is known, or when certain aspects of a model cannot be determined exactly. In this context, one often speaks of imprecision or polymorphic uncertainty. In mathematical finance, model uncertainty or ambiguity is a frequent phenomenon since financial markets usually do not allow for repetition, whereas, in other disciplines, experiments can be repeated under similar conditions arbitrarily often. The most prominent example for ambiguity in mathematical finance is uncertainty with respect to certain parameters (drift, volatility, etc.) of the stochastic process describing the value of an underlying asset. This leads to the task of modelling stochastic processes under model uncertainty.

In mathematical finance, model uncertainty is often being described via nonlinear expectations, introduced by Peng [29]. Some of the most prominent examples of nonlinear expectations include the g-expectation, see Coquet et al. [10], describing a Brownian Motion with uncertainty in the drift parameter, and the G-expectation or G-Brownian Motion introduced by Peng [30],[31], describing a Brownian Motion with uncertain volatility. There is a close connection between

Financial support through the German Research Foundation via CRC 1283 is gratefully acknowledged.

Date: November 28, 2019.

g-expectations, backward stochastic differential equations (BSDEs) and semilinear partial differential equations. We refer to and Coquet et al. [10] and Pardoux and Peng [27] for more details on this topic. We also refer to Cheridito et al. [5] and Soner et al. [32],[33] for the connection of the G-expectation to 2BSDEs and fully nonlinear partial differential equations. Moreover, there is a one-to-one relation between sublinear expectations and coherent monetary risk measure as introduced by Artzner et al. [1] and Delbaen [13],[14]. Another related concept is the concept of a (Choquet) capacity (see e.g. Dellacherie-Meyer [15]) leading to Choquet integrals (see Choquet [6]).

On the other side, there is a large community working on similar questions related to model uncertainty in the field of imprecise probability. Here, the central objects are upper and lower previsions introduced by Walley [37]. In the sublinear case, there is a one-to-one relation between sublinear expectations and coherent upper previsions, which creates a huge intersection between the communities working on nonlinear expectations and upper/lower previsions. Within the field of imprecise probability, many work has been done in the direction of defining, axiomatizing, and computing transition operators of, both, discrete-time and continuous-time imprecise Markov chains, see e.g. De Bock [11], de Cooman et al. [12], Krak et al. [23], and Škulj [34],[35]. Concepts that are related to imprecise Markov chains include Markov set-chains, see Hartfiel [22], and, in the field of mathematical finance, BSDEs on Markov chains, see Cohen and Elliott [8],[9], and Markov chains under nonlinear expectations, see Nendel [24] and Peng [29].

The aim of this paper is link and compare the concepts and results obtained in the fields of imprecise probability and mathematical finance. Since Markov chains under model uncertainty form the largest intersection between both communities, we put a special focus on the latter. The paper is organized as follows: In Section 2, we start by introducing nonlinear expectations, and discussing basic properties and relations to upper/lower previsions, monetary risk measures and Choquet integrals. In Section 3, we present extension procedures for pre-expectations due to Denk et al. [16]. Here, we focus on two different extensions, one in terms of finitely additive measures, and the other in terms of countably additive measures. In Section 4, we discuss Kolmogorov-type extension theorems and the existence of stochastic processes under nonlinear expectations due to Denk et al. [16]. We conclude, in Section 5, by constructing imprecise versions of transition operators for families of time-homogeneous continuous-time Markov chains with countable state space. Here, we use an approach due to Nisio [26], which has been used in various contexts to construct imprecise versions of Markov processes, such as Lévy processes, Ornstein-Uhlenbeck processes, geometric Brownian Motions, and finite-state Markov chains, see Denk et al. [17], Nendel [24] and Nendel and Röckner [25]. We conlcude by comparing the Nisio approach to the approaches used for continuous-time imprecise Markov chains in the field of imprecise probability.

2. Nonlinear expectations and related concepts

In this section, we give an introduction into the theory of nonlinear expectations, and discuss related concepts. Throughout this section, let Ω be a nonempty set, and $\mathcal{F} \subset 2^\Omega$ be an arbitrary σ-algebra on Ω, where 2^Ω denotes the power set of Ω. We emphasize that, throughout this section, $\mathcal{F} = 2^\Omega$ is a possible choice for $\mathcal{F}$. We denote the space of all bounded $\mathcal{F}$-$\mathcal{B}(\mathbb{R})$-measurable random variables $X\colon \Omega \to \mathbb{R}$ by $\mathcal{L}^\infty(\Omega, \mathcal{F})$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. The space $\mathcal{L}^\infty(\Omega, \mathcal{F})$ and subspaces thereof are endowed with the standard norm $\|\cdot\|_\infty$, defined by

$$\|X\|_\infty := \sup_{\omega\in\Omega} |X(\omega)| \quad (X \in \mathcal{L}^\infty(\Omega, \mathcal{F})).$$

For $\alpha \in \mathbb{R}$, we make use of the notation $\alpha := \alpha 1_\Omega$, and, for $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$, we write $\mathbb{R} \subset M$ instead of $\{\alpha 1_\Omega\colon \alpha \in \mathbb{R}\} \subset M$. Here, 1_A stands for the indicator function of $A \subset \Omega$.

We write $\mathrm{ba}(\Omega, \mathcal{F})$ for the space of all real-valued and finitely additive measures on $(\Omega, \mathcal{F})$ with finite total variation, and $\mathrm{ca}(\Omega, \mathcal{F})$ for the subspace of all σ-additive signed measures on $(\Omega, \mathcal{F})$. The subsets $\mathrm{ba}^1_+(\Omega, \mathcal{F})$ and $\mathrm{ca}^1_+(\Omega, \mathcal{F})$ stand for all positive elements $\mu \in \mathrm{ba}(\Omega, \mathcal{F})$ and $\mu \in \mathrm{ca}(\Omega, \mathcal{F})$ with $\mu(\Omega) = 1$, respectively. Using the identification $\mathrm{ba}(\Omega, \mathcal{F}) = (\mathcal{L}^\infty(\Omega, \mathcal{F}))'$ via $\mu X := \int_\Omega X \,\mathrm{d}\mu$ for $\mu \in \mathrm{ba}(\Omega, \mathcal{F})$ and $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ (cf. [18, p. 258]), every monotone linear functional $\mathbb{E}\colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ with $\mathbb{E}(\alpha) = \alpha$, for all $\alpha \in \mathbb{R}$, is the expectation of a finitely additive probability measure $\mu \in \mathrm{ba}^1_+(\Omega, \mathcal{F})$. This motivates the following definition, which is due to Peng [29].

Definition 2.1. Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$. A *(nonlinear) pre-expectation* $\mathcal{E}$ on M is a functional $\mathcal{E}\colon M \to \mathbb{R}$ with the following properties:

(i) Monotonicity: $\mathcal{E}(X) \le \mathcal{E}(Y)$ for all $X, Y \in M$ with $X \le Y$.
(ii) Constant preserving: $\mathcal{E}(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$.

A pre-expectation on $\mathcal{L}^\infty(\Omega, \mathcal{F})$ is called an *expectation*.

Definition 2.2. Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$, and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation on M.

a) If $M + \mathbb{R} := \{X + \alpha\colon X \in M,\ \alpha \in \mathbb{R}\} \subset M$ and $\mathcal{E}(X + \alpha) = \mathcal{E}(X) + \alpha$ for all $X \in M$ and $\alpha \in \mathbb{R}$, we say that $\mathcal{E}$ is *translation invariant*.
b) If $M + M := \{X + Y\colon X, Y \in M\} \subset M$ and $\mathcal{E}(X + Y) \le \mathcal{E}(X) + \mathcal{E}(Y)$ for all $X, Y \in M$, we say that $\mathcal{E}$ is *subadditive*.
c) If M is convex, i.e. $\lambda X + (1 - \lambda)Y \in M$ for all $X, Y \in M$ and $\lambda \in [0, 1]$, and $\mathcal{E}(\lambda X + (1-\lambda Y)) \le \lambda\mathcal{E}(X) + (1-\lambda)Y$ for all $X, Y \in M$ and $\lambda \in [0, 1]$, we say that $\mathcal{E}$ is *convex*.
d) If $[0, \infty) \cdot M := \{\lambda X\colon \lambda \ge 0, X \in M\} \subset M$ and $\mathcal{E}(\lambda X) = \lambda\mathcal{E}(X)$ for all $\lambda \ge 0$ and $X \in M$, we say that $\mathcal{E}$ is *positive homogeneous (of degreee 1)*.
e) If M is a convex cone, i.e. M is convex and $[0, \infty) \cdot M \subset M$, and $\mathcal{E}$ is convex and positive homogeneous, we say that $\mathcal{E}$ is *sublinear*.

f) If M is a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$, i.e. $\lambda X+Y\in M$ for all $\lambda\in\mathbb{R}$ and $X,Y\in M$, and $\mathcal{E}(\lambda X+Y)=\lambda\mathcal{E}(X)+\mathcal{E}(Y)$ for all $\lambda\in\mathbb{R}$ and $X,Y\in M$, we say that $\mathcal{E}$ is *linear*.

Remark 2.3. Assume that M is a convex cone. Then, $X+Y=2\big(\frac12 X+\frac12 Y\big)\in M$ for all $X,Y\in M$, and therefore, $M+M\subset M$. Since $0\in M$, any two of the following three conditions imply the remaining third:

(i) $\mathcal{E}$ is convex,
(ii) $\mathcal{E}$ is positive homogeneous (of degree 1),
(iii) $\mathcal{E}$ is subadditive.

Remark 2.4. A concept that is very much related to nonlinear expectations is the concept of an upper or lower prevision as introduced by Walley [37]. The latter are very prominent in the context of imprecise probabilities. Walley [37] defines a lower prevision as a real-valued functional $\underline{P}\colon M\to\mathbb{R}$ on an arbitrary set of gambles $M\subset\mathcal{L}^\infty(\Omega,\mathcal{F})$. A nonlinear pre-expectation can thus be seen as a lower prevision prevision that is monotone and preserves constants. The concept of a sublinear expectation is equivalent to the concept of a coherent lower prevision. More precisely, if $\mathcal{E}$ is a sublinear expectation, which is defined on a linear subspace M of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$, then $\underline{P}(X):=-\mathcal{E}(-X)$, for $X\in M$, defines a coherent lower prevision. Moreover, if $\mathcal{E}$ is a convex expectation, which is defined on a linear subspace M of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$, then $\underline{P}(X):=-\mathcal{E}(-X)$ is a convex lower prevision, cf. Pelessoni and Vicig [28, Theorem 3.1].

In this work, we mainly focus on expectations that are translation invariant. Prominent examples for such expectations are Choquet integrals, see Choquet [6] and (monetary) risk measures, see e.g. Föllmer and Schied [20]. We start with some observations that help to verify translation invariance.

Lemma 2.5. *Let $M\subset\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$, and let $\mathcal{E}\colon M\to\mathbb{R}$ be a pre-expectation on M.*

a) *If $\mathcal{E}$ is translation invariant, then $\mathcal{E}$ is 1-Lipschitz, i.e. $|\mathcal{E}(X)-\mathcal{E}(Y)|\leq\|X-Y\|_\infty$ for all $X,Y\in M$.*
b) *Let $M+\mathbb{R}\subset M$. Then, the following three statements are equivalent:*
   (i) *$\mathcal{E}$ is translation invariant,*
   (ii) *$\mathcal{E}(X+\alpha)\leq\mathcal{E}(X)+\alpha$ for all $X\in M$ and all $\alpha\in\mathbb{R}$,*
   (iii) *$\mathcal{E}(X+\alpha)\geq\mathcal{E}(X)+\alpha$ for all $X\in M$ and all $\alpha\in\mathbb{R}$.*
c) *If $\mathcal{E}$ is subadditive, then $\mathcal{E}$ is translation invariant.*
d) *If $\mathcal{E}$ is convex, then $\mathcal{E}$ is translation invariant.*

Proof.

a) Assume that $\mathcal{E}$ is translation invariant. Then,
$$\mathcal{E}(X)-\mathcal{E}(Y)\leq\mathcal{E}(Y+\|X-Y\|_\infty)-\mathcal{E}(Y)=\|X-Y\|_\infty$$
for all $X,Y\in M$. By a symmetry argument, we obtain that $|\mathcal{E}(X)-\mathcal{E}(Y)|\leq\|X-Y\|_\infty$ for all $X,Y\in M$.
b) First assume that $\mathcal{E}(X+\alpha)\leq\mathcal{E}(X)+\alpha$ for all $X\in M$ and all $\alpha\in\mathbb{R}$. Then,
$$\mathcal{E}(X)=\mathcal{E}\big((X+\alpha)-\alpha\big)\leq\mathcal{E}(X+\alpha)-\alpha.$$

Now, assume that $\mathcal{E}(X+\alpha) \geq \mathcal{E}(X)+\alpha$ for all $X \in M$ and all $\alpha \in \mathbb{R}$. Then,

$$\mathcal{E}(X+\alpha)-\alpha \leq \mathcal{E}\big((X+\alpha)-\alpha\big)=\mathcal{E}(X).$$

c) This follows directly from part b).

d) Assume that $\mathcal{E}$ is convex. Let $X \in M$ and $\alpha \in \mathbb{R}$. Then, for all $\lambda \in [0,1)$,

$$\begin{aligned}\mathcal{E}(X+\alpha) &\leq \lambda\mathcal{E}(X)+(1-\lambda)\mathcal{E}\big(X+\tfrac{\alpha}{1-\lambda}\big)\\ &\leq \lambda\mathcal{E}(X)+(1-\lambda)\mathcal{E}\big(\|X\|_\infty+\tfrac{\alpha}{1-\lambda}\big)\\ &= \lambda\mathcal{E}(X)+(1-\lambda)\|X\|_\infty+\alpha.\end{aligned}$$

Letting $\lambda \to 1$, we obtain that $\mathcal{E}(X+\alpha) \leq \mathcal{E}(X)+\alpha$. By part b), it follows that $\mathcal{E}$ is translation invariant.

□

Remark 2.6. Since every convex pre-expectation $\mathcal{E}$ is translation invariant, $\rho(X) := \mathcal{E}(-X)$ defines a convex monetary risk measure, cf. Föllmer and Schied [19] and Frittelli and Rosazza Gianin [21]. If $\mathcal{E}$ is sublinear, i.e. convex and subadditive, then ρ is a coherent monetary risk measure as introduced by Artzner et al. [1], see also Delbaen [13], [14].

As in the theory of risk measures, random variables with positive expectation, play a special role, and we refer to them as *acceptable positions*.

Definition 2.7. Let $M \subset \mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R} \subset M$.

a) A set $\mathcal{A} \subset M$ is called an *acceptance set* of M if
  (i) $\inf\{\alpha \in \mathbb{R} \colon \alpha \in \mathcal{A}\}=0$,
  (ii) for all $X \in \mathcal{A}$ and all $Y \in M$ with $Y \geq X$ we have that $Y \in \mathcal{A}$,
  (iii) $\mathcal{A}$ is a closed subset of M.

b) Let $\mathcal{E} \colon M \to \mathbb{R}$ be a pre-expectation. Then, the set

$$\mathcal{A}_\mathcal{E} := \{X \in M \colon \mathcal{E}(X) \geq 0\}$$

is called the *acceptance set* of $\mathcal{E}$.

In the field of imprecise probability, acceptable positions are called *desirable gambles*, and acceptance sets are called *sets of desirable gambles*. Translation invariant expectations are uniquely characterized via their acceptable positions. This is a well-known fact within the theory of risk measures and imprecise prbabilities, and directly carries over to translation invariant expectations. For the reader's convenience, we state the result and provide a short proof.

Proposition 2.8. *Let $M \subset \mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R} \subset M$ and $M+\mathbb{R} \subset M$. Then, the mapping $\mathcal{E} \mapsto \mathcal{A}_\mathcal{E}$ is a bijection between the set of all translation invariant pre-expectations on M and the set of all acceptance sets of M. More precisely, the following holds:*

a) *Let $\mathcal{E} \colon M \to \mathbb{R}$ be a translation invariant pre-expectation. Then, $\mathcal{A}_\mathcal{E}$ is an acceptance set in the sense of Definition 2.7 a), and*

$$\mathcal{E}(X)=\sup\{\alpha \in \mathbb{R} \colon X-\alpha \in \mathcal{A}_\mathcal{E}\}$$

for all $X \in M$.

b) *Let* $\mathcal{A} \subset M$ *be an acceptance set. Then,*

$$\mathcal{E}(X) := \sup\{\alpha \in \mathbb{R} \colon X - \alpha \in \mathcal{A}\}, \quad \textit{for } X \in M,$$

defines a translation invariant pre-expectation $\mathcal{E} \colon M \to \mathbb{R}$ *with* $\mathcal{A}_{\mathcal{E}} = \mathcal{A}$.

Proof.

a) Since $\mathcal{E}$ is translation invariant, for all $X \in M$,

$$\begin{aligned}\mathcal{E}(X) &= \sup\{\alpha \in \mathbb{R} \colon \mathcal{E}(X) \geq \alpha\} = \sup\{\alpha \in \mathbb{R} \colon \mathcal{E}(X - \alpha) \geq 0\} \\ &= \sup\{\alpha \in \mathbb{R} \colon X - \alpha \in \mathcal{A}_{\mathcal{E}}\}.\end{aligned}$$

Clearly, $\mathcal{A}$ satisfies the properties (i) and (ii) in Definition 2.7. Since $\mathcal{E}$ is translation invariant, it is a 1-Lipschitz continuous map $M \to \mathbb{R}$, and therefore $\mathcal{A} = \mathcal{E}^{-1}\big([0, \infty)\big)$ is a closed subset of M.

b) One readily verifies that $\mathcal{E}$ defines a translation invariant pre-expectation on M. By definition, $\mathcal{A} \subset \mathcal{A}_{\mathcal{E}}$. On the other hand, for $X \in \mathcal{A}_{\mathcal{E}}$, there exists a sequence $(\alpha_n)_{n \in \mathbb{N}} \subset (-\infty, 0]$ with $X - \alpha_n \in \mathcal{A}$ and $\alpha_n \to 0$ as $n \to \infty$. Since $\mathcal{A}$ is a closed subset of M, it follows that $X = \lim_{n \to \infty} X - \alpha_n \in \mathcal{A}$.

□

We now specialize on the case, where $\mathcal{E}$ is a convex expectation on a linear subspace of $\mathcal{L}^\infty(\Omega, \mathcal{F})$. Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ be a linear subspace of $\mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$. For a convex function $\mathcal{E} \colon M \to \mathbb{R}$, we write $\mathcal{E}^*$ for its *conjugate function* or *Fenchel-Legendre transform*, i.e. we define

$$\mathcal{E}^*(\mu) := \sup_{X \in M} \mu X - \mathcal{E}(X)$$

for all linear functionals $\mu \colon M \to \mathbb{R}$. Note that the conjugate function $\mathcal{E}^*$ may take the value $+\infty$. In the following proposition, we will see that, for a convex pre-expectation $\mathcal{E}$ on M, its conjugate function $\mathcal{E}^*$ is concentrated on the class of linear pre-expectations on M. That is, $\mathcal{E}^*$ is finite only for linear pre-expectations on M. As every linear pre-expectation on M is continuous, we therefore obtain the representation

$$\mathcal{E}(X) = \sup_{\mu \in M'} \mu X - \mathcal{E}^*(\mu) \quad (X \in \mathcal{L}^\infty(\Omega, \mathcal{F}))$$

for all convex pre-expectations $\mathcal{E}$ on M. Again, this type of representation is well-known for convex risk measures, and the proof relies on a collection of several well-known facts from convex analysis and duality theory. In order to keep this manuscript self-contained, we nevertheless state the proof below.

Proposition 2.9. *Let* $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ *be a linear subspace of* $\mathcal{L}^\infty(\Omega, \mathcal{F})$ *with* $\mathbb{R} \subset M$.

a) *Let* $\mathcal{E} \colon M \to \mathbb{R}$ *be a convex pre-expectation. Then, every linear functional* $\mu \colon M \to \mathbb{R}$ *with* $\mathcal{E}^*(\mu) < \infty$ *is a linear pre-expectation on* M *and therefore,* $\mu \in M'$ *with* $\|\mu\|_{M'} = 1$, *where* $\|\cdot\|_{M'}$ *denotes the operator norm on* M'.

b) *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation, and let $\mathcal{P}_\mathcal{E}$ denote the set of all linear functionals $\mu\colon M \to \mathbb{R}$ with $\mathcal{E}^*(\mu) < \infty$. Then, for all $X \in M$,*

$$\mathcal{E}(X) = \max_{\mu \in \mathcal{P}_\mathcal{E}} \mu X - \mathcal{E}^*(\mu).$$

Moreover, the map $\mathcal{E}^*\colon \mathcal{P}_\mathcal{E} \to \mathbb{R}$ is convex and weak lower semicontinuous with*

$$\min_{\mu \in \mathcal{P}_\mathcal{E}} \mathcal{E}^*(\mu) = 0$$

c) *Let $\mathcal{P} \subset M'$ be a set of linear pre-expectations, and $\sigma\colon \mathcal{P} \to \mathbb{R}$ be a map with $\inf_{\mu \in \mathcal{P}} \alpha(\mu) = 0$. Then,*

$$\mathcal{E}(X) := \sup_{\mu \in \mathcal{P}} \mu X - \sigma(\mu), \quad \textit{for } X \in M,$$

defines a convex pre-expectation on M. It holds $\mathcal{P} \subset \mathcal{P}_\mathcal{E}$ and $\mathcal{E}^*(\mu) \le \sigma(\mu)$ for all $\mu \in \mathcal{P}$.

Proof.

a) Let $\mu\colon M \to \mathbb{R}$ be linear with $\mathcal{E}^*(\mu) < \infty$. Then, for all $\lambda > 0$,

$$\begin{aligned} 1 - \lambda^{-1}\mathcal{E}^*(\mu) &= -\lambda^{-1}\big(\mathcal{E}(-\lambda) + \mathcal{E}^*(\mu)\big) \\ &\le -\lambda^{-1}\mu(-\lambda) = \mu 1 = \lambda^{-1}\mu(\lambda) \\ &\le \lambda^{-1}\big(\mathcal{E}(\lambda) + \mathcal{E}^*(\mu)\big) = 1 + \lambda^{-1}\mathcal{E}^*(\mu). \end{aligned}$$

Letting $\lambda \to \infty$, we obtain that $\mu 1 = 1$. Moreover, for $\lambda > 0$ and all $X, Y \in M$ with $X \le Y$,

$$\begin{aligned} \mu(X - Y) &= \lambda^{-1}\mu\big(\lambda(X - Y)\big) \le \lambda^{-1}\big[\mathcal{E}\big(\lambda(X - Y)\big) + \mathcal{E}^*(\mu)\big] \\ &\le \lambda^{-1}\mathcal{E}^*(\mu) \to 0, \quad \lambda \to \infty. \end{aligned}$$

This shows that $\mu\colon M \to \mathbb{R}$ is a linear pre-expectation on M. Therefore, Lemma 2.5 a) and c) imply that $\mu \in M'$ with $\|\mu\|_{M'} \le 1$. Since $\mu 1 = 1$, it follows that $\|\mu\|_{M'} = 1$.

b) Let $X \in M$ and $\mathcal{E}_0(Y) := \mathcal{E}(X + Y) - \mathcal{E}(X)$ for all $Y \in M$. Then, $\mathcal{E}_0\colon M \to \mathbb{R}$ is convex and $\mathcal{E}_0(0) = 0$. By the extension theorem of Hahn-Banach, there exists a linear functional $\mu\colon M \to \mathbb{R}$ with $\mu Y \le \mathcal{E}_0(Y)$ for all $Y \in M$. That is,

$$\mu Y - \mathcal{E}(Y) \le \mu X - \mathcal{E}(X)$$

for all $Y \in M$. Therefore, by definition of the conjugate function,

$$\mathcal{E}(X) = \mu X - \mathcal{E}^*(\mu).$$

One readily checks that $\mathcal{P}_\mathcal{E} \subset M'$ is convex. Hence, for fixed $X \in M$, the map

$$\mathcal{P}_\mathcal{E} \to \mathbb{R}, \quad \mu \mapsto \mu X - \mathcal{E}(X)$$

is convex and weak* continuous. Taking the pointwise supremum over all $X \in M$, we see that the mapping $\mathcal{P}_\mathcal{E} \to \mathbb{R}$, $\mu \mapsto \sup_{X \in M} \mu X - \mathcal{E}(X)$ is convex and weak* lower semicontinuous. As $0 \in M$ with $\mathcal{E}(0) = 0$, it follows that $\mathcal{E}^*(\mu) \ge 0$ for all $\mu \in \mathcal{P}_\mathcal{E}$. Again, as $0 \in M$, there exists some $\mu \in \mathcal{P}_\mathcal{E}$ with $\mathcal{E}^*(\mu) = \mu 0 - \mathcal{E}(0) = 0$.

c) The map $\mathcal{E}$ is monotone and convex, as it is a supremum over monotone affine linear maps. Moreover,

$$\mathcal{E}(\alpha) = \sup_{\mu\in\mathcal{P}} \mu\alpha - \sigma(\mu) = \alpha - \inf_{\mu\in\mathcal{P}} \sigma(\mu) = \alpha.$$

Therefore, $\mathcal{E}$ defines a convex pre-expectation on M. Let $\nu \in \mathcal{P}$. Then,

$$\nu X - \mathcal{E}(X) = \nu X - \sup_{\mu\in\mathcal{P}} \mu X - \sigma(\mu) \leq \sigma(\mu)$$

for all $X \in M$. Hence, $\mathcal{E}^*(\nu) \leq \sigma(\nu)$ and, in particular, $\mathcal{P} \subset \mathcal{P}_\mathcal{E}$.

□

The set $\mathcal{P}_\mathcal{E}$ from the previous proposition can be used to characterize sublinear and linear pre-expectations. In the field of imprecise probability or, more precisely, for coherent upper previsions the set $\mathcal{P}_\mathcal{E}$ is called the *credal set*.

Lemma 2.10. *Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ be a linear subspace of $\mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$, and let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation.*

a) *$\mathcal{E}$ is sublinear if and only if $\mathcal{P}_\mathcal{E} = \{\mu \in M'\colon \mathcal{E}^*(\mu) = 0\}$.*
b) *$\mathcal{E}$ is linear if and only if $\#\mathcal{P}_\mathcal{E} = 1$. In this case, $\mathcal{P}_\mathcal{E} = \{\mathcal{E}\}$.*

Proof.

a) First, assume that $\mathcal{E}$ is sublinear. Let $\mu \in \mathcal{P}_\mathcal{E}$ and $X \in M$ be arbitrary. Then,

$$\lambda\big(\mu X - \mathcal{E}(X)\big) = \mu(\lambda X) - \mathcal{E}(\lambda X) \leq \mathcal{E}^*(\mu) < \infty$$

for all $\lambda > 0$, and therefore, $\mu X - \mathcal{E}(X) \leq 0$. Taking the supremum over all $X \in M$, we obtain that $\mathcal{E}^*(\mu) \leq 0$. By Proposition 2.9 b), this implies that $\mathcal{E}^*(\mu) = 0$.
Now assume that $\mathcal{P}_\mathcal{E} = \{\mu \in M'\colon \mathcal{E}^*(\mu) = 0\}$. Then, for all $X \in M$ and all $\lambda \geq 0$ we have that

$$\mathcal{E}(\lambda X) = \max_{\mu\in\mathcal{P}} \mu(\lambda X) = \lambda \max_{\mu\in\mathcal{P}} \mu X = \lambda\mathcal{E}(X).$$

b) Assume that $\mathcal{E}$ is linear and let $\mu \in \mathcal{P}_\mathcal{E}$. Then, by part a), $\mathcal{E}^*(\mu) = 0$, and therefore, $\mu X \leq \mathcal{E}(X)$ for all $X \in M$. As $\mathcal{E}$ is linear, we thus obtain that $\mu = \mathcal{E}$. This shows that $\mathcal{P} = \{\mathcal{E}\}$.
Now let $\#\mathcal{P} = 1$ and let $\mu \in \mathcal{P}$. Then, by Proposition 2.9 b), $\mathcal{E}^*(\mu) = 0$, and therefore, $\mathcal{E} = \mu$ is linear.

□

Proposition 2.9 together with Lemma 2.10 yields the following characterization of sublinear pre-expectations.

Proposition 2.11. *Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ be a linear subspace of $\mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$. Then, the map $\mathcal{E} \mapsto \mathcal{P}_\mathcal{E}$ is a bijection between the set of all sublinear expectations and the set of all nonempty, convex and weak* compact sets of linear pre-expectations. More precisely, the following holds:*

a) *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a sublinear pre-expectation. Then,*
$$\mathcal{E}(X) = \max_{\mu \in \mathcal{P}_\mathcal{E}} \mu X.$$
Moreover, the set $\mathcal{P}_\mathcal{E}$ is nonempty, convex, and weak compact.*

b) *Let $\mathcal{P}$ be a nonempty set of linear pre-expectations. Then,*
$$\mathcal{E}(X) := \max_{\mu \in \mathcal{P}} \mu X, \quad \textit{for } X \in M,$$
gives rise to a sublinear pre-expectation on M. If $\mathcal{P}$ is convex and weak compact, then $\mathcal{P} = \mathcal{P}_\mathcal{E}$.*

Proof.

a) By Lemma 2.10, $\mathcal{P}_\mathcal{E} = \{\mu \in M'\colon \mathcal{E}^*(\mu) = 0\}$. Therefore, by Proposition 2.9 b),
$$\mathcal{E}(X) = \max_{\mu \in \mathcal{P}_\mathcal{E}} \mu X.$$
Clearly, $\mathcal{P}_\mathcal{E}$ is convex and, by Proposition 2.9, nonempty. Since
$$\mathcal{P}_\mathcal{E} = \bigcap_{X \in M} \{\mu \in M'\colon \mu X \le \mathcal{E}(X)\}$$
is closed, and $\mathcal{P}_\mathcal{E} \subset \{\mu \in M'\colon \|\mu\|_{M'} \le 1\}$, we obtain that $\mathcal{P}_\mathcal{E}$ is weak* compact by the Banach-Alaoglu Theorem.

b) One readily verifies that $\mathcal{E}$ defines a sublinear pre-expectation on M. Assume that $\mathcal{P}$ is convex and weak* compact. Then, by Proposition 2.9 c), $\mathcal{P} \subset \mathcal{P}_\mathcal{E}$. In order to prove the other inclusion, let $\nu \in M' \setminus \mathcal{P}$. Then, by the separation theorem of Hahn-Banach, there exists some $X \in M$ with
$$\mathcal{E}(X) = \sup_{\mu \in \mathcal{P}_\mathcal{E}} \mu X < \nu X,$$
where we used the fact that the topological dual of M' (endowed with the weak* topology) is M. Hence, $\mathcal{E}^*(\nu) > 0$ and therefore, by Lemma 2.10 a), $\nu \notin \mathcal{P}_\mathcal{E}$. This shows that $\mathcal{P} = \mathcal{P}_\mathcal{E}$.

□

Example 2.12. Let $c\colon \mathcal{F} \to [0,1]$ be a capacity, i.e. function with $c(\emptyset) = 0$, $c(\Omega) = 1$, and $c(A) \le c(B)$ for all $A, B \in \mathcal{F}$ with $A \subset B$. Then, c gives rise to a translation invariant expectation $\mathcal{E}\colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ via
$$\mathbb{E}(X) := \int X \,\mathrm{d}c := \int_0^\infty c(X \ge x)\,\mathrm{d}x, \quad \text{for } X \in \mathcal{L}^\infty(\Omega, \mathcal{F}) \text{ with } X \ge 0,$$
and $\mathcal{E}(X) := \int X \,\mathrm{d}c := \mathcal{E}\big(X + \|X\|_\infty\big) - \|X\|_\infty$ for $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$. The nonlinear expectation $\mathcal{E}$ is called the Choquet integral of X w.r.t. c, see Choquet [6]. By definition, $\mathcal{E}$ is positive homogeneous, translation invariant, and satisfies $\mathcal{E}(1_A) = c(A)$ for all $A \in \mathcal{F}$. Assume that c is 2-*alternating*, i.e.
$$c(A \cup B) + c(A \cap B) \le c(A) + c(B) \quad \text{for all } A, B \in \mathcal{F}.$$
Then it is well-known that $\mathcal{E}$ is subadditive and thus sublinear. By definition, $\mu X \le \mathcal{E}(X)$ for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ and every linear expectation $\mu \in \mathrm{ba}^1_+(\Omega, \mathcal{F})$

with $\mu(A) \leq c(A)$ for all $A \in \mathcal{F}$. Recall that linear expectations can be identified by finitely additive measures via $\mu(A) := \mu 1_A$, for $A \in \mathcal{F}$. This implies that

$$\mathcal{P}_{\mathcal{E}} = \big\{\mu \in \mathrm{ba}^1_+(\Omega, \mathcal{F}) \colon \forall A \in \mathcal{F} \colon \mu(A) \leq c(A)\big\},$$

and thus $\tilde{\mathcal{E}}(X) \leq \mathcal{E}(X)$ for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ and every sublinear expectation $\tilde{\mathcal{E}} \colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ with $\tilde{\mathcal{E}}(1_A) \leq c(A)$ for all $A \in \mathcal{F}$. That is, $\mathcal{E}$ is the largest sublinear expectation with $\mathcal{E}(1_A) \leq c(A)$ for all $A \in \mathcal{F}$. The following corollary shows that a sublinear expectation that coincides with a finitely additive measure on all $\mathcal{F}$-measurable sets is already linear.

Corollary 2.13. *Let $\mathcal{E} \colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ be a sublinear expectation and μ be a finitely additive measure with $\mathcal{E}(1_A) \leq \mu(A)$ for all $A \in \mathcal{F}$. Then, $\mathcal{E}(X) = \mu X$ for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ and therefore, $\mathcal{E}$ is a linear expectation.*

Proof. Let $\nu \in \mathcal{P}_{\mathcal{E}}$. Then, ν is a linear expectation, i.e. a finitely additive measure, with $\nu(A) \leq \mathcal{E}(1_A) \leq \mu(A)$ for all $A \in \mathcal{F}$. Hence, $\mu = \nu$, i.e. $\mathcal{P}_{\mathcal{E}} = \{\mu\}$, which, by Lemma 2.10 b), implies that $\mathcal{E} = \mu$. □

For sublinear expectations, the following version of Jensen's inequality holds.

Lemma 2.14. *Let $\mathcal{E} \colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ be a sublinear expectation and $h \colon \mathbb{R} \to \mathbb{R}$ be a convex function. Then, for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$,*

$$h\big(\mathcal{E}(X)\big) \leq \mathcal{E}\big(h(X)\big).$$

Proof. Since $h \colon \mathbb{R} \to \mathbb{R}$ is convex, h is continuous, and therefore, $h(X) \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$. Moreover,

$$h(x) = \max_{\lambda \in U} \lambda x - h^*(\lambda), \quad \text{for all } x \in \mathbb{R},$$

where $h^* \colon \mathbb{R} \to \mathbb{R}$ is the conjugate function or Fenchel-Legendre transform of h and $U := \{\lambda \in \mathbb{R} \colon h^*(\lambda) < \infty\}$. Now. let $X \in M$ and $\lambda \in U$. If $\lambda \geq 0$, then

$$\lambda \mathcal{E}(X) - h^*(\lambda) = \mathcal{E}\big(\lambda X - h^*(\lambda)\big) \leq \mathcal{E}\big(h(X)\big).$$

If $\lambda < 0$, then

$$\lambda \mathcal{E}(X) - h^*(\lambda) \leq \mathcal{E}(\lambda X) - h^*(\lambda) = \mathcal{E}\big(\lambda X - h^*(\lambda)\big) \leq \mathcal{E}\big(h(X)\big),$$

where, in the first inequality, we used the sublinearity of $\mathcal{E}$. Taking the supremum over all $\lambda \in U$ yields the assertion. □

As in the linear case, one can define the concept of a distribution for nonlinear expectations.

Remark 2.15. Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E} \colon M \to \mathbb{R}$ be a pre-expectation. Moreover, let $S \neq \emptyset$ be a set and $T \colon \Omega \to S$ be an arbitrary map. Then, $N := \{Y \in \mathcal{L}^\infty(S, 2^S) \colon Y \circ T \in M\}$ contains all constant functions $S \to \mathbb{R}$, and one readily verifies that

$$\mathcal{E} \circ T^{-1} \colon N \to \mathbb{R}, \quad Y \mapsto \mathcal{E}(Y \circ T)$$

defines a pre-expectation on N. We call $\mathcal{E} \circ T^{-1}$ the *distribution* of T under $\mathcal{E}$. Note that if $M = \mathcal{L}^\infty(\Omega, \mathcal{F})$, then $N = \mathcal{L}^\infty(S, \mathscr{S})$, where $\mathscr{S} := \{B \in 2^S \colon T^{-1}(B) \in \mathcal{F}\}$. In particular, $N \subset \mathcal{L}^\infty(S, \mathscr{S})$ for all $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$.

Now, if M is, additionally, a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ and $\mathcal{E}$ is sublinear, then, N is a linear subspace of $\mathcal{L}^\infty(S,\mathscr{S})$ and $\mathcal{E}\circ T^{-1}$ is sublinear expectation. In this case,

$$\mathcal{P}_{\mathcal{E}\circ T^{-1}}=\big\{\nu\in N'\colon (\mathcal{E}\circ T^{-1})^*(\nu)<\infty\big\}=\{\mu\circ T^{-1}\colon \mu\in\mathcal{P}_\mathcal{E}\}=:\mathcal{P}_\mathcal{E}\circ T^{-1}.$$

In fact, as the map $M'\to N'$, $\mu\mapsto\mu\circ T^{-1}$ is linear and weak* continuous, the set $\mathcal{P}_\mathcal{E}\circ T^{-1}$ is convex and weak* compact. Moreover, for all $Y\in N$,

$$(\mathcal{E}\circ T^{-1})(Y)=\mathcal{E}(Y\circ T)=\max_{\mu\in\mathcal{P}_\mathcal{E}}\mu(Y\circ T)=\max_{\mu\in\mathcal{P}_\mathcal{E}}(\mu\circ T^{-1})Y=\max_{\nu\in\mathcal{P}_\mathcal{E}\circ T^{-1}}\nu Y.$$

By Lemma 2.11 b), it follows that $\mathcal{P}_{\mathcal{E}\circ T^{-1}}=\mathcal{P}_\mathcal{E}\circ T^{-1}$.

3. Extension of pre-expectations

Let $M\subset\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$. Given a pre-expectation $\mathcal{E}\colon M\to\mathbb{R}$, we are looking for extensions of $\mathcal{E}$ to an expectation on $\mathcal{L}^\infty(\Omega,\mathcal{F})$. Here, the main challenge is to preserve monotonicity. We start with the extension of linear pre-expectations.

Remark 3.1. Let $M\subset\mathcal{L}^\infty(\Omega,\mathcal{F})$ be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $1\in M$. We denote by $\mathrm{ba}^1_+(M)$ the space of all linear pre-expectations on M. A natural question is if the mapping

$$\mathrm{ba}^1_+(\Omega,\mathcal{F})\to\mathrm{ba}^1_+(M),\quad \nu\mapsto\nu|_M \tag{3.1}$$

is bijective. The following theorem by Kantorovich shows that this mapping is surjective, i.e. any linear pre-expectation on M can be extended to a linear expectation on $\mathcal{L}^\infty(\Omega,\mathcal{F})$. For the reader's convenience, we state this theorem and provide a short sketch of the proof. For more details we refer to [36, p. 277]. However, in Example 3.15, we will see that, in general, the mapping in (3.1) is not injective, not even if $\mathcal{F}=\sigma(M)$, i.e. even if $\mathcal{F}=\sigma(M)$, a linear pre-expectation on M usually admits various extensions to an expectation on $\mathcal{L}^\infty(\Omega,\mathcal{F})$.

Theorem 3.2 (Kantorovich)**.** *Let $M\subset\mathcal{L}^\infty(\Omega,\mathcal{F})$ be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $1\in M$ and $\mu\colon M\to\mathbb{R}$ be a linear pre-expectation on M. Then, there exists a linear expectation $\nu\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ with $\nu|_M=\mu$.*

Proof. Let

$$\widehat{\mu}(X):=\inf\{\mu X_0\colon X_0\in M,\ X_0\geq X\}$$

for all $X\in\mathcal{L}^\infty(\Omega,\mathcal{F})$. Then, $\widehat{\mu}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ is a sublinear expectation with $\widehat{\mu}|_M=\mu$. By the extension theorem of Hahn-Banach, there exists a linear functional $\nu\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ with $\nu|_M=\mu$ and $\nu X\leq\widehat{\mu}(X)$ for all $X\in\mathcal{L}^\infty(\Omega,\mathcal{F})$. Thus,

$$\nu X-\nu Y=\nu(X-Y)\leq\widehat{\mu}(X-Y)\leq\widehat{\mu}(0)=0$$

for all $X,Y\in\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $X\leq Y$. □

In the proof of the previous theorem, before applying the Hahn-Banach Theorem, the linear pre-expectation $\mu\colon M\to\mathbb{R}$ is extended to a sublinear expectation $\widehat{\mu}\colon \mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ via

$$\widehat{\mu}(X):=\inf\{\mu X_0\colon X_0\in M,\ X_0\geq X\}\quad (X\in\mathcal{L}^\infty(\Omega,\mathcal{F})).$$

The idea for the first extension procedure therefore is, to extend a pre-expectation $\mathcal{E}\colon M\to\mathbb{R}$ on $M\subset\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ via

$$\widehat{\mathcal{E}}(X):=\inf\{\mathcal{E}(X_0)\colon X_0\in M,\ X_0\geq X\}\quad (X\in\mathcal{L}^\infty(\Omega,\mathcal{F})).$$

The following proposition is due to Denk et al. [16], and shows that $\widehat{\mathcal{E}}$ is an expectation with $\widehat{\mathcal{E}}|_M=\mathcal{E}$. Moreover, if M is assumed to be convex or a convex cone, then convexity or sublinearity of $\mathcal{E}$ carry over to the extension $\widehat{\mathcal{E}}$, respectively. For related extension results on niveloids, we refer to Maccheroni et al. [4]. In the context of lower previsions, for $\mathcal{F}=2^\Omega$, the extension $\widehat{\mathcal{E}}$ is usually referred to as the *natural extension*, cf. Walley [37, Section 3.1].

Proposition 3.3. *Let $M\subset\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ and $\mathcal{E}\colon M\to\mathbb{R}$ be a pre-expectation on M. Further, let*

$$\widehat{\mathcal{E}}(X):=\inf\{\mathcal{E}(X_0)\colon X_0\in M,\ X_0\geq X\}$$

for all $X\in\mathcal{L}^\infty(\Omega,\mathcal{F})$. Then, the following assertions hold:

a) *$\widehat{\mathcal{E}}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ is an expectation with $\widehat{\mathcal{E}}|_M=\mathcal{E}$.*
b) *If $\mathcal{E}$ is translation invariant, then $\widehat{\mathcal{E}}$ is translation invariant.*
c) *If $\mathcal{E}$ is convex, then $\widehat{\mathcal{E}}$ is convex.*
d) *If $\mathcal{E}$ is sublinear, then $\widehat{\mathcal{E}}$ is sublinear.*

Proof.

a) Let $X\in\mathcal{L}^\infty(\Omega,\mathcal{F})$. As $\mathbb{R}\subset M$, we have that $\|X\|_\infty\in M$ with $\|X\|_\infty\geq X$. Thus, the set $\{\mathcal{E}(X_0)\colon X_0\in M,\ X_0\geq X\}$ is nonempty. Since $X_0\geq -\|X\|_\infty$ for all $X_0\in M$ with $X_0\geq X$, we obtain that

$$\mathcal{E}(X_0)\geq\mathcal{E}(-\|X\|_\infty)=-\|X\|_\infty$$

for all $X_0\in M$ with $X_0\geq X$. Hence, $\widehat{\mathcal{E}}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ is well-defined. Further, if $X\in M$ we have that $\mathcal{E}(X)\leq\mathcal{E}(X_0)$ for all $X_0\in M$ with $X_0\geq X$ and, therefore, $\widehat{\mathcal{E}}(X)=\mathcal{E}(X)$. Since $\mathbb{R}\subset M$, we thus obtain that $\widehat{\mathcal{E}}(\alpha)=\alpha$ for all $\alpha\in\mathbb{R}$. Now, let $X,Y\in\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $X\leq Y$. Then, $Y_0\geq X$ for all $Y_0\in M$ with $Y_0\geq Y$ and, therefore, $\widehat{\mathcal{E}}(X)\leq\widehat{\mathcal{E}}(Y)$.

b) Assume that $M+\mathbb{R}\subset M$ and that $\mathcal{E}$ is translation invariant. Let $X\in\mathcal{L}^\infty(\Omega,\mathcal{F})$ and $\alpha\in\mathbb{R}$. Then, for all $X_0\in M$ with $X_0\geq X+\alpha$ we have that $X_0-\alpha\geq X$ and therefore,

$$\mathcal{E}(X_0)=\mathcal{E}(X_0-\alpha)+\alpha\geq\widehat{\mathcal{E}}(X)+\alpha.$$

Taking the infimum over all $X_0\in M$ with $X_0\geq X+\alpha$ yields that $\widehat{\mathcal{E}}(X+\alpha)\geq\widehat{\mathcal{E}}(X)+\alpha$, which, by Lemma 2.5 b) implies that $\widehat{\mathcal{E}}$ is translation invariant.

c) Assume that M is convex and that $\mathcal{E}$ is convex. Let $X, Y \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ and $\lambda \in [0,1]$. Moreover, let $X_0, Y_0 \in M$ with $X_0 \geq X$ and $Y_0 \geq Y$. Since M is convex, $\lambda X_0 + (1-\lambda)Y_0 \in M$ with

$$\lambda X_0 + (1-\lambda)Y_0 \geq \lambda X + (1-\lambda)Y.$$

Due to convexity of $\mathcal{E}$, we thus obtain that

$$\widehat{\mathcal{E}}(\lambda X + (1-\lambda)Y) \leq \mathcal{E}(\lambda X_0 + (1-\lambda)Y_0) \leq \lambda \mathcal{E}(X_0) + (1-\lambda)\mathcal{E}(Y_0).$$

Taking the infimum over all $X_0, Y_0 \in M$ with $X_0 \geq X$ and $Y_0 \geq Y$, we get that

$$\widehat{\mathcal{E}}(\lambda X + (1-\lambda)Y) \leq \lambda \widehat{\mathcal{E}}(X) + (1-\lambda)\widehat{\mathcal{E}}(Y).$$

d) Now assume that M is a convex cone and that $\mathcal{E}$ is sublinear. Then, $\mathcal{E}$ is convex and part b) yields that $\widehat{\mathcal{E}}$ is convex as well. Moreover, as $\lambda X_0 \in M$ for all $X_0 \in M$ and $\lambda > 0$ we have that

$$\begin{aligned}\widehat{\mathcal{E}}(\lambda X) &= \inf\{\mathcal{E}(\lambda X_0) \colon X_0 \in M,\ X_0 \geq X\} \\ &= \inf\{\lambda \mathcal{E}(X_0) \colon X_0 \in M,\ X_0 \geq X\} = \lambda \widehat{\mathcal{E}}(X)\end{aligned}$$

for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ and all $\lambda > 0$. Hence, $\widehat{\mathcal{E}}$ is convex and positive homogeneous, and therefore sublinear.

□

Remark 3.4.

a) Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E} \colon M \to \mathbb{R}$ be a pre-expectation on M. Let $\tilde{\mathcal{E}} \colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ be an expectation with $\tilde{\mathcal{E}}|_M = \mathcal{E}$ and $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$. Then,

$$\widetilde{\mathcal{E}}(X) \leq \widetilde{\mathcal{E}}(X_0) = \mathcal{E}(X_0)$$

for all $X_0 \in M$ with $X_0 \geq X$. Taking the infimum over all $X_0 \in M$ with $X_0 \geq X$, we see that $\widetilde{\mathcal{E}}(X) \leq \widehat{\mathcal{E}}(X)$. That is, $\widehat{\mathcal{E}}$ is the largest expectation, which extends $\mathcal{E}$.

b) Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E} \colon M \to \mathbb{R}$ be a pre-expectation on M. For $X \in M$, let

$$\check{\mathcal{E}}(X) := \sup\{\mathcal{E}(X_0) \colon X_0 \in M,\ X_0 \leq X\}.$$

Then, one readily verifies that $\check{\mathcal{E}} \colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ is the smallest expectation, which extends $\mathcal{E}$. However, convexity of $\mathcal{E}$ usually does not carry over to $\check{\mathcal{E}}$.

Lemma 3.5. *Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$ and let $\mathcal{E} \colon M \to \mathbb{R}$ be a translation invariant pre-expectation on M. Let*

$$\widehat{\mathcal{A}}_{\mathcal{E}} := \{X \in \mathcal{L}^\infty(\Omega, \mathcal{F}) \colon X_0 \in \mathcal{A}_{\mathcal{E}} \text{ for all } X_0 \in M \text{ with } X_0 \geq X\}.$$

Then, $\widehat{\mathcal{A}}_{\mathcal{E}} = \mathcal{A}_{\widehat{\mathcal{E}}}$, i.e. $\widehat{\mathcal{A}}_{\mathcal{E}}$ is the acceptance set of $\widehat{\mathcal{E}}$, and therefore,

$$\widehat{\mathcal{E}}(X) = \sup\left\{\alpha \in \mathbb{R} \colon X - \alpha \in \widehat{\mathcal{A}}_{\mathcal{E}}\right\} \tag{3.2}$$

for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$. Thus, (3.2) provides a second extension procedure for $\mathcal{E}$, which is extending $\mathcal{E}$ via its acceptable positions $\mathcal{A}_\mathcal{E}$.

Proof. Let $X \in \mathcal{A}_{\widehat{\mathcal{E}}}$, i.e. $\widehat{\mathcal{E}}(X) \geq 0$. By definition of $\widehat{\mathcal{E}}$, it follows that $\mathcal{E}(X_0) \geq 0$ for all $X_0 \in M$ with $X_0 \geq X$, i.e. $X \in \widehat{\mathcal{A}}_\mathcal{E}$. Now assume that $X \in \widehat{\mathcal{A}}_\mathcal{E}$, i.e. $\mathcal{E}(X_0) \geq 0$ for all $X_0 \in M$ with $X_0 \geq X$. Then,

$$\widehat{\mathcal{E}}(X) = \inf\{\mathcal{E}(X_0) \colon X_0 \in M,\ X_0 \geq X\} \geq 0,$$

i.e. $X \in \mathcal{A}_{\widehat{\mathcal{E}}}$. Now, (3.2) follows directly from Proposition 2.8. □

Although Proposition 3.3 implies the existence of an extension $\widehat{\mathcal{E}}$ for every pre-expectation $\mathcal{E} \colon M \to \mathbb{R}$, this extension is not neccessarily unique. However, as translation invariant pre-expectations are 1-Lipschitz by Lemma 2.5 a), the extension is uniquely determined on the closure $\overline{M}$ of M for translation invariant pre-expectations.

Proposition 3.6. *Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E} \colon M \to \mathbb{R}$ be a translation invariant pre-expectation. Then, there exists exactly one translation invariant pre-expectation $\widehat{\mathcal{E}} \colon \overline{M} \to \mathbb{R}$ with $\widehat{\mathcal{E}}|_M = \mathcal{E}$. Here, $\overline{M}$ denotes the closure of M as a subset of $\mathcal{L}^\infty(\Omega, \mathcal{F})$ and $\widehat{\mathcal{E}}$ is given as in Proposition 3.3. If $\mathcal{E}$ is convex or sublinear, then $\widehat{\mathcal{E}}$ is convex or sublinear, respectively.*

Proof. As $M + \mathbb{R} \subset M$ it follows that $\overline{M} + \mathbb{R} \subset \overline{M}$. Hence, by Proposition 3.3, there exists a translation invariant pre-expectation $\widehat{\mathcal{E}} \colon \overline{M} \to \mathbb{R}$ with $\widehat{\mathcal{E}}|_M = \mathcal{E}$. Since, by Lemma 2.5 a), every translation invariant pre-expectation on $\overline{M}$ is 1-Lipschitz, it is uniquely determined by its values on M. □

Lemma 3.7. *Let M and N be two linear subspaces of $\mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M \subset N$, and let $\mathcal{E} \colon N \to \mathbb{R}$ a convex pre-expectation. Then,*

$$\big\{\mu \in M' \colon (\mathcal{E}|_M)^*(\mu) < \infty\big\} = \big\{\nu|_M \colon \nu \in \mathcal{P}_\mathcal{E}\big\}$$

and, for all $\mu \in M'$ with $(\mathcal{E}|_M)^*(\mu) < \infty$,

$$(\mathcal{E}|_M)^*(\mu) = \min_{\nu \in \mathcal{P}_\mathcal{E}, \nu|_M = \mu} \mathcal{E}^*(\nu).$$

Proof. By Proposition 2.9 c),

$$\big\{\mu \in M' \colon (\mathcal{E}|_M)^*(\mu) < \infty\big\} \supset \big\{\nu|_M \colon \nu \in N',\ \mathcal{E}^*(\nu) < \infty\big\}$$

and $(\mathcal{E}|_M)^*(\nu|_M) \leq \mathcal{E}^*(\nu)$ for all $\nu \in \mathcal{P}_\mathcal{E}$. Therefore, let $\mu \in M'$ with $(\mathcal{E}|_M)^*(\mu) < \infty$. Then, we have that

$$\mu X \leq \mathcal{E}(X) - \mathcal{E}^*(\mu)$$

for all $X \in M$. Hence, by the extension theorem of Hahn-Banach, there exists a linear functional $\nu \colon N \to \mathbb{R}$ with $\nu|_M = \mu$ and

$$\nu X \leq \mathcal{E}(X) + \mathcal{E}^*(\mu)$$

for all $X \in N$. Therefore, $\mathcal{E}^*(\nu) \leq (\mathcal{E}|_M)^*(\mu)$, i.e. $\nu \in \mathcal{P}_\mathcal{E}$ with $\nu|_M = \mu$ and $\mathcal{E}^*(\nu) = (\mathcal{E}|_M)^*(\mu)$. □

We apply Lemma 3.7 to the case $N = \mathcal{L}^\infty(\Omega, \mathcal{F})$ and obtain the following corollary.

Corollary 3.8. *Let M be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ and $\mathcal{E}\colon M\to\mathbb{R}$ be a convex pre-expectation. If $\widetilde{\mathcal{E}}\colon \mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ be a convex expectation, which extends $\mathcal{E}$, then,*

$$\mathcal{P}_\mathcal{E}=\big\{\nu|_M\colon \nu\in\mathcal{P}_{\widetilde{\mathcal{E}}}\big\}$$

and, for all $\mu\in\mathcal{P}_\mathcal{E}$,

$$\mathcal{E}^*(\mu)=\min_{\nu\in\mathcal{P}_{\widetilde{\mathcal{E}}},\nu|_M=\mu}\widetilde{\mathcal{E}}^*(\nu).$$

In view of Proposition 2.9 and Corollary 3.8, another natural approach to extend a convex pre-expectation $\mathcal{E}$ on M would be to consider $\widehat{\mathcal{P}}_\mathcal{E}:=\{\nu\in\mathrm{ba}^1_+(\Omega,\mathcal{F})\colon \nu|_M\in\mathcal{P}_\mathcal{E}\}$, and to define $\widetilde{\mathcal{E}}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ by

$$\widetilde{\mathcal{E}}(X):=\sup_{\nu\in\widehat{\mathcal{P}}_\mathcal{E}}\nu X-\mathcal{E}^*(\nu|_M)\tag{3.3}$$

for all $X\in\mathcal{L}^\infty(\Omega,\mathcal{F})$. By Proposition 2.9 c) and Theorem 3.2, $\widetilde{\mathcal{E}}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ is an expectation with $\widetilde{\mathcal{E}}|_M=\mathcal{E}$. In the following proposition, we will prove that $\widehat{\mathcal{E}}=\widetilde{\mathcal{E}}$ and that the supremum in (3.3) is attained.

Proposition 3.9. *Let M be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ and $\mathcal{E}\colon M\to\mathbb{R}$ be a convex pre-expectation. Further, let*

$$\widehat{\mathcal{P}}_\mathcal{E}:=\{\nu\in\mathrm{ba}^1_+(\Omega,\mathcal{F})\colon \nu|_M\in\mathcal{P}_\mathcal{E}\},$$

Then, $\widehat{\mathcal{P}}_\mathcal{E}=\mathcal{P}_{\widehat{\mathcal{E}}}$ and $\widehat{\mathcal{E}}^*(\nu)=\mathcal{E}^*(\nu|_M)$ for all $\nu\in\widehat{\mathcal{P}}_\mathcal{E}$. In particular,

$$\widehat{\mathcal{E}}(X)=\max_{\nu\in\widehat{\mathcal{P}}_\mathcal{E}}\nu X-\mathcal{E}^*(\nu|_M)\quad (X\in\mathcal{L}^\infty(\Omega,\mathcal{F})).$$

Proof. By Corollary 3.8, $\mathcal{P}_{\widehat{\mathcal{E}}}\subset\widehat{\mathcal{P}}_\mathcal{E}$ and $\widehat{\mathcal{E}}^*(\nu)\geq\mathcal{E}^*(\nu|_M)$ for all $\nu\in\mathcal{P}_{\widehat{\mathcal{E}}}$. Proposition 2.9 c), thus implies that $\widehat{\mathcal{P}}_\mathcal{E}=\mathcal{P}_{\widehat{\mathcal{E}}}$ with $\widehat{\mathcal{E}}^*(\nu)=\mathcal{E}^*(\nu|_M)$ for all $\nu\in\widehat{\mathcal{P}}_\mathcal{E}$ □

Corollary 3.10. *Let M be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ and $\mathcal{E}\colon M\to\mathbb{R}$ be a sublinear pre-expectation. Then, $\widehat{\mathcal{P}}_\mathcal{E}=\mathcal{P}_{\widehat{\mathcal{E}}}$, and therefore,*

$$\widehat{\mathcal{E}}(X)=\max_{\nu\in\widehat{\mathcal{P}}_\mathcal{E}}\nu X\quad\textit{for all } X\in\mathcal{L}^\infty(\Omega,\mathcal{F}).$$

Corollary 3.11. *Let M be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ and $\mathcal{E}\colon M\to\mathbb{R}$ be a sublinear pre-expectation. Then, there exists a convex weak* compact set $\mathcal{P}\subset\mathrm{ba}^1_+(\Omega,\mathcal{F})$ such that*

$$\mathcal{E}(X)=\max_{\mu\in\mathcal{P}}\mu X\quad\textit{for all } X\in M.$$

We return to the setting of Theorem 3.2. For a given linear pre-expectation $\mu\colon M\to\mathbb{R}$ on M, we consider the sublinear expectation $\widehat{\mu}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ which extends μ. Then, for every $\nu\in M'$ we have that $\mu^*(\nu)=0$ if and only if $\nu=\mu$. With the previous results, we therefore obtain the following corollary, which states that $\widehat{\mu}$ is the pointwise maximum of all linear expectations $\nu\in\mathrm{ba}^1_+(\Omega,\mathcal{F})$ that extend μ.

Corollary 3.12. *Let M be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ and $\mu\colon M\to\mathbb{R}$ be a linear pre-expectation. Further, let*

$$\widehat{\mathcal{P}}:=\{\nu\in \mathrm{ba}_+^1(\Omega,\mathcal{F})\colon \nu|_M=\mu\}.$$

Then, $\widehat{\mathcal{P}}=\{\nu\in\mathrm{ba}(\Omega,\mathcal{F})\colon \widehat{\mu}^*(\nu)=0\}$, and therefore,

$$\widehat{\mu}(X)=\max_{\nu\in\widehat{\mathcal{P}}}\nu X\quad \text{for all } X\in\mathcal{L}^\infty(\Omega,\mathcal{F}).$$

Let $\mathcal{E}\colon M\to\mathbb{R}$ be a convex pre-expectation. Considering the sublinear expectation $\widehat{\mu}$, which extends $\mu\in\mathcal{P}_\mathcal{E}$, one could also think of

$$\widetilde{\mathcal{E}}(X):=\sup_{\mu\in\mathcal{P}_\mathcal{E}}\widehat{\mu}(X)-\mathcal{E}^*(\mu),\quad \text{for all } X\in\mathcal{L}^\infty(\Omega,\mathcal{F}),\tag{3.4}$$

as another possible extension of $\mathcal{E}$. Clearly, we have that $\widetilde{\mathcal{E}}|_M=\mathcal{E}$ and therefore, $\widetilde{\mathcal{E}}(\alpha)=\alpha$ for all $\alpha\in\mathbb{R}$. Moreover, as $\widehat{\mu}$ is monotone for all $\mu\in\mathcal{P}_\mathcal{E}$, we also have that $\widetilde{\mathcal{E}}$ is monotone. Hence, $\widetilde{\mathcal{E}}$ is an expectation which extends $\mathcal{E}$. In the following proposition, we will use Corollary 3.12 to show that the expectation $\widetilde{\mathcal{E}}$ coincides with $\widehat{\mathcal{E}}$ and that the supremum in (3.4) is attained.

Proposition 3.13. *Let M be a linear subspace of $\mathcal{L}^\infty(\Omega,\mathcal{F})$ with $\mathbb{R}\subset M$ and $\mathcal{E}\colon M\to\mathbb{R}$ be a convex pre-expectation. Then,*

$$\widehat{\mathcal{E}}(X)=\max_{\mu\in\mathcal{P}_\mathcal{E}}\widehat{\mu}(X)-\mathcal{E}^*(\mu)\quad \text{for all } X\in\mathcal{L}^\infty(\Omega,\mathcal{F}).$$

Proof. Let $X\in\mathcal{L}^\infty(\Omega,\mathcal{F})$ and $\widetilde{\mathcal{E}}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ be given by (3.4). We have already seen that $\widetilde{\mathcal{E}}\colon\mathcal{L}^\infty(\Omega,\mathcal{F})\to\mathbb{R}$ is an expectation that extends $\mathcal{E}$. Remark 3.4 thus yields that $\widetilde{\mathcal{E}}(X)\le\widehat{\mathcal{E}}(X)$. By Proposition 3.9, there exists a linear expectation $\nu\in\mathrm{ba}_+^1(\Omega,\mathcal{F})$ which satisfies $\mu:=\nu|_M\in\mathcal{P}_\mathcal{E}$ and $\widehat{\mathcal{E}}(X)=\nu X-\mathcal{E}^*(\mu)$. Hence, by Corollary 3.12, we get that

$$\widehat{\mathcal{E}}(X)=\nu X-\mathcal{E}^*(\mu)\le\widehat{\mu}(X)-\mathcal{E}^*(\mu)\le\widetilde{\mathcal{E}}(X)\le\widehat{\mathcal{E}}(X).$$

This shows that $\widehat{\mathcal{E}}(X)=\nu X-\mathcal{E}^*(\mu)=\widehat{\mu}(X)-\mathcal{E}^*(\mu)$. ☐

Remark 3.14. Let $\mathcal{E}\colon M\to\mathbb{R}$ be a convex pre-expectation on M. In Lemma 3.5, Proposition 3.3, Proposition 3.9 and Proposition 3.13, we have seen that the following four extension procedures all lead to the same (maximal) expectation extending $\mathcal{E}$.

(i) $\mathcal{E}\mapsto\widehat{\mathcal{E}}$,
(ii) $\mathcal{E}\mapsto\mathcal{A}_\mathcal{E}\mapsto\widehat{\mathcal{A}}_\mathcal{E}\mapsto\Big[X\mapsto\sup\big\{\alpha\in\mathbb{R}\colon X-\alpha\in\widehat{\mathcal{A}}_\mathcal{E}\big\}\Big]$,
(iii) $\mathcal{E}\mapsto(\mathcal{P},\mathcal{E}^*)\mapsto\big(\widehat{\mathcal{P}},[\nu\mapsto\mathcal{E}^*(\nu|_M)]\big)\mapsto\big[X\mapsto\max_{\nu\in\widehat{\mathcal{P}}}\nu X-\mathcal{E}^*(\nu|_M)\big]$,
(iv) $\mathcal{E}\mapsto(\mathcal{P},\mathcal{E}^*)\mapsto\big(\{\widehat{\mu}\colon\mu\in\mathcal{P}\},\mathcal{E}^*\big)\mapsto[X\mapsto\max_{\mu\in\mathcal{P}}\widehat{\mu}X-\mathcal{E}^*(\mu)]$.

Example 3.15. Let $\Omega:=\mathbb{R}$, $\mathcal{F}$ be the Borel σ-algebra on $\mathbb{R}$, and $M:=C_b(\Omega)$ be the space of all bounded continuous functions $\Omega\to\mathbb{R}$. Let δ_0 be the Dirac measure with center 0, $\mathcal{E}:=\delta_0|_M$ be the restriction of the Dirac measure to M, and $X:=1_{(0,\infty)}$. Then,

$$\widehat{\mathcal{E}}(X)=\inf\{Y\in M\colon Y\ge X\}=1,$$

while $\delta_0(X) = 0$. This shows that $\mathcal{P}_{\widehat{\mathcal{E}}} \neq \delta_0$, i.e. $\mathcal{E}$ admits several extensions in terms of finitely additive measures.

The previous examples shows that the extension $\widehat{\mathcal{E}}$, already in the linear case, ususally does not reslut in uniqueness of any kind. In order to obtain uniqueness, one therefore needs additional continuity properties of the expectation as an analogue of σ-additivity in the linear case

Definition 3.16.

a) Let $M \subset \mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation. Then, we say that $\mathcal{E}$ is *continuous from above* or *below* if $\mathcal{E}(X_n) \to X$ as $n \to \infty$ for all sequences $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \geq X_{n+1}$ or $X_n \leq X_{n+1}$ for all $n \in \mathbb{N}$ and $X := \lim_{n\to\infty} X_n \in M$, respectively.

b) Let $\mathcal{E}\colon \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$ be a convex expectation. Assume that there exists a set $\mathcal{P}$ of probability measures, i.e. linear expectations, which are continuous from above, and a function $\rho\colon \mathcal{P} \to \mathbb{R}$ with $\inf_{\mu\in\mathcal{P}} \rho(\mu) = 0$ such that

$$\mathcal{E}(X) = \sup_{\mu\in\mathcal{P}} \mu X - \rho(\mu), \quad \text{for all } X \in \mathcal{L}^\infty(\Omega, \mathcal{F}),$$

then, we say that $(\Omega, \mathcal{F}, \mathcal{E})$ is a *convex expectation space*. If $\mathcal{E}$ is sublinear or linear, we say that $(\Omega, \mathcal{F}, \mathcal{E})$ is a *sublinear* or *linear expectation space*, respectively.

Notice that every linear expectation space is a probability space and vice versa.

Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation on a linear subspace M of $\mathcal{L}^\infty(\Omega, \mathcal{F})$ with $\mathbb{R} \subset M$. Then, one can show that the following statements are equivalent:

(i) $\mathcal{E}$ is continuous from above.
(ii) Every $\mu \in \mathcal{P}_{\mathcal{E}}$ is continuous from above.

Remark 3.17. Assume that M is a Riesz subspace of $\mathcal{L}^\infty(\Omega, \mathcal{F})$, i.e. a linear subspace of $\mathcal{L}^\infty(\Omega, \mathcal{F})$ with $X \vee Y \in M$ and $X \wedge Y \in M$ for all $X, Y \in M$, with $\mathbb{R} \subset M$. Then, by the Daniell-Stone Theorem, every linear pre-expectation on M, which is continuous from above, can uniquely be extended to a linear expectation $\bar{\mu}$ on $\mathcal{L}^\infty\big(\Omega, \sigma(M)\big)$, which is continuous from above. Now, assume that $\mathcal{E}$ is a convex expectation, and assume that there exists a set $\mathcal{P} \subset M'$ of linear pre-expectations, which are continuous from above, and a map $\rho\colon \mathcal{P} \to \mathbb{R}$ with $\inf_{\mu\in\mathcal{P}} \rho(\mu) = 0$. Then,

$$\bar{\mathcal{E}}(X) := \sup_{\mu\in\mathcal{P}} \bar{\mu} X - \rho(\mu), \quad \text{for } X \in \mathcal{L}^\infty\big(\Omega, \sigma(M)\big),$$

defines a convex expectation, which extends $\mathcal{E}$ to $\mathcal{L}^\infty\big(\Omega, \sigma(M)\big)$ (see Bartl et al. [3]). If $\mathcal{E}\colon M \to \mathbb{R}$ is continuous from above, using a modification of Choquet's capacibility theorem [7], one can show that

$$\bar{\mathcal{E}}(X) := \sup\left\{\inf_{n\in\mathbb{N}} \mathcal{E}(X_n)\colon (X_n)_{n\in\mathbb{N}} \in M^{\mathbb{N}}, X_n \geq X_{n+1}\,(n\in\mathbb{N}), X \geq \inf_{n\in\mathbb{N}} X_n\right\}$$

for $\mathcal{P} = \mathcal{P}_{\mathcal{E}}$ and $\rho := \mathcal{E}^*$, and that uniqueness within a certain class of expectations can be achieved. We refer to Bartl [2] and Denk et al. [16] for more details on the uniqueness of this extension.

4. Stochastic processes under nonlinear expectation

In this section, we apply the extension results of the previous chapter to a Kolmogorov-type setting. That is, given a consistent family of finite-dimensional marginal expectations, we are looking for an expectation on a suitable path space with these marginals.

Throughout this section, let $I \neq \emptyset$ be an index set, $\mathscr{H} := \{J \subset I \colon |J| \in \mathbb{N}\}$ the set of all finite, nonempty subsets of I and $(S, \mathscr{S})$ be a measurable (state) space. For each $J \in \mathscr{H}$ let $M_J \subset \mathcal{L}^\infty(S^J, \mathcal{B}^J)$ be a linear subspace with $1 \in M_J$, where $\mathcal{B}^J$ is the product σ-algebra on S^J. For all $K \subset J \subset I$ let

$$\operatorname{pr}_{JK} \colon S^J \to S^K, \quad (x_i)_{i\in J} \mapsto (x_i)_{i\in K}$$

and $\operatorname{pr}_J := \operatorname{pr}_{IJ}$. Throughout this section, we assume that

$$M_K \circ \operatorname{pr}_{JK} := \{f \circ \operatorname{pr}_{JK} \colon f \in M_K\} \subset M_J$$

for all $J, K \in \mathscr{H}$ with $K \subset J$. Typical examples for the family $(M_J)_{J\in\mathscr{H}}$ are:

(i) the space $\mathcal{L}^\infty(S^J) := \mathcal{L}^\infty(S^J, \mathcal{B}^J)$ of all bounded $\mathcal{B}^J$-$\mathcal{B}(\mathbb{R})$-measurable functions, where $\mathcal{B}^J$ denotes the product σ-algebra on S^J,
(ii) the space $C_b(S^J)$ of all bounded continuous functions $S^J \to \mathbb{R}$, where S^J is endowed with the product topology,
(iii) the space $\mathrm{BUC}(S^J)$ of all bounded uniformly continuous functions $S^J \to \mathbb{R}$ w.r.t. a fixed metric, which generates the topology on S.

Let $J, K \in \mathscr{H}$ with $K \subset J$. For a pre-expectation $\mathcal{E}_J \colon M_J \to \mathbb{R}$ we then denote by $\mathcal{E}_J \circ \operatorname{pr}_{JK}^{-1}$ the restriction of the distribution (see Remark 2.15) of $\mathcal{E}_J$ under $\operatorname{pr}_{JK}^{-1}$ to M_K, i.e.

$$\mathcal{E}_J \circ \operatorname{pr}_{JK}^{-1} \colon M_K \to \mathbb{R}, \quad f \mapsto \mathcal{E}_J(f \circ \operatorname{pr}_{JK}).$$

In [29], Peng defines a consistency condition for nonlinear expectations and proves an extension to the subspace

$$M := \{f \circ \operatorname{pr}_J \colon J \in \mathscr{H}, f \in \mathcal{L}^\infty(S^J, \mathcal{B}^J)\}$$

of $\mathcal{L}^\infty(S^I, \mathcal{B}^I)$. Here $\mathcal{B}^I$ denotes the product σ-algebra of $\mathcal{B}$, i.e. the σ-algebra generated by the sets of the form $\operatorname{pr}_J^{-1}(B_J)$ with $J \in \mathscr{H}$ and $B_J \in \mathcal{B}^J$. In the sequel, we use the same notion of consistency as Peng and apply the extension results from the previous section in order to obtain an extension to $\mathcal{L}^\infty(S^I, \mathcal{B}^I)$.

Definition 4.1. For all $J \in \mathscr{H}$ let $\mathcal{E}_J \colon M_J \to \mathbb{R}$ be a pre-expectation. Then, the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is called *consistent* if, for all $J, K \in \mathscr{H}$ with $K \subset J$,

$$\mathcal{E}_K(f) = \mathcal{E}_J(f \circ \operatorname{pr}_{JK}) \quad \text{for all } f \in M_K,$$

i.e. if $\mathcal{E}_K = \mathcal{E}_J \circ \operatorname{pr}_{JK}^{-1}$.

Remark 4.2. For all $J \in \mathscr{H}$ let $\mathcal{E}_J \colon M_J \to \mathbb{R}$ be a pre-expectation. Then, the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent if and only if

$$\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$$

for all $J, K \in \mathscr{H}$ with $K \subset J$ and $|J| = |K| + 1$. In fact, assume that $\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$ for all $J, K \in \mathscr{H}$ with $K \subset J$ and $|J| = |K| + 1$. We prove that

$$\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$$

for all $J \in \mathscr{H}$ with $K \subset J$ by induction on $n = |J| - |K| \in \mathbb{N}_0$. For $n = 0$ the statement is trivial. Now, assume that there exists some $n \in \mathbb{N}_0$ such that

$$\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$$

for all $J \in \mathscr{H}$ with $K \subset J$ and $|J| = |K| + n$. Let $J \in \mathscr{H}$ with $|J| = |K| + n + 1$, $J' := J \setminus \{i\}$ for some $i \in J \setminus K$ and $f \in M_K$. Then, we have that $g := f \circ \mathrm{pr}_{J'K} \in M_{J'}$ with $g \circ \mathrm{pr}_{JJ'} = f \circ \mathrm{pr}_{JK}$. Therefore, by the induction hypothesis, we have that

$$\mathcal{E}_J(f \circ \mathrm{pr}_{JK}) = \mathcal{E}_J(g \circ \mathrm{pr}_{JJ'}) = \mathcal{E}_{J'}(g) = \mathcal{E}_{J'}(f \circ \mathrm{pr}_{J'K}) = \mathcal{E}_K(f).$$

The following theorem due to Denk et al. [16] is a finitely additive and nonlinear version of Kolmogorov's extension theorem.

Theorem 4.3. *Let $(\mathcal{E}_J)_{J\in\mathscr{H}}$ be a consistent family of pre-expectations $\mathcal{E}_J \colon M_J \to \mathbb{R}$. Then, there exists an expectation $\widehat{\mathcal{E}} \colon \mathcal{L}^\infty(S^I, \mathcal{B}^I) \to \mathbb{R}$ such that*

$$\widehat{\mathcal{E}}(f \circ \mathrm{pr}_J) = \mathcal{E}_J(f) \quad \text{for all } J \in \mathscr{H} \text{ and all } f \in M_J.$$

If the pre-expectations $\mathcal{E}_J$ are translation invariant, convex or sublinear for all $J \in \mathscr{H}$, then $\widehat{\mathcal{E}}$ is translation invariant, convex or sublinear, respectively.

Proof. Let $M := \{f \circ \mathrm{pr}_J \colon f \in M_J,\ J \in \mathscr{H}\}$. Then M is a linear subspace of $\mathcal{L}^\infty(S^I, \mathcal{B}^I)$ with $1 \in M$. For every $J \in \mathscr{H}$ and $f \in M_J$ let $\mathcal{E}(f \circ \mathrm{pr}_J) := \mathcal{E}_J(f)$. Since the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent, the functional $\mathcal{E} \colon M \to \mathbb{R}$ is well-defined. Moreover, $\mathcal{E} \colon M \to \mathbb{R}$ is a pre-expectation on M. The assertion now follows from Proposition 3.3. ☐

In Theorem 4.3, we proved the existence of an extension without any continuity properties and any structural assumptions. The following theorem due to Denk et al. [16] can be viewed as a continuous and convex version of Theorem 4.3.

Theorem 4.4. *Let S be a Polish space and $\mathscr{S}$ be the Borel σ-algebra on S. For all $J \in \mathscr{H}$ let M_J be a Riesz subspace of $\mathcal{L}^\infty(S^J, \mathcal{B}^J)$ with $\sigma(M_J) = \mathcal{B}^J$ and $\mathcal{E}_J \colon M_J \to \mathbb{R}$ be a convex pre-expectation, which is continuous from above. If the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent, then there exists a convex expectation space $(S^I, \mathcal{B}^I, \bar{\mathcal{E}})$ with*

$$\mathcal{E}_J(f) = \bar{\mathcal{E}}(f \circ \mathrm{pr}_J) \quad \text{for all } J \in \mathscr{H} \text{ and } f \in M_J.$$

If the pre-expectations $\mathcal{E}_J$ are sublinear or linear for all $J \in \mathscr{H}$, then $\bar{\mathcal{E}}$ is sublinear or linear, respectively.

Proof. Let $M := \{f\circ \mathrm{pr}_J\colon f\in M_J,\ J\in\mathscr{H}\}$, and define $\mathcal{E}(f\circ\mathrm{pr}_J) := \mathcal{E}_J(f)$ for all $f\in M_J$ and all $J\in\mathscr{H}$. Since the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent, $\mathcal{E}\colon M\to\mathbb{R}$ defines a convex pre-expectation on M. Let $\mu\in M'$ with $\mathcal{E}^*(\mu)<\infty$. We will first show that $\mu\colon M\to\mathbb{R}$ is continuous from above. Let $\mu_J := \mu\circ\mathrm{pr}_J^{-1}$ for all $J\in\mathscr{H}$. Then, $\mathcal{E}_J^*(\mu_J)\le\mathcal{E}^*(\mu)<\infty$, and therefore, $\mu_J\colon M_J\to\mathbb{R}$ is continuous from above. By the theorem of Daniell-Stone, there exists a unique $\nu_J\in\mathrm{ca}_+^1(S^J,\mathcal{B}^J)$ with $\nu_J|_{M_J}=\mu_J$ for all $J\in\mathscr{H}$. Let $J,K\in\mathscr{H}$ with $K\subset J$ and $f\in M_K$. Then,

$$\mu_K f = \mu_J(f\circ\mathrm{pr}_{JK}) = \nu_J(f\circ\mathrm{pr}_{JK})\quad\text{for all } f\in M_K$$

and therefore,

$$\nu_K f = \nu_J(f\circ\mathrm{pr}_{JK})\quad\text{for all } f\in\mathcal{L}^\infty(S^K,\mathcal{B}^K)$$

as the extension of μ_K to a probability measure is unique. By Kolmogorov's extension theorem, there exists a unique $\nu\in\mathrm{ca}_+^1(S^I,\mathcal{B}^I)$ with $\nu(f\circ\mathrm{pr}_J)=\nu_J f$ for all $J\in\mathscr{H}$ and $f\in\mathcal{L}^\infty(S^J,\mathcal{B}^J)$. Hence, we get that $\nu|_M=\mu$, which implies that $\mu\colon M\to\mathbb{R}$ is continuous from above as well. Using Remark 3.17, we thus obtain that there exists an expectation $\bar{\mathcal{E}}\colon\mathcal{L}^\infty(\Omega,\sigma(M))$, which extends $\mathcal{E}$ and results in a convex expectation space $\big(S^I,\sigma(M),\bar{\mathcal{E}}\big)$,
It remains to show that $\mathcal{B}^I=\sigma(M)$. It is clear that $\mathcal{B}^I\supset\sigma(M)$. In order to show the inverse inclusion, let $J\in\mathscr{H}$ and $B_J\in\mathcal{B}^J$. Then, by assumption, $B_J\in\sigma(M_J)$ and therefore, we obtain that $\mathrm{pr}_J^{-1}(B_J)\in\sigma(M_J\circ\mathrm{pr}_J)\subset\sigma(M)$. □

In the situation of Theorem 4.4, considering the canonical process $(\mathrm{pr}_{\{i\}})_{i\in I}$ on the convex expectation space $\big(S^I,\mathcal{B}^I,\overline{\mathcal{E}}\big)$, we obtain the following two corollaries on the existence of stochastic processes under nonlinear expectations.

Corollary 4.5. *Let S be a Polish space and $\mathscr{S}$ be the Borel σ-algebra on S. For all $J\in\mathscr{H}$ let M_J be a Riesz subspace of $\mathcal{L}^\infty(S^J,\mathcal{B}^J)$ with $\sigma(M_J)=\mathcal{B}^J$ and $\mathcal{E}_J\colon M_J\to\mathbb{R}$ be a convex pre-expectation which is continuous from above. Then, the following two statements are equivalent:*

(i) *The family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent,*
(ii) *There exists a convex expectation space $(\Omega,\mathcal{F},\mathcal{E})$ and a stochastic process $(X_i)_{i\in I}$ with*
$$\mathcal{E}(f(X_J))=\mathcal{E}_J(f)$$
for all $J\in\mathscr{H}$ and $f\in M_J$, where $X_J:=(X_i)_{i\in J}$.

Corollary 4.6. *Let S be a Polish space and $\mathscr{S}$ be the Borel σ-algebra on S. For all $J\in\mathscr{H}$ let M_J be a Riesz subspace of $\mathcal{L}^\infty(S^J,\mathcal{B}^J)$ with $\sigma(M_J)=\mathcal{B}^J$ and $\mathcal{E}_J\colon M_J\to\mathbb{R}$ be a sublinear pre-expectation which is continuous from above. Then, the following two statements are equivalent:*

(i) *The family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent,*
(ii) *There exists a sublinear expectation space $(\Omega,\mathcal{F},\mathcal{E})$ and a stochastic process $(X_i)_{i\in I}$ with*
$$\mathcal{E}(f(X_J))=\mathcal{E}_J(f)$$
for all $J\in\mathscr{H}$ and $f\in M_J$, where $X_J:=(X_i)_{i\in J}$.

We conclude this section with the following example on discrete-time Markov chains, which is taken from Denk et al. [16].

Example 4.7. Let S be a finite state space and $\mathscr{S} := 2^S$, so that $\mathcal{L}^\infty(S,\mathcal{B}) = \mathbb{R}^S$. Let

$$\mathcal{P}\colon \mathcal{L}^\infty(S,\mathcal{B}) \to \mathcal{L}^\infty(S,\mathcal{B}) \quad \textit{and} \quad \mu_0\colon \mathcal{L}^\infty(S,\mathcal{B}) \to \mathbb{R}$$

be convex, constant preserving, i.e. $\mathcal{P}(\alpha) = \alpha$ and $\mu_0(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$, and monotone, i.e. $\mathcal{P}(f) \leq \mathcal{P}(g)$ and $\mu_0(f) \leq \mu_0(g)$ for all $f, g \in \mathcal{L}^\infty(S,\mathcal{B})$ with $f \leq g$. For every $k, l \in \mathbb{N}_0$ with $k < l$, we define

$$\mathcal{E}_{k,l}(\,\cdot\,, f) := \mathcal{P}^{l-k}(f) \quad \text{for all } f \in \mathcal{L}^\infty(S,\mathcal{B}).$$

Let $\mathscr{H} := \{J \subset \mathbb{N}_0 \colon \#J \in \mathbb{N}\}$ be the set of all finite, nonempty subsets of $\mathbb{N}_0$. For $k \in \mathbb{N}_0$ we define

$$\mathcal{E}_{\{k\}}(f) := \mu_0(\mathcal{P}^k(f)) \quad \text{for all } f \in \mathcal{L}^\infty(S,\mathcal{B}),$$

where $\mathcal{P}^0$ is the identity. For $n \in \mathbb{N}$, $k_1, \ldots, k_{n+1} \in \mathbb{N}_0$ with $k_1 < \ldots < k_{n+1}$ and $f \in \mathcal{L}^\infty(S^{n+1}, \mathcal{B}^{n+1})$ we now define recursively

$$\mathcal{E}_{\{k_1,\ldots,k_{n+1}\}}(f) := \mathcal{E}_{\{k_1,\ldots,k_n\}}(g)$$

where $g(x_1, \ldots, x_n) := \mathcal{E}_{k_n,k_{n+1}}(x_n, f(x_1, \ldots, x_n, \cdot\,))$ for all $x_1, \ldots, x_n \in S$. Then, $\mathcal{E}_J\colon \mathcal{L}^\infty(S^J, \mathcal{B}^J) \to \mathbb{R}$ is a convex expectation and, therefore, continuous from above since S^J is finite for all $J \in \mathscr{H}$. By Remark 4.2, the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent. Hence, Theorem 4.4 or, more precisely, Corollary 4.5 implies that there exists a convex expectation space $(\Omega, \mathcal{F}, \mathcal{E})$ and a stochastic process $(X_i)_{i\in\mathbb{N}_0}$ with

$$\mathcal{E}(f(X_J)) = \mathcal{E}_J(f)$$

for all $J \in \mathscr{H}$ and $f \in M_J$, where $X_J := (X_i)_{i\in J}$. The process $(X_i)_{i\in\mathbb{N}_0}$ can be viewed as a convex version of a discrete-time Markov chain. If $\mathcal{P}$ is sublinear, the set

$$\left\{\mu \in \mathbb{R}^{S\times S} \colon \mu f \leq \mathcal{P}(f) \text{ for all } f \in \mathbb{R}^S\right\}$$

induces a Markov-set chain, see Hartfiel [22], and the operator $\mathcal{P}$ coincides with the concept of a *conditional coherent upper prevision* or *upper transition operator*, cf. de Cooman et al. [12] and Škulj [34].

5. Continuous-time Markov chains under nonlinear expectation

In this section, we consider time-homogeneous continuous-time Markov chains with a countable state space S (endowed with the discrete topology 2^S). We identify (measurable) functions $S \to \mathbb{R}$ via sequences of the form $(u_i)_{i\in S} \in \mathbb{R}^S$, and use the notation $\ell^\infty := \mathcal{L}^\infty\big(S, 2^S\big)$. We call a (possibly nonlinear) map $\mathcal{P}\colon \ell^\infty \to \ell^\infty$ a *kernel* if

(i) $\mathcal{P}u \leq \mathcal{P}v$ for all $u, v \in \ell^\infty$,
(ii) $\mathcal{P}\alpha = \alpha$ for all $\alpha \in \mathbb{R}$.

Here and throughout, we use the notation $\mathcal{P}u = \mathcal{P}(u)$ for a kernel $\mathcal{P}$ and $u \in \ell^\infty$. In this section, we consider an arbitrary familiy of linear Markov chains and construct a sublinear Markov chain, which can be interpreted as an imprecise version of the familiy of Markov chains. The proofs and statements in this section rely on an approach proposed by Nisio [26], which has been applied in various settings in order to construct Markov processes under model uncertainty, see e.g. Denk et al. [17], Nendel [24], and Nendel and Röckner [25].

Throughout, we consider the following setup: Let Λ be a nonempty index set. For each $\lambda \in \Lambda$ let $S_\lambda = \big(S_\lambda(t)\big)_{t\geq 0}$ be the family of transition probabilites of a time-homogeneous Markov chain. i.e.

(i) $S_\lambda(t)_{ij} \geq 0$ for all $t \geq 0$ and $i, j \in S$,
(ii) $\sum_{j\in S} S_\lambda(t)_{ij} = 1$ for all $t \geq 0$ and $i \in S$,
(iii) $S_\lambda(0) = I$, where I is the identity,
(iv) $S_\lambda(s)S_\lambda(t) = S_\lambda(s+t)$ for all $s, t \geq 0$.

For the family $(S_\lambda)_{\lambda\in\Lambda}$, we construct a semigroup of sublinear kernels, i.e. a family $\big(S(t)\big)_{t\geq 0}$ of sublinear kernels $\ell^\infty \to \ell^\infty$ with

(i) $S(0) = I$,
(ii) $S(s)S(t)u = S(s+t)u$ for all $s, t \geq 0$ and $u \in \ell^\infty$.

To this end, we consider the set of finite partitions

$$P := \big\{\pi \subset [0,\infty)\colon 0 \in \pi, |\pi| < \infty\big\}.$$

For a partition $\pi \in P$, $\pi = \{t_0, t_1, \ldots, t_m\}$ with $0 = t_0 < t_1 < \ldots < t_m$ we set

$$|\pi|_\infty := \max_{j=1,\ldots,m} (t_j - t_{j-1}).$$

Moreover, we define $|\{0\}|_\infty := 0$. The set of partitions with end-point t will be denoted by P_t, i.e. $P_t := \{\pi \in P\colon \max\pi = t\}$. Note that

$$P = \bigcup_{t\geq 0} P_t.$$

For all $h \geq 0$ and $u \in \ell^\infty$ we define

$$\mathcal{E}_h u := \sup_{\lambda\in\Lambda} S_\lambda(h)u,$$

where the supremum is taken componentwise. Then, $\mathcal{E}_h$ is well-defined since

$$\|S_\lambda(h)u\|_\infty \leq \|u\|_\infty$$

for all $\lambda \in \Lambda$. Moreover, $\mathcal{E}_h$ is a sublinear kernel as it is monotone and $\mathcal{E}_h\alpha = \alpha$ for all $\alpha \in \mathbb{R}$. For a partition $\pi = \{t_0, t_1, \ldots, t_m\} \in P$ with $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_m$, we set

$$\mathcal{E}_\pi := \mathcal{E}_{t_1-t_0} \ldots \mathcal{E}_{t_m-t_{m-1}}.$$

Moreover, we set $\mathcal{E}_{\{0\}} := \mathcal{E}_0$. Then, $\mathcal{E}_\pi$ is a sublinear kernel for all $\pi \in P$ as it is a concatenation of sublinar kernels. We define

$$\mathscr{S}(t)u := \sup_{\pi\in P_t} \mathcal{E}_\pi u$$

for all $u \in \ell^\infty$ and $t \geq 0$, and call $\big(\mathscr{S}(t)\big)_{t\geq 0}$ the *Nisio semigroup* or *(upper) semigroup envelope* of $(S_\lambda)_{\lambda\in\Lambda}$. Note that $\mathscr{S}(t)\colon \ell^\infty \to \ell^\infty$ is well-defined and a sublinear kernel for all $t \geq 0$ since $\mathcal{E}_\pi$ is a sublinear kernel for all $\pi \in P$.

For $h_1, h_2 \geq 0$,

$$\begin{aligned}\mathcal{E}_{h_1+h_2}u &= \sup_{q\in\mathcal{P}} S_\lambda(h_1+h_2)u = \sup_{q\in\mathcal{P}} S_\lambda(h_1)S_\lambda(h_2)u \\ &\leq \sup_{q\in\mathcal{P}} S_\lambda(h_1)\mathcal{E}_{h_2}u = \mathcal{E}_{h_1}\mathcal{E}_{h_2}u,\end{aligned}$$

which implies the inequality

$$\mathcal{E}_{\pi_1}u \leq \mathcal{E}_{\pi_2}u \tag{5.1}$$

for $\pi_1, \pi_2 \in P$ with $\pi_1 \subset \pi_2$. The following lemma shows that $\mathscr{S}(t)$ can be obtained by a pointwise monotone approximation with finite partitions letting the mesh size tend to zero.

Lemma 5.1. *Let $u \in \ell^\infty$ and $t \geq 0$. Then, there exists a sequence $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with $\mathcal{E}_{\pi_n}u \leq \mathcal{E}_{\pi_{n+1}}u$ for all $n \in \mathbb{N}$ and*

$$\mathcal{E}_{\pi_n}u \to \mathscr{S}(t)u \quad \text{as } n \to \infty.$$

Proof. For $t = 0$ the statement is trivial. We therefore assume that $t > 0$. Then, for every $i \in S$, there exists a sequence $(\pi_n^i)_{n\in\mathbb{N}} \subset P_t$ with $\pi_n^i \subset \pi_{n+1}^i$ for all $n \in \mathbb{N}$ and

$$\big(\mathcal{E}_{\pi_n^i}u\big)(i) \to \big(\mathscr{S}(t)u\big)(i) \quad \text{as } n \to \infty.$$

Since S is countable, there exists a sequence $(S_n)_{n\in\mathbb{N}}$ with $S_n \subset S_{n+1} \subset S$ for all $n \in \mathbb{N}$ and $S = \bigcup_{n\in\mathbb{N}} S_n$. Let

$$\pi_n := \bigcup_{i\in S_n} \pi_n^i$$

for all $n \in \mathbb{N}$. Then, $\pi_n^i \subset \pi_n \subset \pi_{n+1}$ for all $n \in \mathbb{N}$ and $i \in S_n$, and therefore,

$$\mathcal{E}_{\pi_n^i} \leq \mathcal{E}_{\pi_n} \leq \mathcal{E}_{\pi_{n+1}}$$

for all $n \in \mathbb{N}$ and $i \in S_n$, which implies that $\mathscr{S}(t)u = \sup_{n\in\mathbb{N}} \mathcal{E}_{\pi_n}u$. ☐

Proposition 5.2. *$\mathscr{S}(s+t) = \mathscr{S}(s)\mathscr{S}(t)$ for all $s, t \geq 0$.*

Proof. Let $u \in \ell^\infty$. If $s = 0$ or $t = 0$ the statement is trivial. Therefore, let $s, t > 0$, $\pi_0 \in P_{s+t}$ and $\pi := \pi_0 \cup \{s\}$. Then, $\pi \in P_{s+t}$ with $\pi_0 \subset \pi$. Hence, by (5.1),

$$\mathcal{E}_{\pi_0}u \leq \mathcal{E}_\pi u.$$

Let $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots t_m = s + t$ with $\pi = \{t_0, \ldots, t_m\}$ and $k \in \{1, \ldots, m\}$ with $t_k = s$. Then, $\pi_1 := \{t_0, \ldots, t_k\} \in P_s$ and $\pi_2 := \{t_k - s, \ldots, t_n - s\} \in P_t$ with

$$\mathcal{E}_{\pi_1} = \mathcal{E}_{t_1-t_0} \cdots \mathcal{E}_{t_i-t_{i-1}}$$

and

$$\mathcal{E}_{\pi_2} = \mathcal{E}_{t_{i+1}-t_i} \cdots \mathcal{E}_{t_m-t_{m-1}}.$$

Therefore,

$$\begin{aligned}\mathcal{E}_{\pi_0}u \le \mathcal{E}_{\pi}u = \mathcal{E}_{t_1-t_0}\cdots\mathcal{E}_{t_m-t_{m-1}}u = \big(\mathcal{E}_{t_1-t_0}\cdots\mathcal{E}_{t_i-t_{i-1}}\big)\big(\mathcal{E}_{t_{i+1}-t_i}\cdots\mathcal{E}_{t_m-t_{m-1}}u\big)\\ = \mathcal{E}_{\pi_1}\mathcal{E}_{\pi_2}u \le \mathcal{E}_{\pi_1}\mathscr{S}(t)u \le \mathscr{S}(s)\mathscr{S}(t)u.\end{aligned}$$

Taking the supremum over all $\pi_0 \in P_{s+t}$, we obtain that $\mathscr{S}(s+t)u \le \mathscr{S}(s)\mathscr{S}(t)u$.

Now, let $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with $\mathcal{E}_{\pi_n}u \le \mathcal{E}_{\pi_{n+1}}u$ for all $n \in \mathbb{N}$ and $\mathcal{E}_{\pi_n}u \to \mathscr{S}(t)u$ as $n \to \infty$ (see Lemma 5.1), and fix $\pi_0 \in P_s$. Then, for all $n \in \mathbb{N}$,

$$\pi_n' := \pi_0 \cup \{s+\tau\colon \tau \in \pi_n\} \in P_{s+t}$$

with $\mathcal{E}_{\pi_n'} = \mathcal{E}_{\pi_0}\mathcal{E}_{\pi_n}$. As $\mathcal{E}_{\pi_0}$ is continuous from below, we obtain that

$$\mathcal{E}_{\pi_0}\big(\mathscr{S}(t)u\big) = \lim_{n\to\infty}\mathcal{E}_{\pi_0}\mathcal{E}_{\pi_n}u = \lim_{n\to\infty}\mathcal{E}_{\pi_n'}u \le \mathscr{S}(s+t)u.$$

Taking the supremum over all $\pi_0 \in P_s$, we see that $\mathscr{S}(s)\mathscr{S}(t)u \le \mathscr{S}(s+t)u$. □

Remark 5.3. The semigroup $\mathscr{S}$ is the smallest semigroup that dominates the family $(S_\lambda)_{\lambda\in\Lambda}$. In fact, let T be an upper bound of the family $(S_\lambda)_{\lambda\in\Lambda}$, i.e.

$$S_\lambda(t)u \le T(t)u$$

for all $\lambda \in \Lambda$, $u \in \ell^\infty$ and $t \ge 0$. Then,

$$S_\lambda(h)u \le \mathcal{E}_h u \le T(h)u$$

for all $\lambda \in \Lambda$, $u \in \ell^\infty$ and $h \ge 0$. Since S_λ and T are semigroups, it follows that

$$S_\lambda(t)u \le \mathcal{E}_\pi u \le T(t)u$$

for all $\lambda \in \Lambda$, $u \in \ell^\infty$, $t \ge 0$ and $\pi \in P_t$. Taking the supremum over all $\pi \in P_t$, we obtain that

$$S_\lambda(t)u \le \mathscr{S}(t)u \le T(t)u$$

for all $\lambda \in \Lambda$, $u \in \ell^\infty$ and $t \ge 0$.

The Nisio semigroup $\big(\mathscr{S}(t)\big)_{t\ge 0}$ has been used by various authors in the field of imprecise probability in order to describe imprecise Markov chains via conditional upper previsions, cf. Škulj [35] and Krak et al. [23]. In fact, the operator $\mathscr{S}(t)$ coincides with the conditional upper prevision over a time interval of the length $t \ge 0$. However, one can actually go one step further and extend the family of transition operators to a Markov chain on a canonical path space. In [24], this has been done in the finite-state-case using the Theorem 4.4, which is due to Denk et al. [16]. However, in the case of infinitely many states, the continuity from above is not trivial and leads to restrictions for the family of Markov chains indexed by the set Λ. Nevertheless, using an explicit dual representation of the Nisio semigroup one can extend the family of transition operators to a Markov chain on the canonical path space without requiring the continuity from above.

In the following remark, we start by deriving a dual representation of the Nisio semigroup by viewing it as the cost functional of an optimal control problem, where, roughly speaking, “nature” tries to control the system into the worst possible scenario (using contols within the set Λ).

Remark 5.4. For $\lambda=(\lambda^1,\ldots\lambda^d)\in\Lambda^d$ and $t\geq 0$, let $S_\lambda(t)\colon \ell^\infty\to\ell^\infty$ be given by

$$\big(S_\lambda(t)u\big)_i:=\big(S_{\lambda^i}(t)u\big)_i \tag{5.2}$$

for all $u_0\in\ell^\infty$ and $i\in S$. That is, $S_\lambda(t)$ is the matrix whose i-th row is the i-th row of $S_{\lambda^i}(t)$ for all $i\in S$. Here, the interpretation is that, in every state $i\in S$, "nature" is allowed to choose a different Markov chain, which is indexed by $\lambda\in\Lambda$. We now add a dynamic component, and define

$$Q_t:=\bigg\{(\lambda_k,h_k)_{k=1,\ldots,m}\in\big(\Lambda^d\times[0,t]\big)^m\colon m\in\mathbb{N},\ \sum_{k=1}^m h_k=t\bigg\}.$$

Roughly speaking, the set Q_t corresponds to the set of all (space-time discrete) admissible controls for the control set Λ. For $\theta=(\lambda_k,h_k)_{k=1,\ldots,m}\in Q_t$ with $m\in\mathbb{N}$ and $u\in\ell^\infty$, we define

$$S_\theta u:=S_{\lambda_1}(h_1)\cdots S_{\lambda_m}(h_m)u,$$

where $S_{\lambda_k}(h_k)$ is defined as in (5.2) for $k=1,\ldots,m$. Then, for all $u\in\ell^\infty$,

$$\mathscr{S}(t)u=\sup_{\pi\in P_t}\mathcal{E}_\pi u=\sup_{\theta\in Q_t}S_\theta u. \tag{5.3}$$

In fact, by definition of Q_t, it follows that $S_\lambda(t)u_0\leq\sup_{\theta\in Q_t}S_\theta u_0\leq\mathscr{S}(t)u_0$ for all $\lambda\in\Lambda$, $t\geq 0$ and $u_0\in\mathbb{R}^d$. On the other hand, one readily verifies that $\mathscr{T}(t)u_0:=\sup_{\theta\in Q_t}S_\theta u_0$, for $t\geq 0$ and $u_0\in\mathbb{R}^d$, gives rise to a semigroup $\big(\mathscr{T}(t)\big)_{t\geq 0}$. Since $\big(\mathscr{S}(t)\big)_{t\geq 0}$ is the smallest semigroup that dominates the family $(S_\lambda)_{\lambda\in\Lambda}$ of $(\mathcal{P},f)$, it follows that $\mathscr{T}(t)=\mathscr{S}(t)$ for all $t\geq 0$.

The explicit dual representation from the previous remark can be used as in [17, Proposition 5.12] in order to obtain a sublinear Markov chain in the following sense:

Definition 5.5. A (time-homogeneous) *sublinear Markov chain* is a quadruple $\big(\Omega,\mathcal{F},\mathcal{E},(X_t)_{t\geq 0}\big)$, where

- *(i)* $(\Omega,\mathcal{F})$ is a measurable space.
- *(ii)* $X_t\colon\Omega\to\mathbb{N}$ is $\mathcal{F}$-measurable for all $t\geq 0$.
- *(iiii)* $\mathcal{E}=(\mathcal{E}_i)_{i\in S}$, where $(\Omega,\mathcal{F},\mathcal{E}_i)$ is a sublinear expectation space for all $i\in S$ with $\mathcal{E}_i\big(u(X_0)\big)=u(i)$ for all $u\in\ell^\infty$.
- *(iv)* For all $s,t\geq 0$, $n\in\mathbb{N}$, $0\leq t_1<\ldots<t_n\leq s$ and $v\in\mathcal{L}^\infty\big(S^{n+1},2^{S^{n+1}}\big)$,

  $$\mathcal{E}\big(v_0(X_{t_1},\ldots,X_{t_n},X_{s+t})\big)=\mathcal{E}\left[\big(S(t)v_0(X_{t_1},\ldots,X_{t_n},\,\cdot\,)\big)(X_s)\right]$$

  with $\big(S(t)u\big)(i):=\mathcal{E}_i\big(u(X_t)\big)$ for all $u\in\ell^\infty$ and $i\in S$. The family of kernels $\big(S(t)\big)_{t\geq 0}$ is called the *transition semigroup* of the Markov chain $\big(\Omega,\mathcal{F},\mathcal{E},(X_t)_{t\geq 0}\big)$.

Remark 5.6. Assume that for all $\lambda\in\Lambda$, there exists an infinite matrix $q^\lambda=(q_{ij}^\lambda)_{i,j\in S}\in\mathbb{R}^{S\times S}$ such that, for all $u\in\ell^\infty$

$$\bigg\|\frac{S_\lambda(h)u-u}{h}-q^\lambda u\bigg\|_\infty\to 0\quad\text{as } h\searrow 0.$$

Then, for each $\lambda\in\Lambda$, q^λ satisfies the following:

(i) $q_{ii} \leq 0$ for all $i \in S$,
(ii) $q_{ij} \geq 0$ for all $i, j \in S$ with $i \neq j$,
(iii) $\sum_{j\in S} q_{ij} = 0$ for all $i \in S$.

Then, the above conditions imply that

$$\sum_{j\in S} |q_{ij}| = -q_{ii} + \sum_{j\in\mathbb{N}\setminus\{i\}} q_{ij} = -2q_{ii} = 2|q_{ii}| < \infty. \tag{5.4}$$

In particular, $q^\lambda\colon \ell^\infty \to \ell^\infty$ is a bounded linear operator if and only if $\|q^\lambda\| := 2\sup_{i\in S} |q_{ii}^\lambda| < \infty$. If

$$\sup_{\lambda\in\Lambda}\sup_{i\in S} |q_{ii}^\lambda| = \frac{1}{2}\sup_{\lambda\in\Lambda} \|q^\lambda\| < \infty, \tag{5.5}$$

then, by [25, Example 6.7], it follows that

$$\left\|\frac{\mathscr{S}(h)u-u}{h} - \mathcal{Q}u\right\|_\infty \to 0 \quad \text{as } h \searrow 0$$

with $\mathcal{Q}u := \sup_{\lambda\in\Lambda} q^\lambda u$ for all $u \in \ell^\infty$. Moreover, the function $v\colon [0,\infty) \to \ell^\infty$, $t \mapsto \mathscr{S}(t)u$ is the unique classical solution to the differential equation

$$v'(t) = \mathcal{Q}v(t), \quad \text{for all } t \geq 0, \quad v(0) = u.$$

Therefore, $\mathscr{S}(t)u$ can be computed by solving the above differential equation. In the finite-state case, this approach has been used by Škulj [35] and by Nendel [24] for Markov chains under convex expectations in order to define and compute the transition operator as a solution to an ordinary differential equation. Krak et al. [23] use the explicit Euler method in order to discretize the time interval, and come up with the transition operator as the limit of the Euler approximation. Both approaches result in the same transition operator as shown in [23] and [24]. Examples for a family $(q^\lambda)_{\lambda\in\Lambda}$ that satisfies (5.5) are given by $q^\lambda = \lambda q$ for $\lambda \in \Lambda$, $\Lambda \subset [0,\infty)$ compact, and a fixed bounded operator q satisfying the above conditions (i) - (iii). The operator q could, for example, be the generator of a Poisson process (with intensity 1). In this case, the Nisio semigroup $\big(\mathscr{S}(t)\big)_{t\geq 0}$ can be viewed as the family of transition operators for an imprecise Poisson process with imprecision in the intensity of the process.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[2] D. Bartl. Conditional nonlinear expectations. *Stochastic Process. Appl.*, Forthcoming.

[3] D. Bartl, P. Cheridito, M. Kupper, and L. Tangpi. Duality for increasing convex functionals with countably many marginal constraints. *Banach J. Math. Anal.*, 11(1):72–89, 2017.

[4] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and A. Rustichini. Niveloids and their extensions: risk measures on small domains. *J. Math. Anal. Appl.*, 413(1):343–360, 2014.

[5] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math.*, 60(7):1081–1110, 2007.

[6] G. Choquet. Theory of capacities. *Ann. Inst. Fourier, Grenoble*, 5:131–295 (1955), 1953–1954.

[7] G. Choquet. Forme abstraite du théorème de capacitabilité. *Ann. Inst. Fourier. Grenoble*, 9:83–89, 1959.
[8] S. N. Cohen and R. J. Elliott. Solutions of backward stochastic differential equations on Markov chains. *Commun. Stoch. Anal.*, 2(2):251–262, 2008.
[9] S. N. Cohen and R. J. Elliott. Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. *Ann. Appl. Probab.*, 20(1):267–311, 2010.
[10] F. Coquet, Y. Hu, J. Mémin, and S. Peng. Filtration-consistent nonlinear expectations and related g-expectations. *Probab. Theory Related Fields*, 123(1):1–27, 2002.
[11] J. De Bock. The limit behaviour of imprecise continuous-time Markov chains. *J. Nonlinear Sci.*, 27(1):159–196, 2017.
[12] G. de Cooman, F. Hermans, and E. Quaeghebeur. Imprecise Markov chains and their limit behavior. *Probab. Engrg. Inform. Sci.*, 23(4):597–635, 2009.
[13] F. Delbaen. *Coherent risk measures*. Cattedra Galileiana. [Galileo Chair]. Scuola Normale Superiore, Classe di Scienze, Pisa, 2000.
[14] F. Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, Berlin, 2002.
[15] C. Dellacherie and P.-A. Meyer. *Probabilities and potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam-New York; North-Holland Publishing Co., Amsterdam-New York, 1978.
[16] R. Denk, M. Kupper, and M. Nendel. Kolmogorov-type and general extension results for nonlinear expectations. *Banach J. Math. Anal.*, 12(3):515–540, 2018.
[17] R. Denk, M. Kupper, and M. Nendel. A semigroup approach to nonlinear Lévy processes. *Stochastic Process. Appl.*, Forthcoming.
[18] N. Dunford and J. T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.
[19] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance Stoch.*, 6(4):429–447, 2002.
[20] H. Föllmer and A. Schied. *Stochastic finance*. Walter de Gruyter & Co., Berlin, extended edition, 2011. An introduction in discrete time.
[21] M. Frittelli and E. R. Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486, 2002.
[22] D. J. Hartfiel. *Markov set-chains*, volume 1695 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998.
[23] T. Krak, J. De Bock, and A. Siebes. Imprecise continuous-time Markov chains. *Internat. J. Approx. Reason.*, 88:452–528, 2017.
[24] M. Nendel. Markov chains under nonlinear expectation. *Preprint*, 2018.
[25] M. Nendel and M. Röckner. Upper envelopes of families of Feller semigroups and viscosity solutions to a class of nonlinear Cauchy problems. *Preprint arXiv:1906.04430*, 2019.
[26] M. Nisio. On a non-linear semi-group attached to stochastic optimal control. *Publ. Res. Inst. Math. Sci.*, 12(2):513–537, 1976/77.
[27] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lect. Notes Control Inf. Sci.*, pages 200–217. Springer, Berlin, 1992.
[28] R. Pelessoni and P. Vicig. Convex imprecise previsions. *Reliab. Comput.*, 9(6):465–485, 2003. Special issue on dependable reasoning about uncertainty.
[29] S. Peng. Nonlinear expectations and nonlinear Markov chains. *Chinese Ann. Math. Ser. B*, 26(2):159–184, 2005.
[30] S. Peng. G-expectation, G-Brownian motion and related stochastic calculus of Itô type. In *Stochastic analysis and applications*, volume 2 of *Abel Symp.*, pages 541–567. Springer, Berlin, 2007.

[31] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Process. Appl.*, 118(12):2223–2253, 2008.

[32] H. M. Soner, N. Touzi, and J. Zhang. Martingale representation theorem for the G-expectation. *Stochastic Process. Appl.*, 121(2):265–287, 2011.

[33] H. M. Soner, N. Touzi, and J. Zhang. Quasi-sure stochastic analysis through aggregation. *Electron. J. Probab.*, 16:no. 67, 1844–1879, 2011.

[34] D. Škulj. Discrete time Markov chains with interval probabilities. *Internat. J. Approx. Reason.*, 50(8):1314–1329, 2009.

[35] D. Škulj. Efficient computation of the bounds of continuous time imprecise Markov chains. *Appl. Math. Comput.*, 250:165–180, 2015.

[36] B. Z. Vulikh. *Introduction to the theory of partially ordered spaces.* Translated from the Russian by Leo F. Boron, with the editorial collaboration of Adriaan C. Zaanen and Kiyoshi Iséki. Wolters-Noordhoff Scientific Publications, Ltd., Groningen, 1967.

[37] P. Walley. *Statistical reasoning with imprecise probabilities*, volume 42 of *Monographs on Statistics and Applied Probability.* Chapman and Hall, Ltd., London, 1991.

[1]Center for Mathematical Economics, Bielefeld University, 33615 Bielefeld, Germany

E-mail address: Max.Nendel@uni-bielefeld.de