CONDITIONAL NONLINEAR EXPECTATIONS

DANIEL BARTL

ABSTRACT. Given two Polish spaces $U$ and $V$, denote by $\mathcal{L}(U \times V)$ and $\mathcal{L}(U)$ the set of all bounded upper semianalytic functions from $U \times V$ and $U$ to the real line, respectively. Let $\mathcal{E}(:|U) : \mathcal{L}(U \times V) \to \mathcal{L}(U)$ be a sublinear increasing functional which leaves $\mathcal{L}(U)$ invariant. We prove that there exists a set-valued mapping $P_V$ from $U$ to the set of probabilities on $V$ with compact convex values and analytic graph such that $\mathcal{E}(X|U)(u) = \sup_{P \in P_V(u)} \int_V X(u,v) P(du)$ if and only if $\mathcal{E}(:|U)$ is pointwise continuous from below and continuous from above on the continuous functions. Further, given another sublinear increasing functional $\mathcal{E}(\cdot) : \mathcal{L}(U \times V) \to \mathbb{R}$ which leaves the constants invariant, the tower property $\mathcal{E}(\cdot) = \mathcal{E}(\mathcal{E}(\cdot|U))$ is characterized via a pasting property of the representing sets of probabilities. As applications, we characterize under which conditions the product of a set of probabilities and a set of kernels is compact, and under which conditions a nonlinear version of Fubini’s theorem holds true.

1. INTRODUCTION

The Daniell-Stone theorem is a basic but essential result in measure- and integration theory and states that an abstract expectation (i.e. a linear increasing functional preserving the constants) has an integral representation w.r.t. a probability measure if and only if it satisfies the monotone convergence property. Its nonlinear version is significantly more involved, builds on capacity theory, and is mainly due to Choquet [12]. Let $\Omega$ be a Polish space, denote by $\mathcal{P}(\Omega)$ the set of Borel-probabilities on $\Omega$, and write $C_b(\Omega)$ and $L(\Omega)$ for the set of bounded functions from $\Omega$ to $\mathbb{R}$ which are continuous and upper semianalytic (i.e. all upper level sets are analytic, see Appendix A for a short summary), respectively.

Theorem (Choquet). Let $\mathcal{E}(\cdot) : \mathcal{L}(\Omega) \to \mathbb{R}$ be a sublinear expectation (i.e. a sublinear increasing functional satisfying $\mathcal{E}(X) = X$ for all constant functions $X \in \mathbb{R}$). Then there exists a convex and weakly compact set $\mathcal{P} \subset \mathcal{P}(\Omega)$ such that

$$\mathcal{E}(X) = \sup_{P \in \mathcal{P}} E_P[X] \quad \text{for all } X \in \mathcal{L}(\Omega)$$

if and only if $\mathcal{E}(X_n) \downarrow \mathcal{E}(X)$ for every sequence $X_n \in C_b(\Omega)$ with $X_n \downarrow X \in \mathcal{L}(\Omega)$ pointwise, and $\mathcal{E}(X_n) \uparrow \mathcal{E}(X)$ for every sequence $X_n \in \mathcal{L}(\Omega)$ with $X_n \uparrow X \in \mathcal{L}(\Omega)$ pointwise.
See Choquet’s original work [12] for the theorem in a different form, and e.g. [3] [22] as well as [4] Section 2 for applications and the statement in precisely this form. For convenience, a sketch of the proof is given in Appendix B.

The first goal of this article is to obtain a result of this type for conditional sublinear expectations. Here however, one needs to settle first on the right formulation of “sup$_{P \in \mathcal{P}} E_P[\cdot | U]”$ because $E_P[\cdot | U]$ is defined only up to $P$-zero sets which are usually not the same over the class $\mathcal{P}$, and the uncountable supremum over $P \in \mathcal{P}$ may fail to be measurable. We will see that both problems can be solved if $\Omega$ is (homeomorphic to) a product space $\Omega = U \times V$ of two Polish spaces $U$ and $V$. Indeed, in this case every probability $P \in \Psi(\Omega)$ can be written as $P = Q \otimes R$ for a probability $Q \in \Psi(U)$ and a kernel $R: U \to \Psi(V)$, and it holds

$$E_P[X|U](u) = \int_V X(u,v) R(u)(dv) =: E_{R(u)}[X(u,\cdot)]$$

for every $X \in \mathcal{L}(\Omega)$ and $Q$-almost every $u \in U$. In particular, if for every $u \in U$ one is given a set $\mathcal{P}_V(u)$ of probabilities on $V$ instead of only one $R(u)$, then

$$(2) \quad \mathcal{E}(\cdot|U): \mathcal{L}(\Omega) \to \mathbb{R}^U, \quad \mathcal{E}(X|U)(u) := \sup_{P \in \mathcal{P}_V(u)} E_P[X(u,\cdot)]$$

defines a sublinear increasing functional satisfying $\mathcal{E}(X|U) = X$ for all $X \in \mathcal{L}(U)$. Further, under the assumption that graph $\mathcal{P}_V := \{(u,P) \in U \times \Psi(V): P \in \mathcal{P}_V(u)\}$ is an analytic set, the powerful theory of Luzin and Suslin applies and guarantees that $\mathcal{E}(\cdot|U)$ is in fact a mapping from $\mathcal{L}(\Omega)$ to $\mathcal{L}(U)$. This pointwise construction of conditional sublinear expectations (i.e. sublinear increasing functionals $\mathcal{E}(\cdot|U): \mathcal{L}(\Omega) \to \mathcal{L}(U)$ satisfying $\mathcal{E}(X|U) = X$ for every $X \in \mathcal{L}(U)$) was first used in Nutz and van Handel [27] and Bouchard and Nutz [8] in a continuous and discrete time framework, respectively. The reverse question, whether an arbitrary conditional sublinear expectation has an associated family of probabilities $(\mathcal{P}_V(u))_u$ such that (2) holds, and if graph $\mathcal{P}_V$ needs to be analytic, was not studied so far.

**Theorem 1.1.** Let $\mathcal{E}(\cdot|U): \mathcal{L}(\Omega) \to \mathcal{L}(U)$ be a conditional sublinear expectation. Then there exists a set-valued mapping $\mathcal{P}_V: U \to \Psi(V)$ with analytic graph and nonempty, convex, and weakly compact values such that

$$(3) \quad \mathcal{E}(X|U)(u) = \sup_{P \in \mathcal{P}_V(u)} E_P[X(u,\cdot)] \quad \text{for every } u \in U \text{ and } X \in \mathcal{L}(\Omega)$$

if and only if $\mathcal{E}(X_n|U) \downarrow \mathcal{E}(X|U)$ pointwise for every sequence $X_n \in C_b(\Omega)$ with $X_n \downarrow X \in \mathcal{L}(\Omega)$ pointwise, and $\mathcal{E}(X_n|U) \uparrow \mathcal{E}(X|U)$ pointwise for every sequence $X_n \in \mathcal{L}(\Omega)$ with $X_n \uparrow X \in \mathcal{L}(\Omega)$ pointwise.

Note that this is a generalization of Choquet’s theorem: If $U$ is a singleton, then $\Omega$ is naturally homeomorphic to $U \times \Omega$ and $\mathcal{L}(U)$ to $\mathbb{R}$, in particular the continuity conditions coincide.

In case of linear conditional expectations, next to monotonicity and preservation of $U$-measurable functions, the most important property is the tower property. While the first two properties are part of the definition of conditional sublinear expectations, the tower property $\mathcal{E}(\cdot) = \mathcal{E}(\mathcal{E}(\cdot|U))$ (also called dynamic programming principle) does not hold in general. However, it is possible to characterize when it does hold true on the level of representing probabilities.
**Theorem 1.2.** Assume that $\mathcal{E}(\cdot)$ and $\mathcal{E}(\cdot|U)$ satisfies the assumptions of the previous theorems and therefore have the representation (1) and (3), respectively. If further $\mathcal{E}(X|U)$ is Borel for every $X \in C_b(\Omega)$, then $\mathcal{E}(\cdot) = \mathcal{E}\left(\mathcal{E}(\cdot|U)\right)$ if and only if

$$P = P_U \otimes P_V := \{Q \otimes R : Q \in P_U \text{ and } R(\cdot) \in P_V(\cdot) \text{ Q-almost surely}\},$$

where $P_U := \{Q(\cdot \times V) : Q \in P\} \subset \mathfrak{P}(U)$ is the restriction of $P$ to $U$, and $R : U \to \mathfrak{P}(V)$ is a kernel.

Note that it follows rather directly from results on analytic sets that $\mathcal{E}(\cdot) = \mathcal{E}(\mathcal{E}(\cdot|U))$ whenever $P = P_U \otimes P_V$. The actual statement of the theorem is the reverse implication, which has the following application.

**Proposition 1.3.** For any set-valued mapping $P_V : U \rightharpoonup \mathfrak{P}(V)$ with convex values,

$$P_U \otimes P_V \subset \mathfrak{P}(\Omega) \text{ is compact for every compact convex set } P_U \subset \mathfrak{P}(U)$$

if and only if $P_V$ has compact values and $u \mapsto \max_{P \in P_V(u)} E_P[|X|]$ is upper semi-continuous for every $X \in C_b(V)$.

Further we characterize in which cases a nonlinear version of Fubini’s theorem (on the interchanging of the order of integration) holds true in Proposition 3.1 and provide examples by dynamic risk measures in the presence of Knightian uncertainty in discrete time as well as the controlled Brownian motion in Example 3.2 and Example 3.3, respectively.

Finally, let us point out that the choice of upper semianalytic functions as opposed to canonical ones as Borel- or universally measurable functions is important (in fact, fundamental): As Theorem 1.1 should be an extension of Choquet’s theorem, it should at least include the case where $P_V(u)$ in (3) does not depend on $u \in U$. However, already then one can not expect $\mathcal{E}(X|U)$ to be Borel for $X$ Borel. As a matter of fact it is not possible (in general) to work with vector spaces containing all bounded Borel functions. This is discussed in detail in Remark 2.12.

The systematic treatment of nonlinear expectations with domain $C_b(\Omega)$ (and its completion with respect to a seminorm induced by a capacity) started in [16, 30], in particular with connections to the nonlinear Brownian motion introduced by Peng [29]. The extension of the nonlinear Brownian motion to measurable (upper semianalytic) functions was carried out in [27] and generalized to nonlinear Levy-processes in [26]. In the framework of robust mathematical finance (in discrete time), conditional nonlinear expectations where introduced in the seminal work of Bouchard and Nutz [8] and successfully applied several times, see Section 3.3 for references. Further applications can be found e.g. in the context of (non-exponential) large deviations [18, 23] where the tower property (therein referred to as tensorization) plays a crucial role, or the context of fully nonlinear PDE’s where the tower property is the flow/semigroup property, see [17] and references therein. Let us also mention that if $\mathcal{L}(\Omega)$ is replaced by the quotient space $L^\infty(\Omega, P^*)$ with respect to some reference measure $P^*$, nonlinear expectations and their dual representation where already studied in detail due to the relation to risk measures, see e.g. [14, 20]. In a similar manner, [13] works in a setting where essential suprema are assumed to exist; see [24] for the characterization of this (severe) assumption.

The rest of this article is organized as follows: All results and proofs for conditional nonlinear expectations (in a more general convex instead of sublinear form) are presented in Section 2. Applications, the proof of Proposition 1.3 and examples
are given in Section 4. A summary together with basic facts about analytic sets is
given in Appendix A and Appendix B contains a sketch of Choquet’s theorem in
form of nonlinear expectations.

2. Main results

We first fix notation. For a Polish space $\Omega$, denote by $\mathcal{L}(\Omega)$ the set of all upper
semianalytic bounded functions from $\Omega$ to $\mathbb{R}$, by $usc_b(\Omega)$ and $C_b(\Omega)$ the subsets of
upper semicontinuous and continuous functions, respectively. A short summary of
analytic (and universally measurable) sets and functions is given in Appendix A. For
a function $X : \Omega \to \mathbb{R}$, write $\|X\|_{\infty} := \sup_{\omega \in \Omega} |X(\omega)|$ for the maximum norm. The
set of $\sigma$-additive probability measures on the Borel $\sigma$-field $\mathcal{B}(\Omega)$ of $\Omega$ is denoted
by $\mathfrak{P}(\Omega)$ and endowed with the weak topology $\sigma(\mathfrak{P}(\Omega), C_b(\Omega))$, i.e. the coarsest
topology making the mappings $P \mapsto E_P[X]$ continuous for every $X \in C_b(\Omega)$. This
renders $\mathfrak{P}(\Omega)$ a Polish space. Given a second Polish space $\Omega'$, the term kernel refers
to universally measurable mappings $R : \Omega \to \mathfrak{P}(\Omega')$. Given such a kernel $R$ and
a measure $Q \in \mathfrak{P}(\Omega)$, the formula $Q \otimes R(\cdot) := \int_{\Omega} \int_{\Omega'} 1_A(\omega, \omega') R(\omega)(d\omega') Q(d\omega)$
defines a probability measure $Q \otimes R$ on $\Omega \times \Omega'$. Conversely, by the disintegration
theorem, every probability $P \in \mathfrak{P}(\Omega \times \Omega')$ can be written in this form (even with $R$
Borel) and we simply write $P = Q \otimes R \in \mathfrak{P}(\Omega \times \Omega')$. Product spaces are endowed with
the product topology and when functions are in consideration, (in-)equalities and
convergence is to be understood in a pointwise sense, unless stated otherwise.

For the remainder of this section, fix a Polish space $\Omega$ which is assumed to be
homeomorphic to the product of two Polish spaces $U$ and $V$; written as $\Omega =
U \times V$. For notational convenience, $\mathcal{L}(U)$ and $\mathcal{L}(V)$ are identified with the sets of
functions in $\mathcal{L}(\Omega)$ which depend only on the first component and second component,
respectively, and viewed as a subspace of $\mathcal{L}(\Omega)$.

**Definition 2.1.** A mapping $E(\cdot|U) : \mathcal{L}(\Omega) \to \mathcal{L}(U)$ is called conditional nonlinear
expectation, if for all $X, Y \in \mathcal{L}(\Omega)$ it holds

- $E(X|U) \leq E(Y|U)$ whenever $X \leq Y$,
- $E(X|U) = X$ whenever $X \in \mathcal{L}(U)$,
- $E(\lambda X + (1 - \lambda)Y|U) \leq \lambda E(X|U) + (1 - \lambda)E(Y|U)$ for all $\lambda \in [0, 1]$.

Further $E(\cdot|U)$ is said to be a conditional sublinear expectation if in addition

- $E(\lambda X|U) = \lambda E(X|U)$ for all $\lambda \in [0, +\infty)$.

If $U$ is a singleton, then $\Omega$ is homeomorphic to $U \times \Omega$ and $\mathcal{L}(U)$ to $\mathbb{R}$. In this case
we write $E(\cdot) = E(\cdot|U)$ and drop the phrase conditional.

2.1. Continuity and representation. The goal of this section is to establish a
conditional version of Choquet’s theorem for nonlinear expectations. The sublinear
case, stated in Theorem 1.1, will be a special case; its proof is given at the end of
this section.

**Definition 2.2.** A conditional nonlinear expectation $E(\cdot|U)$ is said to be continuous
from above (on $C_b(\Omega)$) if

(A) $E(X_n|U) \downarrow E(X|U)$ for all sequences $X_n \in C_b(\Omega)$ with $X_n \downarrow X \in \mathcal{L}(\Omega)$.

Similar, $E(\cdot|U)$ is said to be continuous from below (on $\mathcal{L}(\Omega)$) if

(B) $E(X_n|U) \uparrow E(X|U)$ for all sequences $X_n \in \mathcal{L}(\Omega)$ with $X_n \uparrow X \in \mathcal{L}(\Omega)$.
Remark 2.3. If $\mathcal{E}(\cdot|U)$ is a conditional nonlinear expectation which satisfies (A) and (B), then $\mathcal{E}(X_n|U) \downarrow \mathcal{E}(X|U)$ for every sequence $X_n \in \text{uscs}_b(\Omega)$ with $X_n \downarrow X \in \mathcal{L}(\Omega)$. This is shown in the proof of Theorem 2.6.

Remark 2.4. A conditional nonlinear expectation $\mathcal{E}(\cdot|U)$ satisfies (A) if and only if it satisfies both
\begin{align*}
(A') & \mathcal{E}(X_n|U) \downarrow 0 \text{ for all sequences } X_n \in C_b(\Omega) \text{ with } X_n \downarrow 0, \\
(A'') & \sup_{X \in C_b(\Omega)} (E_P[X] - \mathcal{E}(X|U)) = \sup_{X \in \text{uscs}_b(\Omega)} (E_P[X] - \mathcal{E}(X|U)) \text{ for every probability } P \in \mathcal{P}(\Omega).
\end{align*}

Proof. Since $\Omega$ is a Polish space, every upper semicontinuous function can be written as the decreasing limit of a sequence of continuous functions. Therefore it is clear that (A) implies $(A')$ and $(A'')$. The other direction will be shown within the proof of Theorem 2.6. □

Remark 2.5. In the non-conditional case (i.e., when $U$ is a singleton) condition $(A')$ is equivalent to the well-known “tightness” condition: There exists a sequence of compact sets $K_n \subset \Omega$ such that $\mathcal{E}(m1_{K_n}) \downarrow 0$ for every positive real number $m$. In the conditional case this no longer holds true.

Proof. We only show that in general there is no sequence of compact sets $K_n \subset \Omega$ such that $\mathcal{E}(1_{K_n}|U) \downarrow 0$. The equivalence of $(A')$ and the tightness condition in the non-conditional case is shown within the proof of Theorem 2.6. Let $U := V := \mathbb{R}^N$ and define
\[
\mathcal{E}(X|U)(u) := X(u, u) \quad \text{for } u \in U \text{ and } X \in \mathcal{L}(\Omega).
\]

Then $\mathcal{E}(\cdot|U)$ is a conditional sublinear expectation which clearly satisfies (A) and (B). Assume that there exists a sequence of compact sets $K_n \subset \Omega$ such that $\mathcal{E}(1_{K_n}|U) \downarrow 0$ pointwise, where one may assume without loss of generality that $K_n = C_n \times C_n$ for $C_n \subset U = V$ compact. Since $\mathcal{E}(1_{K_n}|U)(u) = 1_{C_n}(u)$ it follows that $U = \{C_n : n\}$ and thus, by the Baire category theorem, there exists some $n$ such that $C_n$ has nonempty interior. However, every compact subset of $U = \mathbb{R}^N$ clearly has empty interior, and thus such a sequence $K_n$ cannot exist. □

Theorem 2.6. Let $\mathcal{E}(\cdot|U)$ be a conditional nonlinear expectation which satisfies (A) and (B). Then there exists a lower semianalytic function $\alpha_V : U \times \mathcal{P}(V) \to [0, +\infty]$ such that $\alpha_V(u, \cdot)$ is convex, $\{\alpha_V(u, \cdot) \leq c\}$ is compact for every $c \in \mathbb{R}$, and $\inf_P \alpha_V(u, \cdot) = 0$ for every $u \in U$ and $c \in \mathbb{R}$, which satisfies
\begin{equation}
\mathcal{E}(X|U)(u) = \sup_{P \in \mathcal{P}(V)} (E_P[X(u, \cdot)] - \alpha_V(u, P))
\end{equation}

for every $u \in U$ and $X \in \mathcal{L}(\Omega)$.

Conversely, if $\alpha_V : U \times \mathcal{P}(V) \to [0, +\infty]$ is a lower semianalytic function such that $\inf_P \alpha_V(u, P) = 0$ for every $u \in U$, then $\mathcal{E}(\cdot|U)$ defined by (1) is a conditional nonlinear expectation satisfying (B). If in addition $\{\alpha_V(u, \cdot) \leq c\}$ is compact and $\alpha_V(u, \cdot)$ is convex for every $u \in U$ and $c \in \mathbb{R}$, then $\mathcal{E}(\cdot|U)$ also satisfies (A).

Remark 2.7. For a conditional nonlinear expectation $\mathcal{E}(\cdot|U)$ satisfying (A) and (B) there are in general many functions $\alpha_V : U \times \mathcal{P}(V) \to [0, +\infty]$ such that (4) holds and not every $\alpha_V$ needs to be lower semianalytic. However, if $\alpha_V(u, \cdot)$ is
required to be convex and lower semicontinuous (which particularly is satisfied if all sublevel sets are compact) for every \( u \in U \), then \( \alpha_V \) is unique, lower semianalytic, and in fact given by

\[
(5) \quad \alpha_V(u,P) = \sup_{X \in C_b(\Omega)} (E_P[X(u,\cdot)] - \mathcal{E}(X|U)(u))
\]

for \( u \in U \) and \( P \in \mathfrak{P}(V) \).

**Proof.** To show that \( \alpha_V \) needs not to be unique (and lower semianalytic), let \( U := V := [0,1] \) and \( \alpha_V(u,P) := 0 \) for \( u \in U \) and \( P \in \mathfrak{P}(V) \). By Theorem 2.10

\[
\mathcal{E}(-U) : \mathcal{L}(\Omega) \to \mathcal{L}(U), \quad \mathcal{E}(X|U)(u) := \sup_{P \in \mathfrak{P}(V)} (E_P[X(u,\cdot)] - \alpha_V(u,P))
\]

defines a conditional nonlinear expectation which satisfies (A) and (B). Now define

\[
\tilde{\alpha}_V(u,P) := +\infty 1_{\mathfrak{D}(V) \neq \emptyset} 1_{R(u) \neq \emptyset}(P),
\]

where \( \mathfrak{D}(V) := \{ \delta_v : v \in V \} \) is the set of all Dirak measures and \( R : U \to \mathfrak{P}(V) \) is any non-universally measurable function such that \( R(u) \notin \mathfrak{D}(V) \) for every \( u \in U \). Then evidently \( \mathcal{E}(X|U)(u) = \sup_{P \in \mathfrak{P}(V)} (E_P[X(u,\cdot)] - \tilde{\alpha}_V(u,P)) \) but \( \tilde{\alpha}_V \) is not lower semianalytic as \( \{ \alpha_V \leq 0 \} \) is the disjoint union of \( U \times \mathfrak{D}(V) \) and \( \{(u,R(u)) : u \in U \} \).

As for the second part, notice that (5) follows from the Fenchel-Moreau theorem. It is shown within the proof of Theorem 2.6 that under conditions (A) and (B), the theorem (more precisely, the convex version as stated in [4, Section 2]). Thus

\[
\mathcal{E}(\cdot|U) \text{ is lower semianalytic.}
\]

**Proof of Theorem 2.6.** Let \( \mathcal{E}(-U) \) be a conditional nonlinear expectation which satisfies (A) and (B), and fix some \( u \in U \). Then, the functional \( \mathcal{E}(\cdot|U)(u) \) from \( \mathcal{L}(\Omega) \) to \( \mathbb{R} \) is a nonlinear expectation which satisfies all assumptions of Choquet’s theorem (more precisely, the convex version as stated in [4, Section 2]). Thus

\[
\mathcal{E}(X|U)(u) = \sup_{P \in \mathfrak{P}(V)} (E_P[X(u,\cdot)] - \mathcal{E}(X|U)(u)) \quad \text{for all } X \in \mathcal{L}(\Omega),
\]

where \( \mathcal{E}(X|U)(u) = \sup_{X \in C_b(\Omega)} (E_P[X] - \mathcal{E}(X|U)(u)) \). Let \( \tilde{P} = Q \otimes R \in \mathfrak{P}(V) \) such that \( \mathcal{E}(\cdot|U)(u) < +\infty \). By assumption it holds \( \mathcal{E}(X|U)(u) = X(u) \) for every \( X \in C_b(U) \), therefore \( E_P[X] = X(u) \), otherwise a scaling argument yields \( \mathcal{E}(X|U)(u) = +\infty \). Since \( E_Q[X] = E_P[X] = X(u) \) for all \( X \in C_b(U) \), it follows that \( Q = \delta_u \) and thus \( P := R(u) \in \mathfrak{P}(V) \) satisfies \( P = \delta_u \otimes P \). Thus

\[
\mathcal{E}(X|U)(u) = \sup_{P \in \mathfrak{P}(V)} (E_P[X(u,\cdot)] - \alpha_V(u,P)) \quad \text{for all } X \in \mathcal{L}(\Omega),
\]

where the function \( \alpha_V : U \times \mathfrak{P}(V) \to [-\infty, +\infty] \) defined by

\[
\alpha_V(u,P) := \mathcal{E}(\cdot(u,\delta_u) \otimes P) = \sup_{X \in C_b(\Omega)} (E_P[X(u,\cdot)] - \mathcal{E}(X|U)(u)).
\]

Further, since \( \mathcal{E}(0|U)(u) = 0 \) one clearly has \( \inf_P \alpha_V(u,P) = 0 \) and since \( \{ \mathcal{E}(\cdot(u,\cdot) \leq c \} \) is compact [4, Theorem 2.2] it also follows that \( \{ \alpha_V(u,\cdot) \leq c \} \) is compact.

We are left to show that \( \alpha_V \) is lower semianalytic and convex in \( P \). Since \( \Omega \) is a Polish space, there exists a metric \( d' \) on \( \Omega \) which induces the original topology under which the space \( uc_b(\Omega, d') \) becomes separable [31, Lemma 3.1.4]. Here \( uc_b(\Omega, d') \) denotes the set of all bounded functions from \( \Omega \) to \( \mathbb{R} \) which are uniformly continuous with respect to \( d' \), and this space is endowed with the maximum norm. Let \( D \) be a countable dense subset. Now, since \( -\|X\|_\infty \leq \mathcal{E}(X|U) \leq \|X\|_\infty \) for every
Let $X \in \mathcal{L}(\Omega)$ by monotonicity of $\mathcal{E}(\cdot|U)$, it follows for all $c \in \mathbb{R}$, $u \in U$, and $X \in \mathcal{L}(\Omega)$ that
\begin{equation}
\mathcal{E}(X|U)(u) = \sup_{p \in \Lambda_{2c}(u)} (E_p[X(u, \cdot)] - \alpha_V(u, P)) \text{ if } \|X\|_{\infty} \leq c,
\end{equation}
where $\Lambda_{2c}(u) := \{\alpha_V(u, \cdot) \leq 2c\}$. Fix $u \in U$, $P \in \mathfrak{P}(V)$, $X \in C_b(\Omega)$, $\varepsilon \in (0, 1)$, and let $c := \|X\|_{\infty} + 1$. Since the set $\Lambda_{2c}(u)$ is compact, Prokhorov’s theorem yields the existence of a compact set $C \subset V$ such that
\begin{equation*}
P(C^c) \leq \frac{\varepsilon}{c} \quad \text{and} \quad \sup_{Q \in \Lambda_{2c}(\omega)} Q(C^c) \leq \frac{\varepsilon}{c}.
\end{equation*}

Then $K := \{u\} \times C$ is a compact subset of $\Omega$ and since the metric $d'$ induces the original topology on $\Omega$, the set $K \subset (\Omega, d')$ is still compact and $X : (\Omega, d') \to \mathbb{R}$ still continuous. Thus $X1_K \in uc_b(K, d')$ and by a version of Tietze’s extension theorem [21] Theorem 3 there exists a uniformly continuous function $Y' \in uc_b(\Omega, d')$ such that $Y = X$ on $K$, where one can assume without loss of generality that $\|Y\|_{\infty} \leq \|X\|_{\infty}$. Moreover, as $D \subset uc_b(\Omega, d')$ is dense, there exists $Y' \in D$ such that $\|Y - Y'\|_{\infty} \leq \varepsilon$. Then in particular $\|Y'\|_{\infty} \leq c$, so that for every $Q \in \Lambda_{2c}(u)$, one has
\begin{equation*}
E_Q[Y'(u, \cdot)] \leq E_Q[Y(u, \cdot)1C(u)] + \varepsilon \leq E_Q[X(u, \cdot)1C(u)] + 2\varepsilon = E_Q[X(u, \cdot)] + 3\varepsilon,
\end{equation*}
which, in combination with (6), implies $\mathcal{E}(Y'|U)(u) \leq \mathcal{E}(X|U)(u) + 3\varepsilon$. Further, changing the roles of $X$ and $Y'$ and replacing $Q$ by $P$, one gets that $E_P[X(u, \cdot)] \leq E_P[Y'(u, \cdot)] + 3\varepsilon$ and therefore
\begin{equation*}
E_P[X(u, \cdot)] - \mathcal{E}(X|U)(u) \leq E_P[Y'(u, \cdot)] - \mathcal{E}(Y'|U)(u) + 6\varepsilon \leq \sup_{Z \in D} (E_P[Z(u, \cdot)] - \mathcal{E}(Z|U)(u)) + 6\varepsilon.
\end{equation*}

Since $D$ is a subset of $C_b(\Omega)$ and $X \in C_b(\Omega)$, $\varepsilon \in (0, 1)$ were arbitrary, this yields
\begin{equation*}
\alpha_V(u, P) = \sup_{Z \in D} (E_P[Z(u, \cdot)] - \mathcal{E}(Z|U)(u)).
\end{equation*}

Finally, for every $Z \in D$, the function $(u, P) \mapsto E_P[Z(u, \cdot)] - \mathcal{E}(Z|U)(u)$ is lower semianalytic [6 Proposition 7.29] and convex in $P$. As the countable supremum, $\alpha_V$ inherits both properties.

To prove the second statement, let $\alpha_V : U \times \mathfrak{P}(V) \to [0, +\infty]$ be a given lower semianalytic function which satisfies $\inf_P \alpha_V(u, P) = 0$ for every $u \in U$, and define
\begin{equation*}
\mathcal{E}(X|U)(u) := \sup_{P \in \mathfrak{P}(V)} (E_P[X(u, \cdot)] - \alpha_V(u, P))
\end{equation*}
for $u \in U$ and $X \in \mathcal{L}(\Omega)$. For $X \in \mathcal{L}(\Omega)$, the mapping

\begin{equation*}
U \times \mathfrak{P}(V) \to [-\infty, +\infty), \quad (u, P) \mapsto E_P[X(u, \cdot)] - \alpha_V(u, P)
\end{equation*}

is upper semianalytic [6 Proposition 7.48], and it follows from [6 Proposition 7.47] that $u \mapsto \mathcal{E}(X|U)(u)$ is upper semianalytic. Further, since $\inf_P \alpha_V(u, P) = 0$, one has
\begin{equation*}
-\|X\|_{\infty} \leq \mathcal{E}(X|U)(u) \leq \|X\|_{\infty}.
\end{equation*}

Thus $\mathcal{E}(X|U)$ is bounded and therefore $\mathcal{E}(X|U) \in \mathcal{L}(U)$. The other properties needed for $\mathcal{E}(\cdot|U)$ to be a conditional nonlinear expectation are immediate. Further, condition (B) follows by interchanging two suprema and the monotone convergence theorem (applied to each $E_P[\cdot]$).
Assume in addition that $\alpha_V$ is convex in $\mathbb{P}$ and $\Lambda_c(u) := \{\alpha_V(u, \cdot) \leq c\}$ is compact for every $c \in \mathbb{R}$ and $u \in U$. Fix some $u \in U$ and a sequence $X_n \in \text{usc}_b(\Omega)$ which decreases pointwise to some $X_\infty \in \mathcal{L}(\Omega)$. Then it follows as in (\ref{7}) that
\[
E(X_n[U])(u) = \max_{P \in \Lambda_c(u)} (E_P[X_n(u, \cdot)] - \alpha_V(u, P)) \quad \text{for } n \in \mathbb{N} \cup \{\infty\},
\]
where $c := \max\{|X_1|, |X_\infty|\}$. Since $\Lambda_2m(u)$ is compact and convex, $X_n$ is a decreasing sequence, and
\[
\mathfrak{P}(V) \ni P \mapsto E_P[X_n(u, \cdot)] - \alpha_V(u, P)
\]
is convex and upper semicontinuous for every $n$ (approximate $X_n$ from above by continuous functions), it follows from (\ref{7}), a minimax theorem \cite[Theorem 2]{19}, and the monotone convergence theorem that
\[
\inf_{n \in \mathbb{N}} E(X_n[U])(u) = \max_{P \in \Lambda_2m(u)} \inf_{n \in \mathbb{N}} (E_P[X_n(u, \cdot)] - \alpha_V(u, P)) = E(X_\infty[U])(u).
\]
Thus condition (A), the missing part of Remark \ref{2.4} and Remark \ref{2.3} are proven. □

**Proof of Theorem 1.1.** If $E(\cdot|U)$ satisfies (A) and (B) and is sublinear, then Theorem \ref{2.6} guarantees that $E(\cdot|U)$ has the representation \ref{4} where $\alpha_V$ is defined by \ref{3} (see the proof of the theorem or Remark \ref{2.7}). Now a scaling argument shows that $\alpha_V$ only takes the values $0$ or $+\infty$ and one can define $\mathcal{P}_V(u) := \{\alpha_V(u, \cdot) = 0\}$. Notice that graph $\mathcal{P}_V = \{\alpha_V \leq 0\}$ is an analytic set. Conversely, if $\mathcal{P}_V$ is given, just set $\alpha_V(u, P) := +\infty 1_{\mathcal{P}_V(u)}(P)$. Then, since $\{\alpha_V \leq c\}$ is empty for $c < 0$ and equals graph $\mathcal{P}_V$ otherwise, $\alpha_V$ is lower semianalytic. □

### 2.2. The tower property

Given a nonlinear expectation $E(\cdot)$ it is, in contrast to the linear case, not always possible to find a conditional nonlinear expectation $E(\cdot|U)$ such that the tower property $E(\cdot) = E(E(\cdot|U))$ holds true, as the simple example $U := V := \{0, 1\}$ and
\[
E(\cdot) := \sup_{P \in \mathcal{P}} E_P[\cdot] \quad \text{where } \mathcal{P} := \{\lambda \delta_{(0,0)} + (1 - \lambda) \delta_{(1,1)} : \lambda \in [0, 1]\},
\]
illustrates. The goal of this section to characterize under which conditions of the representing sets of probabilities the tower property holds true.

**Lemma 2.8.** Let $E(\cdot|U) : \mathcal{L}(\Omega) \to \mathcal{L}(U)$ be a conditional nonlinear expectation and $E(\cdot) : \mathcal{L}(U) \to \mathbb{R}$ be a nonlinear expectation such that
\[
E'(X) = \sup_{P \in \mathfrak{P}(U)} (E_P[X] - \alpha_U(P)); \quad E(X[U])(u) = \sup_{P \in \mathfrak{P}(V)} (E_P[X(u, \cdot)] - \alpha_V(u, P)),
\]
for some $\alpha_U : \mathfrak{P}(U) \to [0, +\infty]$ and lower semianalytic $\alpha_V : U \times \mathfrak{P}(V) \to [0, +\infty]$. Then $E(\cdot) := E'(E(\cdot|U))$ defines a nonlinear expectation from $\mathcal{L}(\Omega)$ to $\mathbb{R}$ and
\[
E(X) = \sup_{P = Q \otimes R \in \mathfrak{P}(\Omega)} \left( E_P[X] - (\alpha_U(Q) + E_Q[\alpha_V(\cdot, R(\cdot))] \right)
\]
for $X \in \mathcal{L}(\Omega)$.

**Proof.** It is clear that $E(\cdot) := E'(E(\cdot|U)) : \mathcal{L}(\Omega) \to \mathbb{R}$ defines a nonlinear expectation. For $P = Q \otimes R \in \mathfrak{P}(\Omega)$ and $X \in \mathcal{L}(\Omega)$ one has
\[
E(X) = E'(E(X[U])) \geq E_Q[E_X[U]] - \alpha_U(Q)
\geq E_Q[\delta_{(u)}(E_R[u])(X(u, \cdot]) - \alpha_V(u, R(u))] - \alpha_U(Q)
= E_P[X] - (\alpha_U(Q) + E_Q[\alpha_V(\cdot, R(\cdot))]),
\]
which shows that the left hand side in \( \text{(8)} \) is larger than the right hand side.

To prove the reverse inequality fix some \( X \in \mathcal{L}(\Omega) \) and \( \varepsilon > 0 \), and let \( Q \in \mathfrak{P}(U) \) be such that

\[
\mathcal{E}'(\mathcal{E}(X|U)) \leq E_Q[\mathcal{E}(X|U)] - \alpha_U(Q) + \varepsilon.
\]

By [3] Proposition 7.48 the mapping

\[
U \times \mathfrak{P}(V) \to [-\infty, +\infty), \quad (u, P) \mapsto E_P[X(u, \cdot)] - \alpha_V(u, P)
\]

is upper semianalytic and since \( X \) is bounded from above, by the Jankov-von Neumann theorem [3] Proposition 7.50, there exists a kernel \( R: U \to \mathfrak{P}(V) \) such that

\[
\mathcal{E}(X|U)(u) = \sup_{P \in \mathfrak{P}(V)} (E_P[X(u, \cdot)] - \alpha_V(u, P)) \leq E_{R(u)}[X(u, \cdot)] - \alpha_V(u, R(u)) + \varepsilon
\]

for every \( u \in U \). Together with \( \text{(9)} \) this implies

\[
\mathcal{E}'(\mathcal{E}(X|U)) = E_{Q(du)}(X|U) - \alpha_V(u, R(u)) + \varepsilon - \alpha_U(Q) + \varepsilon
\]

for \( P := Q \otimes R \in \mathfrak{P}(\Omega) \). As \( \varepsilon \) was arbitrary, the claim follows. \( \square \)

**Remark 2.9.** If in the above lemma both \( \mathcal{E}'(\cdot|\cdot) \) and \( \mathcal{E}(\cdot|U) \) satisfy condition \( (A) \), then \( \mathcal{E}'(\mathcal{E}(\cdot|U)) \) does not necessarily satisfy \( (A) \).

**Proof.** Let \( U := V := \mathbb{R} \) and define

\[
\mathcal{E}'(X) := \sup_{u \in [0,1]} X(u) \quad \text{and} \quad \mathcal{E}(X|U)(u) := X\left(u, \frac{1}{u}\right) \quad \text{with} \quad \frac{1}{0} := 0
\]

for \( X \in \mathcal{L}(U) \) and \( X \in \mathcal{L}(\Omega) \), respectively. Since \( u \mapsto (u, 1/u) \) is Borel, it follows from [3] Lemma 7.30 that \( u \mapsto \mathcal{E}(X|U)(u) \) is upper semianalytic. The sequence of functions \( X_n \in C_b \) defined by

\[
X_n(u, v) := (v - n + 1)1_{[n-1,n]}(v) + 1_{(n,\infty)}(v) \quad \text{for} \quad (u, v) \in \Omega
\]

satisfies \( X_n \downarrow 0 \) but

\[
\mathcal{E}'(\mathcal{E}(X_n|U)) = \sup_{u \in [0,1]} X_n\left(u, \frac{1}{u}\right) \geq X_n\left(\frac{1}{n}, n\right) = 1
\]

for all \( n \). Hence \( \mathcal{E}'(\mathcal{E}(\cdot|U)) \) does not satisfy condition \( (A) \), while both \( \mathcal{E}'(\cdot) \) and \( \mathcal{E}(\cdot|U) \) clearly do satisfy \( (A) \). \( \square \)

In Lemma 2.8 is was shown that the composition of nonlinear expectations can be represented by a function which equals the (integrated) sum over \( \alpha_U \) and \( \alpha_V \). However, it was not shown that this function is minimal in the sense that it is lower semicontinuous in \( P \) and thus given by formula \( \text{(5)} \). The following theorem shows that, given additional regularity, this is true.

**Theorem 2.10.** Let \( \mathcal{E}(\cdot): \mathcal{L}(\Omega) \to \mathbb{R} \) and \( \mathcal{E}(\cdot|U): \mathcal{L}(\Omega) \to \mathcal{L}(U) \) be two (conditional) nonlinear expectations which satisfy \( (A) \) and \( (B) \) and therefore

\[
\mathcal{E}(X) = \sup_{P \in \mathfrak{P}(\Omega)} (E_P[X] - \alpha(P)) \quad \text{for} \quad X \in \mathcal{L}(\Omega),
\]

\[
\mathcal{E}(X) = \sup_{P \in \mathfrak{P}(U)} (E_P[X] - \alpha_U(P)) \quad \text{for} \quad X \in \mathcal{L}(U),
\]

\[
\mathcal{E}(X|U)(u) = \sup_{P \in \mathfrak{P}(V)} (E_P[X(u, \cdot)] - \alpha_V(u, P)) \quad \text{for} \quad X \in \mathcal{L}(\Omega),
\]
where \( \alpha : \mathfrak{B}(\Omega) \to [0, +\infty] \), \( \alpha_U : \mathfrak{B}(U) \to [0, +\infty] \), and \( \alpha_V(u, \cdot) : \mathfrak{B}(V) \to [0, +\infty] \) are lower semicontinuous for every \( u \in U \) (by Theorem 2.6). Then
\[
\mathcal{E}(\cdot) \leq \mathcal{E}(\mathcal{E}(|\cdot| U)) \quad \text{if and only if} \quad \alpha(P) \geq \alpha_U(Q) + E_Q[\alpha_V(\cdot, R(\cdot))]
\]
for all \( P = Q \otimes R \in \mathfrak{B}(\Omega) \). Assume further that \( \mathcal{E}(X|U) \) is Borel for every \( X \in C_b(\Omega) \). Then
\[
\mathcal{E}(\cdot) \geq \mathcal{E}(\mathcal{E}(|\cdot| U)) \quad \text{if and only if} \quad \alpha(P) \leq \alpha_U(Q) + E_Q[\alpha_V(\cdot, R(\cdot))]
\]
for all \( P = Q \otimes R \in \mathfrak{B}(\Omega) \).

In particular, this implies that the tower property \( \mathcal{E}(\cdot) = \mathcal{E}(\mathcal{E}(|\cdot| U)) \) is equivalent to the additivity of the function \( \alpha \), i.e. \( \alpha(P) = \alpha_U(Q) + E_Q[\alpha_V(\cdot, R(\cdot))] \) for all probabilities \( P = Q \otimes R \in \mathfrak{B}(\Omega) \).

**Proof of Theorem 2.10** Define
\[
\beta(P) := \alpha_U(Q) + E_Q[\alpha_V(\cdot, R(\cdot))] \quad \text{for} \quad P = Q \otimes R \in \mathfrak{B}(\Omega)
\]
and \( \mathcal{E}'(\cdot) : \mathcal{L}(U) \to \mathbb{R} \) as the restriction of \( \mathcal{E}(\cdot) \) to \( \mathcal{L}(U) \), i.e. \( \mathcal{E}'(X) := \mathcal{E}(X) \) for every \( X \in \mathcal{L}(U) \). It follows from Lemma 2.8 that
\[
(10) \quad \mathcal{E}(\mathcal{E}(X|U)) = \mathcal{E}'(\mathcal{E}(X|U)) = \sup_{P \in \mathfrak{B}(\Omega)} (E_P[X] - \beta(P)) \quad \text{for} \quad X \in \mathcal{L}(\Omega). 
\]
In particular if \( \alpha \leq \beta \), then \( \mathcal{E}(\cdot) \geq \mathcal{E}(\mathcal{E}(|\cdot| U)) \); and if \( \alpha \geq \beta \), then \( \mathcal{E}(\cdot) \leq \mathcal{E}(\mathcal{E}(|\cdot| U)) \).

Now assume that \( \mathcal{E}(\cdot) \geq \mathcal{E}(\mathcal{E}(|\cdot| U)) \) and fix some \( P \in \mathfrak{B}(\Omega) \). By (10) it holds
\[
\mathcal{E}(X) \geq \mathcal{E}(\mathcal{E}(X|U)) \geq E_P[X] - \beta(P) \quad \text{for} \quad X \in \mathcal{L}(\Omega). 
\]
and therefore
\[
\alpha(P) = \sup_{X \in C_b(\Omega)} (E_P[X] - \mathcal{E}(X)) \leq \sup_{X \in C_b(\Omega)} (E_P[X] - (E_P[X] - \beta(P))) = \beta(P),
\]
where the first quality holds due to lower semicontinuity of \( \alpha \), see Remark 2.7. As \( P \) was arbitrary, this shows \( \alpha \leq \beta \).

It remains to prove that if \( \mathcal{E}(\cdot) \leq \mathcal{E}(\mathcal{E}(|\cdot| U)) \) and \( \mathcal{E}(X|U) \) is Borel for every \( X \in C_b(\Omega) \), then \( \alpha \geq \beta \). To that end, one may argue as in the proof of Theorem 2.6 and choose a metric \( d' \) on \( V \) under which the space of bounded and uniformly continuous functions \( uc_b(V, d') \) becomes separable. Let \( D = \{d_n : n\} \) be a countable dense subset and define
\[
D_n := \{0\} \cup \{d_k - q_l : 1 \leq k, l \leq n\} \subset C_b(V)
\]
for every natural number \( n \), where \( \{q_n : n\} \) is an enumeration of the rational numbers. We claim that
\[
(11) \quad \alpha_V(u, P) = \sup_n \alpha_V^n(u, P) \quad \text{for} \quad u \in U \text{ and } P \in \mathfrak{B}(V), \quad \text{where}
\]
\[
\alpha_V^n(u, P) := \max\{E_P[X] : X \in D_n \text{ such that } \mathcal{E}(X|U)(u) \leq 0\}
\]
for every \( n \). Indeed, fix some \( u \in U \) and \( P \in \mathfrak{B}(V) \). Due to lower semicontinuity of \( \alpha_V(u, \cdot) \), one has
\[
\alpha_V(u, P) = \sup_{X \in C_b(\Omega)} (E_P[X(u, \cdot)] - \mathcal{E}(X|U)(u)),
\]
see Remark 2.7. Let $X \in C_b(\Omega)$ and $\varepsilon > 0$ be arbitrary and define $Y := X(u, \cdot) \in C_b(V)$. Then, by the dual representation of $\mathcal{E}(\cdot|U)$, one has $\mathcal{E}(X|U)(u) = \mathcal{E}(Y|U)(u)$. It follows as in the proof of Theorem 2.6 that there exists some $Y' \in D$ such that

$$E_P[Y'] - \mathcal{E}(Y'|U)(u) \geq E_P[Y] - \mathcal{E}(Y|U)(u) - \varepsilon.$$ 

Now let $q$ be rational such that $0 \leq q - \mathcal{E}(Y'|U)(u) \leq \varepsilon$ and define $Y'' := Y' - q$. Then $\mathcal{E}(Y''|U)(u) \leq 0$ and

$$E_P[Y''] \geq E_P[Y'] - \mathcal{E}(Y'|U)(u) - \varepsilon \geq E_P[Y] - \mathcal{E}(Y|U)(u) - 2\varepsilon = E_P[X(u, \cdot)] - \mathcal{E}(X|U)(u) - 2\varepsilon.$$ 

Since $Y'' \in D_n$ for some large $n$, it follows that

$$\sup_n \alpha^n_v(u, P) \geq E_P[Y''] \geq E_P[X(u, \cdot)] - \mathcal{E}(X|U)(u) - 2\varepsilon$$

and, as $X \in C_b(\Omega)$ was arbitrary, $\sup_n \alpha^n_v(u, P) \geq \alpha^v(u, P) - 2\varepsilon$. As $D_n \subset C_b(V)$, the reverse inequality is clear and therefore (11) is established.

For the remainder fix some $P = Q \otimes R \in \mathcal{F}(\Omega)$, where $R$ is taken Borel. Note that $0 \in D_n$ and $D_n \subset D_{n+1}$ imply $0 \leq \alpha^{n+1}_v \leq \alpha^n_v$ for all $n$, hence the monotone convergence theorem applies and yields

$$\beta(P) = \alpha^v(Q) + E_Q[\alpha^v(\cdot, R(\cdot))].$$

(12)

Fix some $Y \in C_b(U)$ and $n$, and define the set-valued mapping $\Psi: U \to D_n$ by

$$\Psi(u) := \{X \in D_n : E_{R(u)}[X] = \alpha^n_v(u, R(u)) \text{ and } \mathcal{E}(X|U)(u) \leq 0\}.$$ 

The set $D_n$ is endowed with the discrete topology which, as a finite set, makes $D_n$ a Polish space. Then $\Psi$ is weakly Borel-measurable, that is,

$$\Psi^I(O) := \{u \in U : \Psi(u) \cap O \neq \emptyset\}$$

is Borel for every open subset $O$ of $D_n$. Indeed, first notice that since every $X \in D_n$ is continuous and bounded, the mapping $U \ni u \mapsto \mathcal{E}(X)(u)$ is Borel by assumption. Moreover $\mathcal{F}(V) \ni P \mapsto E_P[X]$ is continuous which implies that $u \mapsto E_{R(u)}[X]$ is Borel. Therefore

$$U \ni u \mapsto \alpha^n_v(u, R(u)) = \max_u (E_{R(u)}[X] - \infty 1_{(0, \infty)}(\mathcal{E}(X)(u)))$$

is Borel, which implies that

$$\Psi^I(O) = \bigcup_{X \in O} \{u \in U : \alpha^n_v(u, R(u)) = E_{R(u)}[X] - \infty 1_{(0, \infty)}(\mathcal{E}(X)(u))\}$$

is Borel, as the finite union of Borel sets. Since $0 \in D_n$, it follows that $\Psi$ has nonempty values, and since the topology on $D_n$ is the discrete one, the values of $\Psi$ are closed as well. Therefore the Kuratowski–Ryll–Nardzewski theorem 2.7 Theorem 18.13 yields the existence of a Borel measurable mapping

$$X^\Psi: U \to D_n \text{ such that } X^\Psi(u) \in \Psi(u) \text{ for all } u \in U.$$ 

Define the function

$$Z(u, v) := Y(u) + X^\Psi(u)(v) \text{ for } (u, v) \in \Omega.$$ 

For every $u$, the mapping $Z(u, \cdot)$ is continuous by definition. Moreover, since $D_n$ is armed with the discrete topology, the mapping $Z(\cdot, v)$ is Borel for every $v \in V$. A
basic fact on Carathéodory-functions [3, Lemma 4.51] yields that $Z$ is jointly Borel and bounded, where the boundedness follows since $D_n$ is a finite set consisting of bounded functions. In particular $Z \in \mathcal{L}(\Omega)$. Using the dual representation of $\mathcal{E}(\cdot|U)$ and that $X^\Psi(u) \in \Psi(u)$ for each $u$, one obtains

$$\mathcal{E}(Z|U)(u) = Y(u) + \mathcal{E}(X^\Psi(u)|U)(u) \leq Y(u),$$

so that monotonicity of $\mathcal{E}(\cdot)$ together with the assumption that $\mathcal{E}(\cdot) \leq \mathcal{E}(\mathcal{E}(\cdot|U))$ imply $\mathcal{E}(Z) \leq \mathcal{E}(\mathcal{E}(Z|U)) \leq \mathcal{E}(Y)$ and therefore

$$E_P[Z] - \mathcal{E}(Z) \geq E_Q[Y] - \mathcal{E}(Y) + E_{Q(du)}[E_{R(u)}[X^\Psi(u)(\cdot)]]$$

$$= E_Q[Y] - \mathcal{E}(Y) + E_{Q(du)}[\alpha_V(u, R(u))].$$

Since $\mathcal{E}(Z) \geq E_P[Z] - \alpha(P)$ by the dual representation of $\mathcal{E}(\cdot)$, one has $\alpha(P) \geq E_P[Z] - \mathcal{E}(Z)$ and, as $Y$ and $n$ in (12) were arbitrary, it follows that $\alpha(P) \geq \beta(P)$. This concludes the proof. \hfill \square

**Proof of Theorem 2.10.** With the notation of Theorem 2.10 By Remark 2.11 one has $P = \{\alpha(\cdot) = 0\}$ and $P_V(u) = \{\alpha_V(u, \cdot) = 0\}$ and it follows from the definition of $\alpha_U$ that also $P_U := \{P(\cdot \times V) : P \in P\} = \{\alpha_U(\cdot) = 0\}$. Therefore the claim follows from Theorem 2.10. \hfill \square

**Remark 2.11.** Let $\mathcal{E}(\cdot|U)$ be a conditional sublinear expectation which satisfies (A) and (B) and therefore $\mathcal{E}(X|U)(u) = \sup_{P \in P_V(u)} E_P[X(u, \cdot)]$ for some $P_V : U \rightarrow \Psi(V)$ with convex and compact values. If $V$ is compact, then $\mathcal{E}(X|U)$ is Borel for $X \in C_b(\Omega)$ if and only if $P_V$ is weakly Borel-measurable.

**Proof.** Assume first that $P_V : U \rightarrow \Psi(V)$ is weakly Borel-measurable. Then, since it has nonempty, convex, and compact values, it admits a Castaing representation [2, Corollary 18.14]: There are Borel measurable mappings $R_n : U \rightarrow \Psi(V)$ (that is, Borel kernels) such that the closure of $\{R_n(u) : n \in \mathbb{N}\}$ equals $P_V(u)$ for every $u \in U$. Therefore

$$\mathcal{E}(X|U)(u) = \sup_n E_{R_n(u)}[X(u, \cdot)]$$

and as $u \mapsto E_{R_n(u)}[X(u, \cdot)]$ is Borel [6, Lemma 7.30] for every $n$, the claim follows.

On the other hand, assume that $\mathcal{E}(X|U)$ is Borel for every $X \in C_b(V)$ and let $O \subset \Psi(V)$ be open. Since the weak topology on $\Psi(V)$ is locally convex and metrizable, there are closed and convex sets $C_n \subset \Psi(V)$ such that $O = \bigcup\{C_n : n \in \mathbb{N}\}$. Therefore $P_V(u) \cap O = \emptyset$ if and only if $P_V(u) \cap C_n = \emptyset$ for all $n$. Now, as $P_V(u)$ is compact and convex for every $u \in U$ by assumption, the hyper plane separation theorem yields that $P_V(u) \cap C_n = \emptyset$ if and only if

$$\mathcal{E}(X|U)(u) = \sup_{P \in P_V(u)} E_P[X] < \inf_{P \in C_n} E_P[X]$$

for some $X \in C_b(V)$. Since $C_b(V)$ is separable (w.r.t. to the maximum norm), $X$ can in fact be chosen in some fixed (i.e. independent of $u$ and $n$) countable dense set $D \subset C_b(V)$. Therefore

$$\{u \in U : P_V(u) \cap O = \emptyset\} = \bigcap_n \bigcup_{X \in D} \left\{u \in U : \mathcal{E}(X|U)(u) < \inf_{P \in C_n} E_P[X]\right\}.$$

By assumption all sets on the right hand side are Borel, hence the countable union and intersection is Borel, too. \hfill \square
We conclude this section by explaining why the upper semianalytic functions (instead of e.g. Borel functions) are the natural domain and range for conditional nonlinear expectations. For an illustration on the basis of a concrete example (the G-Brownian motion) see [27], in particular Section 5.3 and Section 5.4 therein.

Remark 2.12. As already mentioned in the introduction, Theorem 2.6 should be an extension of Choquet’s theorem to the conditional case and therefore at least cover the case of conditional nonlinear expectations which are represented by a constant set-valued mapping (i.e. \(\mathcal{P}_V(u)\) in [4] does not depend on \(u\)). However, already in this setting it makes little sense to work with linear spaces instead of the semianalytic functions.

For example, the space \(\Omega := U \times V \times W = [0, 1] \times [0, 1] \times [0, 1]\) naturally carries two conditional nonlinear expectations \(\mathcal{E}(\cdot|U \times V)\) and \(\mathcal{E}(\cdot|U)\), say defined by

\[
\mathcal{E}(X|U \times V)(u, v) := \sup_{P \in \mathfrak{P}(W)} E^P[X(u, v, \cdot)], \quad \mathcal{E}(X|U)(u) := \sup_{P \in \mathfrak{P}(V \times W)} E^P[X(u, \cdot, \cdot)]
\]

wherever all integrals make sense. Now assume that there are linear spaces of (bounded) functions \(L(\Omega), L(U \times V),\) and \(L(U)\) such that \(L(\Omega)\) contains all Borel functions, and \(\mathcal{E}(\cdot|U \times V)\) and \(\mathcal{E}(\cdot|U)\) are mappings from \(L(\Omega)\) to \(L(U \times V)\) and \(L(U)\), respectively. By Borel isomorphy and the definition of analytic sets, every analytic set \(A \subset U \times V\) is the projection of some Borel set \(B \subset \Omega\). Since \(\mathcal{E}(1_B|U \times V) = 1_A\) and \(L(U \times V)\) is a linear space, it contains all complements of analytic sets. However, the projection of \(A^c\) on the first component (denoted by \(N \subset U\)) needs not to be universally measurable, and \(1_N = \mathcal{E}(1_{A^c}|U)\) implies that \(L(U)\) contains non-universally measurable functions (in fact even non-Lebesgue measurable ones [27, Section 5.4]). But this implies that even if \(\mathcal{E}(\cdot) := E^P[\cdot]\) for some \(P\), one cannot define \(\mathcal{E}(X)\) for \(X \in L(U)\).

3. Applications, extensions, and examples

3.1. The proof of Proposition 1.3. Assume first that \(u \mapsto \max_{P \in \mathcal{P}_V(u)} E^P[X]\) is upper semicontinuous for every \(X \in C_b(V)\) and the values of \(\mathcal{P}_V\) are compact. Define the functional

\[
\mathcal{E}(\cdot|U) : L(\Omega) \to \mathbb{R}^U, \quad \mathcal{E}(X|U)(u) := \sup_{P \in \mathcal{P}_V(u)} E^P[X(u, \cdot, \cdot)].
\]

Then, by convexity and compactness of each \(\mathcal{P}_V(u)\), the hyperplane separation theorem shows that

\[
\text{graph } \mathcal{P}_V = \{(u, P) \in U \times \mathfrak{P}(V) : E^P[X] \leq \mathcal{E}(X|U)(u) \text{ for all } X \in C_b(V)\}.
\]

Further, the same argumentation as in the proof of Theorem 2.6 show that one can restrict to all \(X\) in a countable set \(D \subset C_b(V)\). For every \(X \in D\) the mapping \((u, P) \to E^P[X] - \mathcal{E}(X|U)(u)\) is Borel by assumption so that graph \(\mathcal{P}_V\), as a countable intersection, is Borel and in particular analytic. Now Theorem 1.1 implies that \(\mathcal{E}(\cdot|U)\) is a conditional nonlinear expectation which satisfies (A) and (B); the same holds true for \(\mathcal{E}'(\cdot) : L(U) \to \mathbb{R}\) defined by \(\mathcal{E}'(X) := \sup_{P \in \mathcal{P}_U} E^P[X]\). Therefore \(\mathcal{E}(\cdot) := \mathcal{E}'(\cdot|U)\) defines a nonlinear expectation, which clearly satisfies (B) and we claim that it also satisfies (A). Indeed, let \(X_n \in C_b(\Omega)\) be a sequence which decreases pointwise to \(X \in L(\Omega)\). By assumption \(\mathcal{E}(X_n|U) \in usc_b(U)\) decreases pointwise to \(\mathcal{E}(X|U)\), therefore \(\mathcal{E}(X_n)\) decreases to \(\mathcal{E}(X)\) by Remark 2.3. Now Theorem 1.1 implies that \(\mathcal{E}(X) = \sup_{P \in \mathcal{P}} E^P[X]\) for a (by Remark 2.7 unique) convex
compact set $\mathcal{P} \subset \mathfrak{P}(\Omega)$. Since $\mathcal{E}(\cdot) = \mathcal{E}(\cdot|U)$ by definition, Theorem 1.2 yields that $\mathcal{P} = \mathcal{P}_U \otimes \mathcal{P}_V$, which proves the claim.

To show the reverse direction, assume that $\mathcal{P}_U \otimes \mathcal{P}_V$ is compact for every compact convex set $\mathcal{P}_U$. If $\mathcal{P}_V(u)$ is not compact for some $u \in U$, then neither is $\mathcal{P}_U \otimes \mathcal{P}_V$ for $\mathcal{P}_U := \{\delta_u\}$. So assume that $\mathcal{P}_V$ has compact values but $u \mapsto \max_{P \in \mathcal{P}_V(u)} E_P[X]$ is not upper semicontinuous for some $X \in C_b(V)$, i.e. there is $u \in U$ and a sequence $u_n \in U$ converging to $u$ such that

$$\limsup_{n} \max_{P \in \mathcal{P}_V(u_n)} E_P[X] > \max_{P \in \mathcal{P}_V(u)} E_P[X].$$

For every $n$, pick some $P_n \in \mathcal{P}_V(u_n)$ which attains the maximum in the left hand side of (13). After passing to a subsequence (still denoted by $P_n$), one may assume that $E_{P_n}[X]$ converges to the left hand side of (13). Since $C := \{u_n : n\} \cup \{u\} \subset U$ is compact, the set $\mathcal{P}_U := \{P \in \mathfrak{P}(U) : P(C) = 1\}$ is also compact (and obviously convex). Now distinguish between two cases. If $P_n$ does not have any convergent subsequence, then neither does $\delta_{(u_n)} \otimes P_n \in \mathcal{P}_U \otimes \mathcal{P}_V$ which implies that the latter set cannot be compact. Otherwise, possibly after passing to a subsequence, $P_n$ converges to some $P$ and one has $P \notin \mathcal{P}_V(u)$ by (13). However, since

$$\mathcal{P}_U \otimes \mathcal{P}_V \ni \delta_{(u_n)} \otimes P_n \rightarrow \delta_{(u)} \otimes P \notin \mathcal{P}_U \otimes \mathcal{P}_V,$$

this implies that $\mathcal{P}_U \otimes \mathcal{P}_V$ is not closed. The proof is complete.

Note that whenever $\mathcal{P}_V : U \rightarrow \mathfrak{P}(V)$ is upper hemi-continuous with nonempty compact values, by a variant of Berge’s maximum theorem \[2\, \text{Lemma 17.30}], the mapping $u \mapsto \max_{P \in \mathcal{P}_V(u)} E_P[X]$ is upper semicontinuous for $X \in C_b(V)$.

3.2. **Fubini’s theorem.** Let $\Omega = U \times V$ be the product of two Polish spaces. While Lemma 2.8 can be seen as a nonlinear version of Fubini’s theorem on the existence of the product of a measure and a kernel, one can also ask if there is a nonlinear version of Fubini’s classical theorem (i.e. on the possibility to interchange the order of integration when two measures are replaced by two sets of measures). In general this is not true any more (take for example $U := V := [0,1]$ as well as $\mathcal{P}_U := \{\{\delta_0 + \delta_1\}/2\}$ and $\mathcal{P}_V := \mathfrak{P}([0,1])$) but it is possible to characterize when interchanging the order is possible.

**Proposition 3.1.** Let $\mathcal{P}_U \subset \mathfrak{P}(U)$ and $\mathcal{P}_V \subset \mathfrak{P}(V)$ be two convex and compact sets of probabilities. Then it holds

$$\sup_{P \in \mathcal{P}_U} \sup_{Q \in \mathcal{P}_V} \int_U \int_V X(u, v) Q(dv) P(du) = \sup_{Q \in \mathcal{P}_V} \sup_{P \in \mathcal{P}_U} \int_U \int_V X(u, v) P(du) Q(dv)$$

for all $X \in \mathcal{L}(\Omega)$ if and only if

$$\{P \otimes R : P \in \mathcal{P}_U \text{ and } R : U \rightarrow \mathfrak{P}(V) \text{ kernel with } R(\cdot) \in \mathcal{P}_V P\text{-as}\}$$

$$= \{(Q \otimes R) \circ \pi^{-1} : Q \in \mathcal{P}_V \text{ and } R : V \rightarrow \mathfrak{P}(U) \text{ kernel with } R(\cdot) \in \mathcal{P}_U Q\text{-as}\},$$

where $\pi : U \times V \rightarrow V \times U$ is given by $\pi(u,v) := (v,u)$.

**Proof.** The proof is similar to the one given for Proposition 1.2 but somewhat notationally involved, and we shall skip it. \qed
3.3. Risk measures under Knightian uncertainty. The notation is as in [8]. Fix some finite time horizon $T \in \mathbb{N}$ and define $\Omega_t := \Omega_1^t$ for $t = 1, \ldots, T$, where $\Omega_1$ is some Polish space. For every $t$, let $\mathcal{E}_t(\cdot): \mathcal{L}(\Omega_{t+1}) \to \mathcal{L}(\Omega_t)$ be a conditional sublinear expectation which satisfies the (A) and (B) and therefore has the representation $\mathcal{E}_t(X)(\omega) = \sup_{P \in \mathcal{P}_t(\omega)} E_P[X(\omega, \cdot)]$ for $\omega \in \Omega_t$ and $X \in \mathcal{L}(\Omega_{t+1})$ by Theorem 1.1 for a (unique) set-valued mapping $\mathcal{P}_t: \Omega_t \rightrightarrows \mathcal{P}(\Omega_t)$ with convex compact values and analytic graph; see e.g. [8] Chapter 2.3 for concrete examples. Then $\mathcal{E}_{t,T}(\cdot) := \mathcal{E}_t \circ \cdots \circ \mathcal{E}_{T-1}(\cdot)$ defines a sublinear expectation from $\mathcal{L}(\Omega_T)$ to $\mathcal{L}(\Omega_t)$ with representation

$$\mathcal{E}_{t,T}(X)(\omega) = \sup_{P \in \mathcal{P}_{t,T}(\omega)} E_P[X(\omega, \cdot)] \quad \text{for } \omega \in \Omega_t \text{ and } X \in \mathcal{L}(\Omega_T),$$

where $\mathcal{P}_{t,T}(\omega) := \mathcal{P}_t(\omega) \otimes \cdots \otimes \mathcal{P}_{T-1}(\omega, \cdot)$, see Lemma 2.8. Now let $l: \mathbb{R} \to \mathbb{R}$, $x \mapsto x^+/\lambda$ for some $\lambda \in (0,1)$, and define the time-consistent robust average value at risk by $\mathcal{R}_{t,T}(\cdot) := \mathcal{R}_t \circ \cdots \circ \mathcal{R}_{T-1}(\cdot)$, where

$$\mathcal{R}_t(X)(\omega) := \inf_{s \in \mathbb{R}} \left( \mathcal{E}_t(l(X-s))(\omega) + s \right)$$

for $\omega \in \Omega_t$ and $X \in \mathcal{L}(\Omega_{t+1})$; see e.g. [11] Example 2.3.1 for the case without Knightian uncertainty. Note that one can replace $l$ by a general loss function and therefore define time-consistent robust optimized certainty equivalents.

**Example 3.2.** For every $t$, the functional $\mathcal{R}_t(\cdot)$ is a conditional sublinear expectation which satisfies condition (A) and (B) and has the representation

$$\mathcal{R}_t(X)(\omega) = \sup_{Q \in \mathcal{Q}_t(\omega)} E_Q[X(\omega, \cdot)] \quad \text{for } \omega \in \Omega_t \text{ and } X \in \mathcal{L}(\Omega_t),$$

where $\mathcal{Q}_t(\omega)$ is the set of all probabilities $Q \in \mathcal{P}(\Omega_t)$ for which there exists $P \in \mathcal{P}_t(\omega)$ such that $Q$ is absolutely continuous w.r.t. $P$ and the Radon-Nykodim derivative $dQ/dP$ is bounded by $1/\lambda$. Moreover

$$\mathcal{R}_{t,T}(X)(\omega) = \sup_{Q \in \mathcal{Q}_{t,T}(\omega)} E_Q[X(\omega, \cdot)] \quad \text{for } \omega \in \Omega_t \text{ and } X \in \mathcal{L}(\Omega),$$

where $\mathcal{Q}_{t,T}(\omega) := \mathcal{Q}_t(\omega) \otimes \cdots \otimes \mathcal{Q}_{T-1}(\omega, \cdot)$.

**Proof.** Since $l$ is increasing, one has $(l(X-s)) \in \mathcal{L}(\Omega_{t+1})$ for every $X \in \mathcal{L}(\Omega_{t+1})$ and $s \in \mathbb{R}$ so that $\mathcal{R}_t(\cdot)$ is well-defined. Moreover, as $s \mapsto \mathcal{E}_t(l(X-s))(\omega) + s$ is convex and real-valued for every $\omega \in \Omega_t$, one may restrict the infimum in (14) to $s \in \mathbb{Q}$. Therefore, as the countable infimum, $\mathcal{R}_t(X)$ is upper semianalytic whenever $X$ is. Elementary computations show that $\mathcal{R}_t(\cdot)$ is increasing, sublinear, and satisfies $\mathcal{R}_t(X) = X$ for $X \in \mathcal{L}(\Omega_t)$. By interchanging two infima and the fact that $\mathcal{E}_t(\cdot)$ satisfies (A), one gets that $\mathcal{R}_t(\cdot)$ satisfies (A) as well. As for condition (B), fix $\omega \in \Omega_t$ and let $X_n \in \mathcal{L}(\Omega_{t+1})$ be a sequence which increases pointwise to $X \in \mathcal{L}(\Omega_{t+1})$. For every $n$, let $s_n \in \mathbb{R}$ such that $\mathcal{E}_t(l(X_n-s_n))(\omega) + s_n \leq \mathcal{R}_t(X_n)(\omega) + 1/n$. Since $(X_n)$ is bounded and $\lambda \in (0,1)$, it follows that $(s_n)$ converges. The dual representation of $\mathcal{E}_t(\cdot)$ now implies that

$$\mathcal{R}_t(X)(\omega) \leq \mathcal{E}_t(l(X-s))(\omega) + s \leq \liminf_n \left( \mathcal{E}_t(l(X_n-s_n))(\omega) + s_n \right) \leq \mathcal{R}_t(X)(\omega),$$

where the last inequality holds due to the choice of $s_n$ and $\mathcal{R}_t(X_n) \leq \mathcal{R}_t(X)$. Hence $\mathcal{R}_t(\cdot)$ satisfies condition (B). The specific form of $\mathcal{Q}_t(\omega)$ follows as in the case without Knightian uncertainty [20] Lemma 4.51 and Theorem 4.52, additionally.
using a suitable minimax theorem [19] Theorem 2. [18] The representation of $\mathcal{R}_{t,T}(\cdot)$ is due to Lemma 2.8.

For further literature which uses dynamic programming and conditional nonlinear expectations in the context of mathematical finance under Knightian uncertainty in discrete time, see e.g. [1] 7 [9] 10 [25] 28.

3.4. Controlled Brownian motion. This last example is neither new nor presented in the most general form [27] 29, its purpose is to illustrate how the results can be applied in a continuous time setting. For some fixed time horizon $T > 0$ let $\Omega := C([0,T], \mathbb{R})$. Then, for every $t \in (0, T)$, there is a natural product structure $\Omega \cong U \times V$ for $U := C([0, t], \mathbb{R})$ and $V := \{v \in C([t, T], \mathbb{R}) : v(t) = 0\}$ via

$$(u \otimes t)(s) := u(s)1_{[0,t]}(s) + (v(s) + u(t))1_{(t,T]}(s) \quad \text{for } s \in [0,T]$$

for $(u, v) \in U \times V$. Let $B$ be the canonical process on $\Omega$, denote by $W$ the Wiener measures, fix two numbers $0 < \alpha < \beta$, and write $\Sigma_t$ for the set of all processes $\sigma : [t, T] \times C([t, T]) \to \mathbb{R}$ adapted to $B$ which satisfy $\sigma \in [\alpha^2, \beta^2] W \times dt$-almost surely. For $\sigma \in \Sigma_t$, denote by $B^{u,t,\sigma} := u1_{[0,t]} + (u(t) + \int_t^T \sigma_s dB_s)1_{(t,T]}$ the Brownian motion with volatility $\sigma$ starting in $(t, u(t))$.

**Example 3.3.** The functional

$$\mathcal{E}(X|U)(u) := \sup_{\sigma \in \Sigma_t} E_W[X(B^{u,t,\sigma})] \quad \text{for } u \in U \text{ and } X \in \mathcal{L}(\Omega).$$

defines a conditional sublinear expectation.

**Proof.** The only non-trivial property is that $\mathcal{E}(\cdot|U)$ is mapping from $\mathcal{L}(\Omega)$ to $\mathcal{L}(U)$. To show this, endow $\Sigma_t$ with the norm $\|\sigma\| := E_W[\int_t^T \sigma_s^2 ds]^{1/2}$ which renders $\Sigma_t$ a Polish space, and notice that the mapping

$$U \times \Sigma_t \ni (u, \sigma) \mapsto P^{u,t,\sigma} := W \circ (B^{u,t,\sigma})^{-1} \in \mathfrak{P}(\Omega)$$

is Borel. Indeed, for Lipschitz-continuous $X : \Omega \to \mathbb{R}$, Doob’s inequality shows that $(u, \sigma) \mapsto E_{P^{u,t,\sigma}}[X]$ is continuous, and a monotone class argument yields that $(u, \sigma) \mapsto E_{P^{u,t,\sigma}}[X]$ is Borel whenever $X$ is. Thus $(u, \sigma) \mapsto P^{u,t,\sigma}$ is Borel [6] Proposition 7.26], hence for every $X \in \mathcal{L}(\Omega)

$$(u, \sigma) \mapsto E_{P^{u,t,\sigma}}[X] = E_W[X(B^{u,t,\sigma})]$$

is upper semianalytic [6] Proposition 7.48]. Since $U \times \Sigma_t$ is a Polish space, the claim follows from [6] Proposition 7.47]. \qed

**APPENDIX A. ANALYTIC SETS**

In the sequel let $\Omega$ and $\Omega'$ be two Polish spaces. Recall that a subset of a Polish space is called analytic if it is the image of a Borel set of another Polish space under a Borel function. A function $f : \Omega \to [-\infty, +\infty]$ is upper semianalytic, if $\{f \geq c\} \subset \Omega$ is an analytic set for every $c \in \mathbb{R}$ and lower semianalytic, if $-f$ is upper semianalytic. The set of universally measurable subsets of $\Omega$ is defined as $\bigcap\{B(\Omega)^P : P \in \mathfrak{P}(\Omega)\}$, where $B(\Omega)^P$ is the completion of the Borel $\sigma$-field $B(\Omega)$ with respect to the probability $P$. A function $f : \Omega \to \Omega'$ is universally measurable, if $\{f \in B\}$ is universally measurable for every $B \in B(\Omega')$. It follows from the definition that every Borel set is analytic, and from Lusin’s theorem [6] Proposition 7.42] that every analytic set is universally measurable. The same of course holds

\begin{itemize}
  \item \text{16 DANIEL BARTL} \text{\smallskip}
\end{itemize}
true if sets are replaced by functions in the previous sentence. Finally, a set-valued mapping $\Phi: \Omega \rightharpoonup \Omega'$ has analytic graph, if

$$\text{graph} \Psi := \{ (\omega, \omega') \in \Omega \times \Omega' : \omega' \in \Phi(\omega) \}$$

is an analytic subset of the Polish space $\Omega \times \Omega'$. For a countable family $\{X_n : n\}$ of upper semianalytic functions, $X_1 + X_2, \sup_n X_n, \text{and} \inf_n X_n$ are again upper semianalytic. Moreover, if $Y$ is a positive Borel function, then $YX_1$ is upper semianalytic, see e.g. [6 Lemma 7.30] for both statements. A comprehensive treatment of analytic sets well suited for the present setting can be found in [6 Chapter 7].

**APPENDIX B. CHOQUET’S THEOREM**

For convenience, we briefly sketch the proof of Choquet’s theorem for nonlinear expectations; a detailed proof is given e.g. in [4 Section 2]. Let $\mathcal{E}(\cdot): \mathcal{L}(\Omega) \rightarrow \mathbb{R}$ be a increasing convex functional which preserves the constants, is continuous from above on $C_b(\Omega)$, and continuous from below on $\mathcal{L}(\Omega)$. By monotonicity, $\mathcal{E}(\cdot)$ is continuous w.r.t. the maximum norm on $C_b(\Omega)$ so that the Fenchel-Moreau / Hahn-Banach theorem implies

$$\mathcal{E}(X) = \max_{P \in C_b(\Omega)^*} \left( (X, P) - \mathcal{E}^*(P) \right) \quad \text{for } X \in C_b(\Omega),$$

where $\mathcal{E}^*(P) := \sup_{X \in C_b(\Omega)} (X, P) - \mathcal{E}(X)$ and $C_b(\Omega)^*$ denotes the topological dual of $(C_b(\Omega), \| \cdot \|_{\infty})$. For every $P$ with $\mathcal{E}^*(P) < +\infty$ a scaling argument implies that $P$ needs to be a increasing functional satisfying $\langle 1, P \rangle = 1$. Moreover, since $\mathcal{E}(\cdot)$ is continuous from above on $C_b(\Omega)$, one can show that $P$ has this property as well. Therefore, by the Daniell-Stone theorem, $P$ can be viewed as a probability on $\sigma(C_b(\Omega)) = \mathcal{B}(\Omega)$ and $\langle X, P \rangle = E_P[X]$ for all $X \in C_b(\Omega)$ (in fact, a combination of the Banach-Alaoglu and Daniell-Stone theorem even gives that $\{\mathcal{E}^* \leq c\} \subset C_b(\Omega)$ is compact for $c \in \mathbb{R}$). This implies that

$$\mathcal{E}(X) = \max_{P \in \mathcal{F}(\Omega)} (E_P[X] - \mathcal{E}^*(P)) \quad \text{for } X \in C_b(\Omega)$$

and, using a minimax theorem (similar as in the proof of Theorem 2.6), this equality extends to $X \in usc_b(\Omega)$. Now notice that $\mathcal{E}(\cdot)$ is a (functional) capacity in the sense of Choquet by assumption, therefore his regularity result yield

$$\mathcal{E}(X) = \sup_{Y \leq X, Y \in usc_b(\Omega)} \mathcal{E}(Y) = \sup_{Y \leq X, Y \in usc_b(\Omega)} \max_{P \in \mathcal{F}(\Omega)} (E_P[Y] - \mathcal{E}^*(P))$$

for every function $X$ which can be written as the nucleus of a Sulin scheme in $C_b(\Omega)$, see [12 Section 3]. Since $\Omega$ is Polish, this is possible precisely when $X$ is upper semianalytic. The representation $\mathcal{E}(X) = \sup_{P \in \mathcal{F}(\Omega)} (E_P[X] - \mathcal{E}^*(P))$ for $X \in \mathcal{L}(\Omega)$ now follows from the representation of $\mathcal{E}(Y)$ for $Y \in usc_b(\Omega)$, interchanging two suprema, and using the fact that $\sup_{Y \leq X, Y \in usc_b(\Omega)} E_P[Y] = E_P[X]$ (to see this apply for example Choquet’s results to $E_P[\cdot]$).

**REFERENCES**

[1] A. Aksamit, S. Deng, J. Obloj, and X. Tan. Robust pricing–hedging duality for American options in discrete time financial markets. arXiv preprint [arXiv:1604.05517] 2016.

[2] C. Aliprantis and K. Border. *Infinite dimensional analysis: a hitchhiker’s guide*. Springer Science & Business Media, 2006.

[3] D. Bartl. Exponential utility maximization under model uncertainty for unbounded endowments. arXiv preprint [arXiv:1610.00999] 2016.
[4] D. Bartl, P. Cheridito, and M. Kupper. Robust expected utility maximization with medial limits. in preparation.

[5] M. Beiglböck, M. Nutz, and N. Touzi. Complete duality for martingale optimal transport on the line. The Annals of Probability forthcoming, arXiv:1507.00671, 2015.

[6] D. Bertsekas and S. Shreve. Stochastic optimal control: The discrete time case, volume 23. Academic Press New York, 1978.

[7] R. Blanchard and L. Carassus. Robust optimal investment in discrete time for unbounded utility function. The Annals of Applied Probability forthcoming, arXiv:1609.09205, 2016.

[8] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. The Annals of Applied Probability, 25(2):823–859, 2015.

[9] B. Bouchard, S. Deng, and X. Tan. Super-replication with proportional transaction cost under model uncertainty. arXiv preprint arXiv:1707.09158, 2017.

[10] M. Burzoni, M. Frittelli, Z. Hou, M. Maggis, and J. Obłój. Pointwise Arbitrage Pricing Theory in Discrete Time. arXiv preprint arXiv:1612.07618, 2016.

[11] P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. International Journal of Theoretical and Applied Finance, 14(01):137–162, 2011.

[12] G. Choquet. Forme abstraite du théorème de capacitaibilité. In Annales de l'institut Fourier, volume 9, pages 83–89, 1959.

[13] S. Cohen. Quasi-sure analysis, aggregation and dual representations of sublinear expectations in general spaces. Electron. J. Probab, 17(62):1–15, 2012.

[14] F. Delbaen. Coherent risk measures on general probability spaces. In Advances in finance and stochastics, pages 1–37. Springer, 2002.

[15] C. Dellacherie and P. Meyer. Probabilities and Potential, C: Potential Theory for Discrete and Continuous Semigroups, volume 151. Elsevier, 2011.

[16] L. Denis, M. Hu, and S. Peng. Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. Potential Analysis, 32(2):139–161, 2011.

[17] R. Denk, M. Kupper, M. Nendel. A semigroup approach to nonlinear Lévy processes. arXiv preprint arXiv:1710.08130, 2017.

[18] S. Eckstein. Extended Laplace Principle for Empirical Measures of a Markov Chain. arXiv preprint arXiv:1709.02278, 2017.

[19] K. Fan. Minimax theorems. Proceedings of the National Academy of Sciences, 39(1):42–47, 1953.

[20] H. Föllmer and A. Schied. Stochastic finance: an introduction in discrete time. Walter de Gruyter, 2011.

[21] M. Katětov. On real-valued functions in topological spaces. Fundamenta Mathematicae, 38(1):85–91, 1951.

[22] H. Kellerer. Duality theorems for marginal problems. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 67(4):399–432, 1984.

[23] D. Lacker. A non-exponential extension of sanov’s theorem via convex duality. arXiv preprint arXiv:1609.04744, 2016.

[24] M. Maggis, T. Meyer-Brandis, and G. Svindland. The fatou property under model uncertainty. arXiv preprint arXiv:1610.02085, 2016.

[25] A. Neufeld and M. Šišić. Robust utility maximization in discrete-time markets with friction. arXiv preprint arXiv:1610.09290, 2016.

[26] A. Neufeld and M. Nutz. Nonlinear Lévy processes and their characteristics. Transactions of the American Mathematical Society, 369(1):69–95, 2017.

[27] M. Nutz and R. van Handel. Constructing sublinear expectations on path space. Stochastic Processes and their Applications, 123(8):3100–3121, 2013.

[28] M. Nutz. Utility maximization under model uncertainty in discrete time. Mathematical Finance, 26(2):252–268, 2016.

[29] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. Stochastic Processes and their Applications, 118(12):2223–2253, 2008.

[30] S. Peng. Nonlinear expectations and stochastic calculus under uncertainty. arXiv preprint arXiv:1002.4546, 2010.

[31] D. Stroock. Probability theory: an analytic view. Cambridge university press, 2010.