

Infinite exchangeability for sets of desirable gambles

Gert de Cooman¹ and Erik Quaeghebeur^{1,2}

¹ Ghent University, SYSTeMS Research Group
Technologiepark–Zwijnaarde 914, 9052 Zwijnaarde, Belgium
{Gert.deCooman,Erik.Quaeghebeur}@UGent.be

² Department of Philosophy, Baker Hall 135, Carnegie Mellon University
Pittsburgh, PA 15213-3890, United States
ErikQ@andrew.cmu.edu

Abstract Sets of desirable gambles constitute a quite general type of uncertainty model with an interesting geometrical interpretation. We study infinite exchangeability assessments for them, and give a counterpart of de Finetti’s infinite representation theorem. We show how the infinite representation in terms of frequency vectors is tied up with multivariate Bernstein (basis) polynomials. We also lay bare the relationships between the representations of updated exchangeable models, and discuss conservative inference (natural extension) under exchangeability.

Key words: desirability, weak desirability, sets of desirable gambles, coherence, exchangeability, representation, natural extension, updating.

1 Introduction

In this paper, we bring together desirability,³ an interesting approach to modelling uncertainty, with infinite exchangeability, a structural assessment for uncertainty models that is important for inference purposes.

Desirability, or the theory of (coherent) sets of desirable gambles, represents the uncertainty of a subject with a set of gambles⁴ that she finds desirable. This theory is more expressive than the theory of (coherent) lower previsions [11], which itself is a generalization of the theory of linear previsions [5]. A complete model for a rational subject’s uncertainty is a coherent set of desirable gambles, or mathematically, a cone satisfying some constraints. This geometric aspect is one of the points that make desirability appealing. Another is its generality [12]. We introduce the necessary desirability-related concepts in Sec. 2.

Here we study infinite exchangeability³ for sets of desirable gambles. An exchangeability assessment expresses that the order of the samples in a sequence of them is irrelevant for inference purposes. This study builds on earlier work on exchangeability for coherent lower previsions [2] and finite exchangeability for

³ For a brief historical overview, see some of our earlier work [1, Sec. 1].

⁴ Gambles are also called bets or random rewards.

sets of desirable gambles [1]. We need to recall the definition and representation results for finite exchangeability of sets of desirable gambles (Sec. 3). After that, we can give our definition and representation result of infinite exchangeability for sets of desirable gambles and give some results about updating and natural extension under exchangeability (Sec. 4).

We end with some conclusions (Sec. 5).

2 Desirability

Consider an experiment with a non-empty set Ω describing its mutually exclusive possible outcomes, and a subject who is uncertain about its outcome.

Sets of desirable gambles A *gamble* f is a bounded real-valued map on Ω , and it is interpreted as an uncertain reward. When the actual outcome of the experiment is ω , then the corresponding (possibly negative) reward is $f(\omega)$. The set of all gambles is $\mathcal{G}(\Omega)$.

We say that a non-zero gamble f is *desirable* to a subject if she accepts to engage in the following transaction, where: (i) the actual outcome ω of the experiment is determined, and (ii) she receives the reward $f(\omega)$. The zero gamble is not considered to be desirable.⁵

We try to model the subject's beliefs about the outcome of the experiment by considering which gambles are desirable for her. We suppose the subject has some set $\mathcal{R} \subseteq \mathcal{G}(\Omega)$ of desirable gambles.

Coherence Not every such set should be considered as a reasonable model, and in what follows, we give an abstract and fairly general treatment of ways to impose 'rationality' constraints on sets of desirable gambles.

The set $\mathcal{G}(\Omega)$ of all gambles on Ω is a linear space with respect to the (point-wise) addition of gambles, and the (point-wise) scalar multiplication of gambles with real numbers. The *positive hull operator* posi generates the set of finite *strictly* positive linear combinations of elements of its argument set. A subset \mathcal{C} of $\mathcal{G}(\Omega)$ is a *convex cone* if $\text{posi}(\mathcal{C}) = \mathcal{C}$.

Consider a linear subspace \mathcal{K} of the linear space $\mathcal{G}(\Omega)$. With any convex cone $\mathcal{C} \subset \mathcal{K}$ such that $0 \in \mathcal{C}$ we can always associate a *vector ordering* \succeq on \mathcal{K} , defined by $f \succeq g \Leftrightarrow f - g \in \mathcal{C} \Leftrightarrow f - g \succeq 0$. The partial ordering \succeq turns \mathcal{K} into an ordered linear space [10, Section 11.44]. We also write $f \succ g$ if $f \succeq g$ and $f \neq g$. Finally, we let $\mathcal{K}_{\preceq 0} := \{f \in \mathcal{K} : f \preceq 0\} = -\mathcal{C}$ and $\mathcal{K}_{\succ 0} := \{f \in \mathcal{K} : f \succ 0\} = \mathcal{C}_0$.⁶

Definition 1 (Avoiding non-positivity and coherence). *Let \mathcal{K} be a linear subspace of $\mathcal{G}(\Omega)$ and let $\mathcal{C} \subset \mathcal{K}$ be a convex cone containing the zero gamble 0.*

A set of desirable gambles $\mathcal{R} \subseteq \mathcal{K}$ avoids non-positivity relative to $(\mathcal{K}, \mathcal{C})$ if $f \not\preceq 0$ for all gambles f in $\text{posi}(\mathcal{R})$, or in other words if $\mathcal{K}_{\preceq 0} \cap \text{posi}(\mathcal{R}) = \emptyset$.

⁵ For clarification on the confusing nomenclature in the literature, see [1, footnote 2].

⁶ Subscripting a set of gambles with zero *removes* the zero gamble, if present.

A set of desirable gambles $\mathcal{R} \subseteq \mathcal{K}$ is coherent relative to $(\mathcal{K}, \mathcal{C})$ if it satisfies the following requirements, for all gambles f , f_1 , and f_2 in \mathcal{K} and all real $\lambda > 0$:

- D1. if $f = 0$ then $f \notin \mathcal{R}$;
- D2. if $f \succ 0$ then $f \in \mathcal{R}$, or equivalently $\mathcal{K}_{\succ 0} \subseteq \mathcal{R}$;
- D3. if $f \in \mathcal{R}$ then $\lambda f \in \mathcal{R}$ [scaling];
- D4. if $f_1, f_2 \in \mathcal{R}$ then $f_1 + f_2 \in \mathcal{R}$ [combination].

We denote by $\mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)$ the set of sets of desirable gambles that are coherent relative to $(\mathcal{K}, \mathcal{C})$.

These requirements make \mathcal{R} a cone that excludes $\mathcal{K}_{\leq 0}$:

- D5. if $f \leq 0$ then $f \notin \mathcal{R}$, or equivalently $\mathcal{K}_{\leq 0} \cap \mathcal{R} = \emptyset$.

We see that \mathcal{K} is never, and $\mathcal{K}_{\succ 0}$ is always coherent relative to $(\mathcal{K}, \mathcal{C})$.

Natural extension If we consider an arbitrary non-empty family of sets of desirable gambles that are coherent relative to $(\mathcal{K}, \mathcal{C})$, then so is their intersection. If a subject gives us an *assessment*, a set $\mathcal{A} \subseteq \mathcal{K}$ of gambles on Ω that she finds desirable, then the following theorem tells us exactly when this assessment can be extended to a coherent set, and how to construct the smallest such set.

Theorem 1 (Natural extension). *Let \mathcal{K} be a linear subspace of $\mathcal{G}(\Omega)$ and let $\mathcal{C} \subset \mathcal{K}$ be a convex cone containing the zero gamble 0.*

Consider an assessment $\mathcal{A} \subseteq \mathcal{K}$, and define its $(\mathcal{K}, \mathcal{C})$ -natural extension: $\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A}) := \bigcap \{ \mathcal{R} \in \mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega) : \mathcal{A} \subseteq \mathcal{R} \}$, with $\bigcap \emptyset = \mathcal{K}$. Then the following statements are equivalent:

- (i) \mathcal{A} avoids non-positivity relative to $(\mathcal{K}, \mathcal{C})$;
- (ii) \mathcal{A} is included in some set of desirable gambles that is coherent relative to $(\mathcal{K}, \mathcal{C})$;
- (iii) $\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A}) \neq \mathcal{K}$;
- (iv) the set of desirable gambles $\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A})$ is coherent relative to $(\mathcal{K}, \mathcal{C})$;
- (v) $\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A})$ is the smallest set of desirable gambles that is coherent relative to $(\mathcal{K}, \mathcal{C})$ and includes \mathcal{A} .

When any of these equivalent statements holds, then $\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A}) = \text{posi}(\mathcal{K}_{\succ 0} \cup \mathcal{A})$.

Point-wise comparison coherence We now turn to the important special case, commonly considered in the literature [12], where $\mathcal{K} := \mathcal{G}(\Omega)$ and the partial order \succeq is the point-wise ordering \geq .⁷ This partial order is associated to $\mathcal{C} := \mathcal{G}_0^+(\Omega) := \mathcal{G}(\Omega)_{> 0}$, the cone of all non-negative gambles.

If \mathcal{R} avoids non-positivity relative to $(\mathcal{G}(\Omega), \mathcal{G}_0^+(\Omega))$, we simply say that \mathcal{R} avoids non-positivity: $\mathcal{G}^-(\Omega) \cap \text{posi}(\mathcal{R}) = \emptyset$, where $\mathcal{G}^-(\Omega) := \mathcal{G}(\Omega)_{\leq 0}$ is the set of all non-positive gambles. Similarly, if \mathcal{R} is coherent relative to $(\mathcal{G}(\Omega), \mathcal{G}_0^+(\Omega))$, we simply say that \mathcal{R} is *coherent*, and we denote the set of coherent sets of desirable gambles by $\mathbb{D}(\Omega)$. In this case, the coherence conditions D1–D5 are to be seen as *rationality criteria*. The $(\mathcal{G}(\Omega), \mathcal{G}_0^+(\Omega))$ -natural extension of an assessment $\mathcal{A} \subseteq \mathcal{G}(\Omega)$ is simply denoted by $\mathcal{E}(\mathcal{A})$, and is called the *natural extension* of \mathcal{A} .

⁷ $f \geq g$ iff $f(\omega) \geq g(\omega)$ for all ω in Ω ; $f > g$ iff $f \geq g$ and $f \neq g$.

Weakly desirable gambles We now define *weak desirability*, a concept that lies at the basis of our definition of exchangeability.

Definition 2 (Weak desirability). *Consider a coherent set \mathcal{R} of desirable gambles. Then a gamble f is called weakly desirable if $f + f'$ is desirable for all desirable f' , i.e., if $f + f' \in \mathcal{R}$ for all f' in \mathcal{R} . So the set of weakly desirable gambles is $\mathcal{D}_{\mathcal{R}} := \{f \in \mathcal{G}(\Omega) : f + \mathcal{R} \subseteq \mathcal{R}\}$.*

Every desirable gamble is also weakly desirable, so $\mathcal{R} \subseteq \mathcal{D}_{\mathcal{R}}$. Moreover, $\mathcal{D}_{\mathcal{R}}$ also satisfies the scaling and combination requirements D3–D4, so it is a cone as well.

Updating sets of desirable gambles Consider a set of desirable gambles \mathcal{R} on Ω . With a non-empty subset B of Ω , we associate an *updated* set of desirable gambles $\mathcal{R} \downarrow B := \{f_B : f = I_B f \in \mathcal{R}\} \subseteq \mathcal{G}(B)$,⁸ where f_B is the restriction of f to B and I_B is the indicator function of B , i.e., 1 on B and 0 elsewhere. $\mathcal{R} \downarrow B$ is our subject's set of desirable gambles contingent on observing the event B .

Proposition 1. *If \mathcal{R} is a coherent set of desirable gambles on Ω , then $\mathcal{R} \downarrow B$ is a coherent set of desirable gambles on B .*

3 Finite exchangeable sequences

Now that we have familiarised ourselves with sets of desirable gambles, we turn to exchangeability. In this section, we recall the basic definitions and results about finite exchangeable sequences from our earlier work [1], and add some new material related to frequency vector representations.

Consider random variables X_1, \dots, X_N taking values in a non-empty finite set \mathcal{X} , where $N \in \mathbb{N}_0$, i.e., a positive integer. The possibility space is $\Omega = \mathcal{X}^N$.

Let $x = (x_1, \dots, x_N)$ be an arbitrary element of \mathcal{X}^N . \mathcal{P}_N is the set of all permutations of the index set $\{1, \dots, N\}$. With any such permutation π , we associate a permutation of \mathcal{X}^N , also denoted by π , and defined by $(\pi x)_k = x_{\pi(k)}$. Similarly, we lift π to a permutation π^t of $\mathcal{G}(\mathcal{X}^N)$ by letting $\pi^t f = f \circ \pi$.

The *counting map* T^N maps a sequence x to its count vector, an \mathcal{X} -tuple with a z -component $T_z^N(x) := |\{k \in \{1, \dots, N\} : x_k = z\}|$ for all z in \mathcal{X} . The set of possible *count vectors* is given by $\mathcal{N}^N := \{m \in \mathbb{N}^{\mathcal{X}} : \sum_{x \in \mathcal{X}} m_x = N\}$. The permutation invariant atoms $[x] := \{\pi x : \pi \in \mathcal{P}_N\}$ are the smallest permutation invariant subsets of \mathcal{X}^N . If $m = T^N(x)$, then $[x] = \{y \in \mathcal{X}^N : T^N(y) = m\}$, so the atom $[x]$ is completely determined by the count vector m of all its the elements, and is therefore also denoted by $[m]$.

Defining exchangeability If a subject assesses that X_1, \dots, X_N are exchangeable, this means that for any gamble f and any permutation π , she finds exchanging $\pi^t f$ for f weakly desirable, because she is indifferent between them [11, Section 4.1.1]. Taking into account that $\mathcal{D}_{\mathcal{R}}$ is a cone, we introduce the linear⁹ space $\mathcal{D}_{\mathcal{U}_N} := \text{posi} \{f - \pi^t f : f \in \mathcal{G}(\mathcal{X}^N) \text{ and } \pi \in \mathcal{P}_N\}$. It holds that $\mathcal{R} \cap \mathcal{D}_{\mathcal{U}_N} = \emptyset$.

⁸ Our definition is different from, but equivalent to the usual one [12].

⁹ Due to the negation invariance of posi 's argument.

Definition 3 (Exchangeability). *A coherent set \mathcal{R} of desirable gambles on \mathcal{X}^N is called exchangeable if one (and hence both) of the following equivalent conditions is (are) satisfied: (i) all gambles in $\mathcal{D}_{\mathcal{U}_N}$ are weakly desirable: $\mathcal{D}_{\mathcal{U}_N} \subseteq \mathcal{D}_{\mathcal{R}}$; and (ii) $\mathcal{D}_{\mathcal{U}_N} + \mathcal{R} \subseteq \mathcal{R}$.*

Updating exchangeable models Consider an exchangeable and coherent set of desirable gambles \mathcal{R} on \mathcal{X}^N . Assume that the subject has observed the values $\check{x} = (\check{x}_1, \check{x}_2, \dots, \check{x}_{\check{n}})$ or a count vector $\check{m} \in \mathcal{N}^{\check{n}}$ of the variables $X_1, \dots, X_{\check{n}}$. She wants to make inferences about the remaining $\hat{n} := N - \check{n}$ variables. Consider the updated models $\mathcal{R} \downarrow \{\check{x}\} \times \mathcal{X}^{\hat{n}}$ and $\mathcal{R} \downarrow [\check{m}] \times \mathcal{X}^{\hat{n}}$, and their restrictions to these \hat{n} variables, $\mathcal{R} \downarrow \check{x}$ and $\mathcal{R} \downarrow \check{m}$.

Proposition 2. *Consider \check{x} in $\mathcal{X}^{\check{n}}$, \check{m} in $\mathcal{N}^{\check{n}}$, and a coherent and exchangeable set of desirable gambles \mathcal{R} on \mathcal{X}^N . Then $\mathcal{R} \downarrow \check{x}$ and $\mathcal{R} \downarrow \check{m}$ are coherent and exchangeable sets of desirable gambles on $\mathcal{X}^{\hat{n}}$. If $\check{m} = T^{\check{n}}(\check{x})$, then $\mathcal{R} \downarrow \check{x} = \mathcal{R} \downarrow \check{m}$.*

Finite representation in terms of count vectors In earlier work [1], we have proved that a coherent and exchangeable set of desirable gambles on sequences can be represented by a coherent set of desirable gambles on count vectors. To move between both spaces of gambles, we were led to the linear map MuHy^N that maps a gamble f on \mathcal{X}^N to the gamble $\text{MuHy}^N(f) := \text{MuHy}^N(f|\cdot)$ on \mathcal{N}^N . Here, for every m in \mathcal{N}^N , $\text{MuHy}^N(f|m) := \sum_{y \in [m]} f(y)/|[m]|$ is the expectation of f under the *multivariate hyper-geometric distribution* [7, Section 39.2] associated with random sampling without replacement from an urn, whose composition is characterised by the count vector m . For the other direction, we use the linear map T^N that maps a gamble g on \mathcal{N}^N to the permutation invariant gamble $T^N(g) := g \circ T^N$ on \mathcal{X}^N assuming the value $g(m)$ on the invariant atom $[m]$.

Theorem 2 (Finite Representation). *A set of desirable gambles \mathcal{R} on \mathcal{X}^N is coherent and exchangeable iff there is some coherent set \mathcal{S} of desirable gambles on \mathcal{N}^N such that $\mathcal{R} = (\text{MuHy}^N)^{-1}(\mathcal{S})$, and in that case this \mathcal{S} is uniquely determined by $\mathcal{S} = \{g \in \mathcal{G}(\mathcal{N}^N) : T^N(g) \in \mathcal{R}\} = \text{MuHy}^N(\mathcal{R})$. We call \mathcal{S} the count representation of the exchangeable set \mathcal{R} .*

Multinomial processes Next, we turn to a number of important ideas related to multinomial processes. They are useful for comparisons with the existing literature [6, for example], and essential for our treatment of countable exchangeable sequences in Section 4.

Consider the \mathcal{X} -simplex $\Sigma_{\mathcal{X}} := \{\theta \in \mathbb{R}^{\mathcal{X}} : \theta \geq 0 \text{ and } \sum_{x \in \mathcal{X}} \theta_x = 1\}$, and, for $N \in \mathbb{N}_0$, the linear map CoMn^N from $\mathcal{G}(\mathcal{N}^N)$ to $\mathcal{G}(\Sigma_{\mathcal{X}})$ defined by $\text{CoMn}^N(g) = \text{CoMn}^N(g|\cdot)$, where for all θ in $\Sigma_{\mathcal{X}}$, $\text{CoMn}^N(g|\theta) := \sum_{m \in \mathcal{N}^N} g(m)B_m(\theta)$ is the expectation associated with the *count multinomial* distribution with parameters N and θ , and where B_m is the multivariate *Bernstein (basis) polynomial of degree N* given by $B_m(\theta) := \binom{N}{m} \prod_{x \in \mathcal{X}} \theta_x^{m_x} = |[m]| \prod_{x \in \mathcal{X}} \theta_x^{m_x}$. We also consider the related linear map Mn^N from $\mathcal{G}(\mathcal{X}^N)$ to $\mathcal{G}(\Sigma_{\mathcal{X}})$ defined by $\text{Mn}^N(f) = \text{Mn}^N(f|\cdot)$,

where for all θ in $\Sigma_{\mathcal{X}}$, $\text{Mn}^N(f|\theta) := \sum_{m \in \mathcal{N}^N} \text{MuHy}^N(f|m)B_m(\theta)$ is the expectation associated with the *multinomial* distribution with parameters N and θ . We then have that $\text{CoMn}^N = \text{Mn}^N \circ \text{T}^N$ and $\text{Mn}^N = \text{CoMn}^N \circ \text{MuHy}^N$.

The Bernstein basis polynomials B_m , $m \in \mathcal{N}^N$ form a basis for the linear space $\mathcal{V}_N(\Sigma_{\mathcal{X}})$ of all polynomials on $\Sigma_{\mathcal{X}}$ of degree up to N [9]. This means that for each polynomial p whose degree $\deg(p)$ does not exceed N , there is a unique gamble b_p^N on \mathcal{N}^N such that $p = \text{CoMn}^N(b_p^N)$. We denote by $\mathcal{V}(\Sigma_{\mathcal{X}})$ the linear space of all polynomials on $\Sigma_{\mathcal{X}}$.

Finite representation in terms of polynomials We see that the range of the linear maps CoMn^N and Mn^N is the linear space $\mathcal{V}_N(\Sigma_{\mathcal{X}})$. Moreover, since for every polynomial p of degree up to N , i.e., for every p in $\mathcal{V}_N(\Sigma_{\mathcal{X}})$, there is a *unique* count gamble $b_p^N \in \mathcal{G}(\mathcal{N}^N)$ such that $p = \text{CoMn}^N(b_p^N)$, CoMn^N is a *linear isomorphism* between the linear spaces $\mathcal{G}(\mathcal{N}^N)$ and $\mathcal{V}_N(\Sigma_{\mathcal{X}})$.

In summary, everything that can be expressed using the language of gambles on \mathcal{N}^N , can also be expressed using the language of polynomial gambles on $\Sigma_{\mathcal{X}}$ of degree up to N , and *vice versa*. The map CoMn^N and its inverse are the tools that take care of the translation between the two languages. This is essentially what is behind the representation theorem for countable exchangeable sequences that we will turn to in Section 4. In order to lay the proper foundations for this work, we now present a version of the finite representation theorem in terms of polynomial gambles of degree N on $\Sigma_{\mathcal{X}}$, rather than count gambles on \mathcal{N}^N .

We first introduce a concept of coherence for Bernstein polynomials:

Definition 4 (Bernstein coherence). *We call a set \mathcal{H} of polynomials in $\mathcal{V}_N(\Sigma_{\mathcal{X}})$ Bernstein coherent at degree N if it satisfies the following properties: for all p, p_1 , and p_2 in $\mathcal{V}_N(\Sigma_{\mathcal{X}})$ and all real $\lambda > 0$,*

- B_N1. *if $p = 0$ then $p \notin \mathcal{H}$;*
- B_N2. *if p is such that $b_p^N > 0$ then $p \in \mathcal{H}$;*
- B_N3. *if $p \in \mathcal{H}$ then $\lambda p \in \mathcal{H}$;*
- B_N4. *if $p_1, p_2 \in \mathcal{H}$ then $p_1 + p_2 \in \mathcal{H}$.*

Bernstein coherence at degree N is very closely related to coherence, the only difference being that we do not consider whether a polynomial p is positive, but whether its Bernstein expansion b_p^N is. Interpretation-wise, this means that models in terms of sequences or count vectors are authoritative over those in terms of frequency vectors. Bernstein coherence at degree N is a special case of the general concept of coherence relative to $(\mathcal{K}, \mathcal{C})$, discussed in Section 2, where $\mathcal{K} := \mathcal{V}_N(\Sigma_{\mathcal{X}})$ and $\mathcal{C} := \{p \in \mathcal{V}_N(\Sigma_{\mathcal{X}}) : b_p^N \geq 0\}$ is the convex cone of all polynomials of degree at most N with a non-negative expansion b_p^N in the Bernstein basis of degree N .

Theorem 3 (Finite Representation). *A set of desirable gambles \mathcal{R} on \mathcal{X}^N , with count representation $\mathcal{S} := \text{MuHy}^N(\mathcal{R})$, is coherent and exchangeable iff there is some subset \mathcal{H} of $\mathcal{V}_N(\Sigma_{\mathcal{X}})$, Bernstein coherent at degree N , such that $\mathcal{R} = (\text{Mn}^N)^{-1}(\mathcal{H})$ or equivalently $\mathcal{S} = (\text{CoMn}^N)^{-1}(\mathcal{H})$, and in that case this \mathcal{H} is*

uniquely determined by $\mathcal{H} = \text{Mn}^N(\mathcal{R}) = \text{CoMn}^N(\mathcal{S})$. We call \mathcal{H} the frequency representation of the coherent and exchangeable set of desirable gambles \mathcal{R} .

4 Countable exchangeable sequences

With the experience gained in investigating finite exchangeable sequences, we are now ready to address reasoning about countably infinite exchangeable sequences. The first step is to use finite frequency representation to find a Representation Theorem for infinite exchangeable sequences. We can then show what updating and natural extension look like in terms of this frequency representation.

Infinite representation We consider a countable sequence X_1, \dots, X_N, \dots of random variables assuming values in the same finite set \mathcal{X} . We call this sequence *exchangeable* if each of its finite subsequences is, or equivalently, if for all n in \mathbb{N}_0 , the random variables X_1, \dots, X_n are exchangeable.

How can we model this? First of all, this means that for each n in \mathbb{N}_0 , there is a coherent and exchangeable set of desirable gambles \mathcal{R}^n on \mathcal{X}^n . Equivalently, we have a coherent set of desirable gambles (count representation) $\mathcal{S}^n := \text{MuHy}^n(\mathcal{R}^n)$ on \mathcal{N}^n , or a set (frequency representation) $\mathcal{H}^n := \text{Mn}^n(\mathcal{R}^n) = \text{CoMn}^n(\mathcal{S}^n)$ of polynomials in $\mathcal{V}_n(\Sigma_{\mathcal{X}})$, Bernstein coherent at degree n .

In addition, there is a time-consistency constraint. Consider the following linear *projection operators* $\text{proj}_{n_2}^{n_1}: \mathcal{X}^{n_2} \rightarrow \mathcal{X}^{n_1}$ defined by $\text{proj}_{n_2}^{n_1}(x_1, \dots, x_{n_2}) := (x_1, \dots, x_{n_1})$, where $n_1 \leq n_2$. With each such operator there corresponds a linear map $\text{ext}_{n_1}^{n_2}$ between the linear spaces $\mathcal{G}(\mathcal{X}^{n_1})$ and $\mathcal{G}(\mathcal{X}^{n_2})$, defined as follows: $\text{ext}_{n_1}^{n_2}(f) = f \circ \text{proj}_{n_2}^{n_1}$. In other words, $\text{ext}_{n_1}^{n_2}(f)$ is the *cylindrical extension* of the gamble f on \mathcal{X}^{n_1} to a gamble on \mathcal{X}^{n_2} .

Time-consistency now means that if we consider a gamble on \mathcal{X}^{n_2} that really only depends on the first n_1 variables, it should not matter, as far as its desirability is concerned, whether we consider it to be a gamble on \mathcal{X}^{n_1} or a gamble on \mathcal{X}^{n_2} . More formally, we require that $(\forall n_1 \leq n_2) \text{ext}_{n_1}^{n_2}(\mathcal{R}^{n_1}) = \mathcal{R}^{n_2} \cap \text{ext}_{n_1}^{n_2}(\mathcal{G}(\mathcal{X}^{n_1}))$.

How can we translate this constraint in terms of the count representations \mathcal{S}^n or the frequency representations \mathcal{H}^n ? If we introduce the linear *extension map* $\text{enl}_{n_1}^{n_2}$ from the linear space $\mathcal{G}(\mathcal{N}^{n_1})$ to the linear space $\mathcal{G}(\mathcal{N}^{n_2})$ by $\text{enl}_{n_1}^{n_2}(g) := \sum_{m \in \mathcal{N}^{n_1}} g(m) |[\cdot - m]| / |[m]| / |[\cdot]|$, then $\text{MuHy}^{n_2} \circ \text{ext}_{n_1}^{n_2} = \text{enl}_{n_1}^{n_2} \circ \text{MuHy}^{n_1}$, and time-consistency is equivalent to $(\forall n_1 \leq n_2) \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1}) = \mathcal{S}^{n_2} \cap \text{enl}_{n_1}^{n_2}(\mathcal{G}(\mathcal{N}^{n_1}))$, which is in turn equivalent to $(\forall n_1 \leq n_2) \mathcal{H}^{n_1} = \mathcal{H}^{n_2} \cap \mathcal{V}_{n_1}(\Sigma_{\mathcal{X}})$. The time-consistency condition is most elegantly expressed using frequency representations.

We call the family \mathcal{R}^n , $n \in \mathbb{N}_0$ *time-consistent, coherent and exchangeable* when each member \mathcal{R}^n is coherent and exchangeable, and when the family \mathcal{R}^n , $n \in \mathbb{N}_0$ satisfies the time-consistency requirement.

We can generalise the concept of Bernstein coherence given in Definition 4 to sets of polynomials of arbitrary degree:

Definition 5 (Bernstein coherence). We call a set \mathcal{H} of polynomials in $\mathcal{V}(\Sigma_{\mathcal{X}})$ Bernstein coherent if it satisfies the following properties: for all p , p_1 , and p_2 in $\mathcal{V}(\Sigma_{\mathcal{X}})$ and all real $\lambda > 0$,

- B1. if $p = 0$ then $p \notin \mathcal{H}$;
- B2. if p is such that $b_p^n > 0$ for some $n \geq \deg(p)$, then $p \in \mathcal{H}$;
- B3. if $p \in \mathcal{H}$ then $\lambda p \in \mathcal{H}$;
- B4. if $p_1, p_2 \in \mathcal{H}$ then $p_1 + p_2 \in \mathcal{H}$.

We can replace B1 by the following requirement, equivalent to it under B2–B4:

B5. If p is such that $b_p^n \leq 0$ for some $n \geq \deg(p)$, then $p \notin \mathcal{H}$.

This type of Bernstein coherence is again very closely related to coherence, the only difference being that not all positive polynomials, but rather all polynomials with some positive Bernstein expansion are required to belong to a Bernstein coherent set. Bernstein coherence is a special case of the general concept of coherence relative to $(\mathcal{K}, \mathcal{C})$, discussed in Section 2, where $\mathcal{K} := \mathcal{V}(\Sigma_{\mathcal{X}})$ and $\mathcal{C} := \mathcal{V}^+(\Sigma_{\mathcal{X}}) := \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : (\exists n \geq 0) b_p^n \geq 0\}$ is the convex cone of all polynomials with some non-negative Bernstein expansion. We also denote the set $\mathbb{D}_{(\mathcal{V}(\Sigma_{\mathcal{X}}), \mathcal{V}^+(\Sigma_{\mathcal{X}}))}(\Sigma_{\mathcal{X}})$ of all Bernstein coherent subsets of $\mathcal{V}(\Sigma_{\mathcal{X}})$ by $\mathbb{D}_{\text{Be}}(\Sigma_{\mathcal{X}})$.

We are now ready to formulate our Infinite Representation Theorem, which is a significant generalisation of de Finetti's representation result for countable sequences [3]. A similar result can be proved for coherent lower previsions [2].

Theorem 4 (Infinite Representation). A family \mathcal{R}^n , $n \in \mathbb{N}_0$ of sets of desirable gambles on \mathcal{X}^n , with associated count representations $\mathcal{S}^n := \text{MuHy}^n(\mathcal{R}^n)$ and frequency representations $\mathcal{H}^n := \text{Mn}^n(\mathcal{R}^n) = \text{CoMn}^n(\mathcal{S}^n)$, is time-consistent, coherent and exchangeable iff there is some Bernstein coherent set \mathcal{H} of polynomials in $\mathcal{V}(\Sigma_{\mathcal{X}})$ such that, for all n in \mathbb{N}_0 , both $\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{H})$ and $\mathcal{R}^n = (\text{Mn}^n)^{-1}(\mathcal{H})$, and in that case this \mathcal{H} is uniquely given by $\mathcal{H} = \bigcup_{n \in \mathbb{N}_0} \mathcal{H}^n$. We call \mathcal{H} the frequency representation of the coherent, exchangeable and time-consistent family of sets of desirable gambles \mathcal{R}^n , $n \in \mathbb{N}_0$.

Updating and infinite representation Suppose we have a coherent, exchangeable and time-consistent family of sets of desirable gambles \mathcal{R}^n , $n \in \mathbb{N}_0$, with associated count representations $\mathcal{S}^n := \text{MuHy}^n(\mathcal{R}^n)$ and frequency representation $\mathcal{H} := \bigcup_{n \in \mathbb{N}} \mathcal{H}^n$ with $\mathcal{H}^n := \text{Mn}^n(\mathcal{R}^n)$. Now suppose we observe the values \check{x} of the first \check{n} variables. It turns out that updating becomes especially easy in terms of the frequency representation.

Theorem 5. Consider a coherent, exchangeable and time-consistent family of sets of desirable gambles \mathcal{R}^n , $n \in \mathbb{N}_0$, with associated frequency representation \mathcal{H} . After updating with a sample with count vector $\check{m} \in \mathcal{N}_{\mathcal{X}}^{\check{n}}$, the family $\mathcal{R}^{\hat{n}} \upharpoonright \check{m}$, $\hat{n} \in \mathbb{N}_0$ is still coherent, exchangeable and time-consistent, and has frequency representation $\mathcal{H} \upharpoonright \check{m} := \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : B_{\check{m}} p \in \mathcal{H}\}$.

Independence: iid sequences We can use Theorem 5 to find an intriguing characterisation of a sequence of *independent identically distributed* (iid) random variables X_1, \dots, X_N, \dots assuming values in a finite set \mathcal{X} . This is an

exchangeable sequence where learning the value of any finite number of variables does not change our subject's beliefs about the remaining, unobserved ones. This is the case iff the frequency representation \mathcal{H} of the sequence satisfies $(\forall \check{n} \in \mathbb{N}_0)(\forall \check{m} \in \mathcal{N}_{\mathcal{X}}^{\check{n}})\mathcal{H} \upharpoonright \check{m} = \mathcal{H}$. This is equivalent to $(\forall \check{n} \in \mathbb{N}_0)(\forall \check{m} \in \mathcal{N}_{\mathcal{X}}^{\check{n}})(\forall p \in \mathcal{V}(\Sigma_{\mathcal{X}}))(p \in \mathcal{H} \Leftrightarrow B_{\check{m}}p \in \mathcal{H})$, and also to $\mathcal{H} = \mathcal{V}^+(\Sigma_{\mathcal{X}})\mathcal{H}$, which is shorthand for $(\forall p \in \mathcal{V}(\Sigma_{\mathcal{X}}))(\forall p^+ \in \mathcal{V}^+(\Sigma_{\mathcal{X}}))(p \in \mathcal{H} \Leftrightarrow p^+p \in \mathcal{H})$.

Bernstein natural extension The intersection of an arbitrary non-empty family of Bernstein coherent sets of polynomials is still Bernstein coherent. This is the idea behind the following theorem, which is a special instance of Theorem 1 with $\mathcal{K} := \mathcal{V}(\Sigma_{\mathcal{X}})$ and $\mathcal{C} := \mathcal{V}^+(\Sigma_{\mathcal{X}})$.

We denote by $\mathcal{V}_0^+(\Sigma_{\mathcal{X}})$ the set of all polynomials on $\Sigma_{\mathcal{X}}$ with some positive Bernstein expansion: $\mathcal{V}_0^+(\Sigma_{\mathcal{X}}) = \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : (\exists n \geq \deg(p))b_p^n > 0\}$ and by $\mathcal{V}^-(\Sigma_{\mathcal{X}}) = \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : (\exists n \geq \deg(p))b_p^n \leq 0\}$ the set of all polynomials on $\Sigma_{\mathcal{X}}$ with some non-positive Bernstein expansion. Moreover, we say that a set \mathcal{A} of polynomials *avoids Bernstein non-positivity* if no polynomial in its positive hull $\text{posi}(\mathcal{A})$ has any non-positive Bernstein expansion, i.e. $\text{posi}(\mathcal{A}) \cap \mathcal{V}^-(\Sigma_{\mathcal{X}}) = \emptyset$; clearly, this is the case iff \mathcal{A} avoids non-positivity relative to $(\mathcal{V}(\Sigma_{\mathcal{X}}), \mathcal{V}^+(\Sigma_{\mathcal{X}}))$. We also call the $(\mathcal{V}(\Sigma_{\mathcal{X}}), \mathcal{V}^+(\Sigma_{\mathcal{X}}))$ -natural extension $\mathcal{E}_{(\mathcal{V}(\Sigma_{\mathcal{X}}), \mathcal{V}^+(\Sigma_{\mathcal{X}}))}(\mathcal{A})$ of \mathcal{A} its Bernstein natural extension, and denote it by $\mathcal{E}_{\text{Be}}(\mathcal{A})$.

Theorem 6 (Bernstein natural extension). *Consider $\mathcal{A} \subseteq \mathcal{V}(\Sigma_{\mathcal{X}})$ and its Bernstein natural extension $\mathcal{E}_{\text{Be}}(\mathcal{A}) := \bigcap \{\mathcal{H} \in \mathbb{D}_{\text{Be}}(\Sigma_{\mathcal{X}}) : \mathcal{A} \subseteq \mathcal{H}\}$. The following statements are then equivalent:*

- (i) \mathcal{A} avoids Bernstein non-positivity;
 - (ii) \mathcal{A} is included in some Bernstein coherent set of polynomials;
 - (iii) $\mathcal{E}_{\text{Be}}(\mathcal{A}) \neq \mathcal{V}(\Sigma_{\mathcal{X}})$;
 - (iv) $\mathcal{E}_{\text{Be}}(\mathcal{A})$ is a Bernstein coherent set of polynomials;
 - (v) $\mathcal{E}_{\text{Be}}(\mathcal{A})$ is the smallest Bernstein coherent set of polynomials including \mathcal{A} .
- When any these equivalent statements holds, then $\mathcal{E}_{\text{Be}}(\mathcal{A}) = \text{posi}(\mathcal{V}_0^+(\Sigma_{\mathcal{X}}) \cup \mathcal{A})$.

Exchangeable natural extension for infinite sequences To finish this discussion of exchangeability for infinite sequences of random variables, we take up the issue of deductive inference, and introduce exchangeable natural extension.

Suppose we have an assessment consisting of a set \mathcal{A}^n of desirable gambles on \mathcal{X}^n for each n in \mathbb{N}_0 . We are looking for the (element-wise) smallest coherent, exchangeable and time-consistent family \mathcal{R}^n , $n \in \mathbb{N}_0$ that includes this assessment in the sense that $\mathcal{A}^n \subseteq \mathcal{R}^n$ for all n in \mathbb{N}_0 , or equivalently $\bigcup_{n \in \mathbb{N}_0} \text{Mn}^n(\mathcal{A}^n) \subseteq \bigcup_{n \in \mathbb{N}_0} \text{Mn}^n(\mathcal{R}^n) =: \mathcal{H}$, a condition formulated in terms of the frequency representation \mathcal{H} of the family \mathcal{R}^n , $n \in \mathbb{N}_0$.

Theorem 7. *Consider an assessment consisting of the sets of desirable gambles \mathcal{A}^n on \mathcal{X}^n for each n in \mathbb{N}_0 , and the corresponding set of 'desirable' polynomials $\mathcal{A} := \bigcup_{n \in \mathbb{N}_0} \text{Mn}^n(\mathcal{A}^n)$. There is a coherent, exchangeable and time-consistent*

family \mathcal{R}^n , $n \in \mathbb{N}_0$ that includes this assessment iff \mathcal{A} avoids Bernstein non-positivity. In that case $\mathcal{E}_{\text{Be}}(\mathcal{A})$ is the frequency representation of the (element-wise) smallest coherent, exchangeable and time-consistent family that includes this assessment.

5 Conclusions

Modelling an infinite exchangeability assessment using sets of desirable gambles is not only possible, but also quite elegant. Our Infinite Representation Theorem reduces reasoning about infinite exchangeable sequences to reasoning about (polynomials of) frequency vectors. This automatically guarantees that, next to the exchangeability of finite subsequences, time-consistency of these subsequences is satisfied. The representation for the natural extension and for updated models can be derived directly from the representation of the original model.

Acknowledgements Erik Quaeghebeur was supported by a Fellowship of the Belgian American Educational Foundation. We wish to thank the reviewer.

References

1. De Cooman, G., Quaeghebeur, E.: Exchangeability for sets of desirable gambles. In: Augustin, T., Coolen, F.P.A., Moral, S., Troffaes, M.C.M. (eds.) ISIPTA '09: proceedings of the Sixth International Symposium on Imprecise Probabilities: Theories and Applications. pp. 159–168. SIPTA, Durham, United Kingdom (2009), <http://hdl.handle.net/1854/LU-718913>
2. De Cooman, G., Quaeghebeur, E., Miranda, E.: Exchangeable lower previsions. *Bernoulli* 15(3), 721–735 (2009), <http://hdl.handle.net/1854/LU-498518>, an expanded version can be found at <http://arxiv.org/abs/0801.1265>
3. de Finetti, B.: La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré* 7, 1–68 (1937), English translation in [8]
4. de Finetti, B.: *Teoria delle Probabilità*. Einaudi, Turin (1970)
5. de Finetti, B.: *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons, Chichester (1974–1975), English translation of [4], two volumes
6. Heath, D.C., Sudderth, W.D.: De Finetti's theorem on exchangeable variables. *The American Statistician* 30, 188–189 (1976)
7. Johnson, N.L., Kotz, S., Balakrishnan, N.: *Discrete Multivariate Distributions*. Wiley Series in Probability and Statistics, John Wiley and Sons, New York (1997)
8. Kyburg Jr., H.E., Smokler, H.E. (eds.): *Studies in Subjective Probability*. Wiley, New York (1964), second edition (with new material) 1980
9. Prautzsch, H., Boehm, W., Paluszny, M.: *Bézier and B-Spline Techniques*. Springer, Berlin (2002)
10. Schechter, E.: *Handbook of Analysis and Its Foundations*. Academic Press, San Diego, CA (1997)
11. Walley, P.: *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London (1991)
12. Walley, P.: Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning* 24, 125–148 (2000)