

# EXCHANGEABILITY AND SETS OF DESIRABLE GAMBLES

GERT DE COOMAN AND ERIK QUAEGHEBEUR

**ABSTRACT.** Sets of desirable gambles constitute a quite general type of uncertainty model with an interesting geometrical interpretation. We give a general discussion of such models and their rationality criteria. We study exchangeability assessments for them, and prove counterparts of de Finetti’s finite and infinite representation theorems. We show that the finite representation in terms of count vectors has a very nice geometrical interpretation, and that the 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 and the extension of exchangeable sequences.

## 1. INTRODUCTION

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

Desirability, or the theory of (coherent) sets of desirable gambles, has been introduced with all main ideas present—so far as our search has unearthed—by Williams [29, 30, 31]. Building on de Finetti’s betting framework [11], he considered the ‘acceptability’ of *one-sided* bets instead of *two-sided* bets. This relaxation leads one to work with cones of bets instead of with linear subspaces of them. The germ of the theory was, however, already present in Smith’s work [25, p. 15], who used a (generally) open cone of ‘exchange vectors’ when talking about currency exchange. Both authors influenced Walley [27, Section 3.7 and App. F], who describes three variants (almost, really, and strictly desirable gambles) and emphasises the conceptual ease with which updated (or posterior) models can be obtained in this framework [28]. Moral [20, 21] then took the next step and applied the theory to study epistemic irrelevance, a structural assessment. He also pointed out how conceptually easy extension, marginalisation, and conditioning are in this framework. de Cooman and Miranda [4] made a general study of transformational symmetry assessments for desirable gambles. Recently, Couso and Moral [1] discussed the relationship with credal sets, computer representation, and maximal sets of desirable gambles.

The structural assessment we are interested in here, is exchangeability. Conceptually, it says that the order of the samples in a sequence of them is irrelevant for inference purposes. The first detailed study of this concept was made by de Finetti [9], using the terminology of ‘equivalent’ events. He proved the now famous Representation Theorem, which is often interpreted as stating that a sequence of random variables is exchangeable if it is conditionally independent and identically distributed. Other important work—all using probabilities or previsions—was done by, amongst many others, Hewitt and Savage [14], Heath and Sudderth [13], and Diaconis and Freedman [12].<sup>1</sup> Exchangeability in the context of imprecise-probability theory—using lower previsions—was studied by Walley [27, Section 9.5] and more in-depth by de Cooman et al. [4, 7, 8]. The first embryonic study of exchangeability using desirability was recently performed by Quaeghebeur [23, Section 3.1.1].

---

*Key words and phrases.* desirability, real desirability, weak desirability, sets of desirable gambles, coherence, exchangeability, representation, natural extension, updating, extending an exchangeable sequence.

<sup>1</sup>See, e.g., Kallenberg [16, 17] for a measure-theoretic discussion of exchangeability.

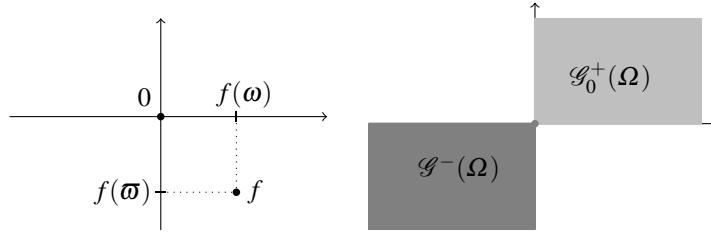
Here we present the results of a more matured study of exchangeability using sets of desirable gambles. First, in Section 2, we give a general discussion of desirability, coherence—the criteria that define which sets of desirable gambles are rational uncertainty models—and the smallest and maximal such sets compatible with some assessment. Next, in Section 3, we introduce the special case of this general theory that will form the basis of our analysis of exchangeability using the theory of desirable gambles. Then, in Section 4, we give a desirability-based analysis of finite exchangeable sequences, presenting a Representation Theorem—both in terms of count and frequency vectors—and treating the issues of natural extension and updating under exchangeability. Building on these results, we extend our scope to countable exchangeable sequences in Section 5, where we present a second Representation Theorem—in terms of frequency vectors—and again also treat updating and natural extension. Finally, in Section 6, we see if and how finite exchangeable sequences can be extended to longer, finite or even infinite exchangeable sequences.

Proofs of this paper’s results are included in Appendix A. Appendix B collects a few relevant facts about multivariate Bernstein basis polynomials.

## 2. A GENERAL DISCUSSION OF DESIRABILITY AND COHERENCE

Consider a non-empty set  $\Omega$  describing the possible and mutually exclusive outcomes of some experiment. We also consider a subject who is uncertain about the outcome of the experiment.

**2.1. Sets of desirable gambles.** A *gamble*  $f$  is a bounded real-valued map on  $\Omega$ , and it is interpreted as an uncertain reward. When the actual outcome of the experiment is  $\omega$ , then the corresponding (possibly negative) reward is  $f(\omega)$ , expressed in units of some pre-determined linear utility. This is illustrated for  $\Omega = \{\omega, \varpi\}$ .  $\mathcal{G}(\Omega)$  denotes the set of all gambles on  $\Omega$ ,  $\mathcal{G}_0^+(\Omega)$  the non-negative non-zero ones, and  $\mathcal{G}^-(\Omega)$  the non-positive ones.



We also use the following notational conventions throughout: subscripting a set with zero corresponds to *removing* zero (or the zero gamble) from the set, if present. For example  $\mathbb{R}^+$  ( $\mathbb{R}_0^+$ ) is the set of non-negative (positive) real numbers including (excluding) zero. Furthermore,  $f \geq g$  iff  $f(\omega) \geq g(\omega)$  for all  $\omega$  in  $\Omega$ ;  $f > g$  iff  $f \geq g$  and  $f \neq g$ .

We say that a non-zero gamble  $f$  is *desirable* to a subject if he accepts to engage in the following transaction, where: (i) the actual outcome  $\omega$  of the experiment is determined, and (ii) he receives the reward  $f(\omega)$ , i.e., his capital is changed by  $f(\omega)$ . The zero gamble is not considered to be desirable.<sup>2</sup>

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

<sup>2</sup>The nomenclature in the literature regarding desirability is somewhat confusing, and we have tried to resolve some of the ambiguity here. Our notion of desirability coincides with Walley’s later [28] notion of desirability, initially (and quite recently [11]) also used by Moral [20]. Walley in his book [27, App. F] and Moral in a later paper [20] use another notion of desirability. The difference between the two approaches resides in whether the zero gamble is assumed to be desirable or not. We prefer to use the non-zero version here, because it is better behaved in conjunction with our notion of weak desirability in Definition 2.

**2.2. 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.

We begin with a few preliminary definitions involving ordered linear spaces.

The set  $\mathcal{G}(\Omega)$  of all gambles on  $\Omega$  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*  $\text{posi}$  generates the set of *strictly* positive linear combinations of elements of its argument set: for any subset  $\mathcal{A}$  of  $\mathcal{G}(\Omega)$ ,

$$\text{posi}(\mathcal{A}) := \left\{ \sum_{k=1}^n \lambda_k f_k : f_k \in \mathcal{A}, \lambda_k \in \mathbb{R}_0^+, n \in \mathbb{N}_0 \right\}. \quad (1)$$

A subset  $\mathcal{C}$  of  $\mathcal{G}(\Omega)$  is a *convex cone* if it is closed under (strictly) positive linear combinations, or in other words, if  $\text{posi}(\mathcal{C}) = \mathcal{C}$ .

Consider a linear subspace  $\mathcal{H}$  of the linear space  $\mathcal{G}(\Omega)$ . With any convex cone  $\mathcal{C} \subset \mathcal{H}$  such that  $0 \in \mathcal{C}$  we can always associate a *vector ordering*  $\succeq$  on  $\mathcal{H}$ , defined as follows:<sup>3</sup>

$$f \succeq g \Leftrightarrow f - g \in \mathcal{C} \Leftrightarrow f - g \succeq 0. \quad (2)$$

The partial ordering  $\succeq$  turns  $\mathcal{H}$  into an ordered linear space [24, Section 11.44]. We also write  $f \succ g$  if  $f - g \in \mathcal{C}_0$ , or in other words, if  $f \succeq g$  and  $f \neq g$ . As usual,  $f \preceq g$  means  $g \succeq f$  and similarly,  $f \prec g$  means  $g \succ f$ . Finally, we let

$$\mathcal{H}_{\preceq 0} := \{f \in \mathcal{H} : f \preceq 0\} = -\mathcal{C} \quad \text{and} \quad \mathcal{H}_{\succ 0} := \{f \in \mathcal{H} : f \succ 0\} = \mathcal{C}_0. \quad (3)$$

**Definition 1** (Avoiding non-positivity and coherence). *Let  $\mathcal{H}$  be a linear subspace of  $\mathcal{G}(\Omega)$  and let  $\mathcal{C} \subset \mathcal{H}$  be a convex cone containing the zero gamble 0. We say that a set of desirable gambles  $\mathcal{R} \subseteq \mathcal{H}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$  if  $f \not\preceq 0$  for all gambles  $f$  in  $\text{posi}(\mathcal{R})$ , or in other words if  $\mathcal{H}_{\preceq 0} \cap \text{posi}(\mathcal{R}) = \emptyset$ .*

*We say that a set of desirable gambles  $\mathcal{R} \subseteq \mathcal{H}$  is coherent relative to  $(\mathcal{H}, \mathcal{C})$  if it satisfies the following requirements, for all gambles  $f, f_1$ , and  $f_2$  in  $\mathcal{H}$  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  $\mathcal{H}_{\succ 0} \subseteq \mathcal{R}$ ;*
- 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  $\mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega)$  the set of sets of desirable gambles that are coherent relative to  $(\mathcal{H}, \mathcal{C})$ .*

Requirements D3 and D4 make  $\mathcal{R}$  a *cone*:  $\text{posi}(\mathcal{R}) = \mathcal{R}$ . Due to D2, it includes  $\mathcal{H}_{\succ 0}$ ; due to D1, D2 and D4, it excludes  $\mathcal{H}_{\preceq 0}$ :

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

The non-triviality requirement  $\mathcal{C} \neq \mathcal{H}$  makes sure that  $\mathcal{H}$  is *never* coherent relative to  $(\mathcal{H}, \mathcal{C})$ . On the other hand,  $\mathcal{H}_{\succ 0}$  is *always* coherent relative to  $(\mathcal{H}, \mathcal{C})$ , and it is the smallest such subset of  $\mathcal{H}$ .

**2.3. Natural extension.** If we consider an arbitrary non-empty family of sets of desirable gambles  $\mathcal{R}_i, i \in I$  that are coherent relative to  $(\mathcal{H}, \mathcal{C})$ , then their intersection  $\bigcap_{i \in I} \mathcal{R}_i$  is still coherent relative to  $(\mathcal{H}, \mathcal{C})$ . This is the idea behind the following result. If a subject gives us an *assessment*, a set  $\mathcal{A} \subseteq \mathcal{H}$  of gambles on  $\Omega$  that he finds desirable, then it tells us exactly when this assessment can be extended to a coherent set, and how to construct the smallest such set.

<sup>3</sup>We require that  $\mathcal{C}$  should be *strictly* included in  $\mathcal{H}$  ( $\mathcal{C} \neq \mathcal{H}$ ) because otherwise the ordering  $\succeq$  would be trivial: we would have that  $f \succeq g$  for all  $f, g \in \mathcal{H}$ .

**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:<sup>4</sup>*

$$\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A}) := \bigcap \{ \mathcal{R} \in \mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega) : \mathcal{A} \subseteq \mathcal{R} \} \quad (4)$$

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 (and hence all) of these equivalent statements hold, then

$$\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A}) = \text{posi}(\mathcal{K}_{>0} \cup \mathcal{A}). \quad (5)$$

This shows that if we have an assessment  $\mathcal{A}$  with a finite description, we can represent its natural extension on a computer by storing a finite description of its extreme rays.

**2.4. Maximal coherent sets.** We see that the set  $\mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)$  of all sets of desirable gambles that are coherent relative to  $(\mathcal{K}, \mathcal{C})$  can be partially ordered by set inclusion  $\subseteq$ .<sup>5</sup> Following in the footsteps of Couso and Moral [1], let us now investigate the maximal elements of this poset in more detail.

An element  $\mathcal{R}$  of  $\mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)$  is called *maximal* if it is not strictly included in any other element of  $\mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)$ , or in other words, if adding any gamble  $f$  to  $\mathcal{R}$  makes sure we can no longer extend the result  $\mathcal{R} \cup \{f\}$  to a set that is still coherent relative to  $(\mathcal{K}, \mathcal{C})$ :

$$\mathbb{M}_{(\mathcal{K}, \mathcal{C})}(\Omega) = \{ \mathcal{R} \in \mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega) : (\forall \mathcal{R}' \in \mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)) (\mathcal{R} \subseteq \mathcal{R}' \Rightarrow \mathcal{R} = \mathcal{R}') \} \quad (6)$$

is the set of all maximal elements of  $\mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)$ .

The following proposition provides a characterisation of such maximal elements.

**Proposition 2.** *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. Let  $\mathcal{R} \in \mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)$ , then  $\mathcal{R}$  is a maximal coherent set relative to  $(\mathcal{K}, \mathcal{C})$  iff*

$$(\forall f \in \mathcal{K}_0)(f \notin \mathcal{R} \Rightarrow -f \in \mathcal{R}). \quad (7)$$

For the following important result(s), a constructive proof can be given in case  $\Omega$  is finite, based on the same ideas as in [1]. They guarantee that  $\mathbb{D}_{(\mathcal{K}, \mathcal{C})}(\Omega)$  has all the useful properties of a *strong belief structure* [3]. In Appendix A, we give a non-constructive proof (based on Zorn's Lemma) for  $\Omega$  that may also be infinite.

**Theorem 3.** *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 any subset  $\mathcal{A}$  of  $\mathcal{K}$ , then  $\mathcal{A}$  avoids non-positivity relative to  $(\mathcal{K}, \mathcal{C})$  iff there is some maximal  $\mathcal{R} \in \mathbb{M}_{(\mathcal{K}, \mathcal{C})}(\Omega)$  such that  $\mathcal{A} \subseteq \mathcal{R}$ .*

**Corollary 4.** *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 any subset  $\mathcal{A}$  of  $\mathcal{K}$ , then*

$$\mathcal{E}_{(\mathcal{K}, \mathcal{C})}(\mathcal{A}) = \bigcap \{ \mathcal{R} \in \mathbb{M}_{(\mathcal{K}, \mathcal{C})}(\Omega) : \mathcal{A} \subseteq \mathcal{R} \}. \quad (8)$$

<sup>4</sup>As usual, in this expression, we let  $\bigcap \emptyset = \mathcal{K}$ .

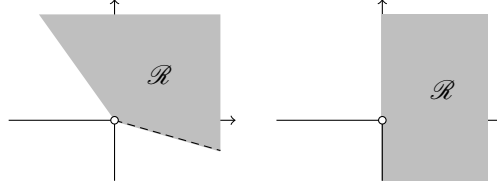
<sup>5</sup>This structure is a complete meet-semilattice, where intersection plays the role of infimum.

## 3. AN IMPORTANT SPECIAL CASE

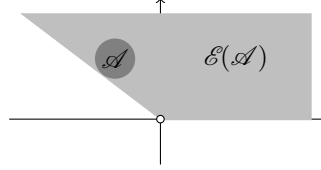
We now turn to the important special case, commonly considered in the literature [20, 28], where  $\mathcal{H} = \mathcal{G}(\Omega)$  and  $\mathcal{C} = \mathcal{G}_0^+(\Omega)$  is the cone of all non-negative gambles, for which the associated partial order  $\succeq$  is the point-wise ordering  $\geq$ .

If  $\mathcal{R}$  avoids non-positivity relative to  $(\mathcal{G}(\Omega), \mathcal{G}_0^+(\Omega))$ , we simply say that  $\mathcal{R}$  *avoids non-positivity*:  $\mathcal{G}^-(\Omega) \cap \text{posi}(\mathcal{R}) = \emptyset$ .<sup>6</sup>

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*. In particular, D2 is now an ‘accepting partial gain’ condition:  $\mathcal{G}_0^+(\Omega) \subseteq \mathcal{R}$ . D5 is an ‘avoiding partial loss’ condition, together with the convention that the zero gamble is never considered desirable. We give two illustrations, the first is a general one and the second models certainty about  $\omega$  happening. The dashed line indicates a non-included border.



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  $\mathcal{A}$ . In that case we can visualise Eq. (5), i.e., natural extension as a positive hull operation, with a small illustration:



More generally, consider a linear subspace  $\mathcal{H}$  of  $\mathcal{G}(\Omega)$ , and let  $\mathcal{C} = \{f \in \mathcal{H} : f \geq 0\} = \mathcal{G}_0^+(\Omega) \cap \mathcal{H}$  be the cone of all non-negative gambles in that subspace. If a set  $\mathcal{R} \subseteq \mathcal{H}$  is coherent relative to  $(\mathcal{H}, \mathcal{G}_0^+(\Omega) \cap \mathcal{H})$ , we also say that it is *coherent relative to  $\mathcal{H}$* .

In Secs. 4.9 and 5, we shall come across other useful types of coherence, relative to more involved structures  $\mathcal{H}$  and  $\mathcal{C}$ .

**3.1. Weakly desirable gambles, previsions & marginally desirable gambles.** We now define *weak desirability*: a useful modification of Walley’s [27, Section 3.7] notion of *almost-desirability*. Our conditions for a gamble  $f$  to be weakly desirable are more stringent than Walley’s for almost-desirability: he only requires that adding any constant strictly positive amount of utility to  $f$  should make the resulting gamble desirable. We require that adding anything desirable (be it constant or not) to  $f$  should make the resulting gamble desirable. Weak desirability is better behaved under updating: we shall see in Proposition 15 that it makes sure that the exchangeability of a set of desirable gambles, whose definition hinges on the notion of weak desirability, is preserved under updating after observing a sample. This is not necessarily true if weak desirability is replaced by almost-desirability in the definition of exchangeability, as was for instance done in our earlier work [4].

**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  $f' \in \mathcal{R}$ . We denote the set of weakly desirable gambles by  $\mathcal{D}_{\mathcal{R}}$ :*

$$\mathcal{D}_{\mathcal{R}} = \{f \in \mathcal{G}(\Omega) : f + \mathcal{R} \subseteq \mathcal{R}\}. \quad (9)$$

<sup>6</sup>A related, but weaker condition, is that  $\mathcal{R}$  *avoids partial loss*, meaning that  $f \not\prec 0$  for all gambles  $f$  in  $\text{posi}(\mathcal{R})$ . We need the stronger condition because we have excluded the zero gamble from being desirable.

In particular, every desirable gamble is also weakly desirable, so  $\mathcal{R} \subseteq \mathcal{D}_{\mathcal{R}}$ .

**Proposition 5.** *Let  $\mathcal{R}$  be a coherent set of desirable gambles, and let  $\mathcal{D}_{\mathcal{R}}$  be the associated set of weakly desirable gambles. Then  $\mathcal{D}_{\mathcal{R}}$  has the following properties, for all gambles  $f$ ,  $f_1$ , and  $f_2$  in  $\mathcal{G}(\Omega)$  and all real  $\lambda \geq 0$ :*

- WD1. *if  $f < 0$  then  $f \notin \mathcal{D}_{\mathcal{R}}$ , or equivalently  $\mathcal{G}_0^-(\Omega) \cap \mathcal{D}_{\mathcal{R}} = \emptyset$  [avoiding partial loss];<sup>7</sup>*
- WD2. *if  $f \geq 0$  then  $f \in \mathcal{D}_{\mathcal{R}}$ , or equivalently  $\mathcal{G}_0^+(\Omega) \subseteq \mathcal{D}_{\mathcal{R}}$  [accepting partial gain];*
- WD3. *if  $f \in \mathcal{D}_{\mathcal{R}}$  then  $\lambda f \in \mathcal{D}_{\mathcal{R}}$  [scaling];*
- WD4. *if  $f_1, f_2 \in \mathcal{D}_{\mathcal{R}}$  then  $f_1 + f_2 \in \mathcal{D}_{\mathcal{R}}$  [combination].*

Like  $\mathcal{R}$ ,  $\mathcal{D}_{\mathcal{R}}$  is a convex cone.

With a set of gambles  $\mathcal{A}$ , we associate a lower prevision  $\underline{P}_{\mathcal{A}}$  and an upper prevision  $\overline{P}_{\mathcal{A}}$  by letting

$$\underline{P}_{\mathcal{A}}(f) := \sup \{ \mu \in \mathbb{R} : f - \mu \in \mathcal{A} \} \quad (10)$$

$$\overline{P}_{\mathcal{A}}(f) := \inf \{ \mu \in \mathbb{R} : \mu - f \in \mathcal{A} \} \quad (11)$$

for all gambles  $f$ . Observe that  $\underline{P}_{\mathcal{A}}$  and  $\overline{P}_{\mathcal{A}}$  always satisfy the conjugacy relation  $\underline{P}_{\mathcal{A}}(-f) = -\overline{P}_{\mathcal{A}}(f)$ . We call a real functional  $\underline{P}$  on  $\mathcal{G}(\Omega)$  a *coherent lower prevision* if there is some coherent set of desirable gambles  $\mathcal{R}$  on  $\mathcal{G}(\Omega)$  such that  $\underline{P} = \underline{P}_{\mathcal{R}}$ .

**Theorem 6.** *Let  $\mathcal{R}$  be a coherent set of desirable gambles. Then  $\underline{P}_{\mathcal{R}}$  is real-valued,  $\underline{P}_{\mathcal{R}} = \underline{P}_{\mathcal{D}_{\mathcal{R}}}$ , and  $\underline{P}_{\mathcal{R}}(f) \geq 0$  for all  $f \in \mathcal{D}_{\mathcal{R}}$ . Moreover, a real functional  $\underline{P}$  is a coherent lower prevision iff it satisfies the following properties, for all gambles  $f_1$  and  $f_2$  in  $\mathcal{G}(\Omega)$  and all real  $\lambda \geq 0$ :*

- P1.  $\underline{P}(f) \geq \inf f$  [accepting sure gain];
- P2.  $\underline{P}(f_1 + f_2) \geq \underline{P}(f_1) + \underline{P}(f_2)$  [super-additivity];
- P3.  $\underline{P}(\lambda f) = \lambda \underline{P}(f)$  [non-negative homogeneity].

A coherent lower prevision  $\underline{P}$  is called a *linear prevision* if it is self-conjugate, in the sense that  $\underline{P} = \overline{P}$ . Such a linear prevision can be seen as an expectation operator associated with a (finitely additive) probability. Using Proposition 2, it is not difficult to prove that the lower prevision  $\underline{P}_{\mathcal{R}}$  associated with a maximal coherent set  $\mathcal{R}$  is a linear prevision.<sup>8</sup>

Finally, we turn to marginal desirability. Given a coherent set of desirable gambles  $\mathcal{R}$ , we define the associated set of *marginally desirable gambles* as

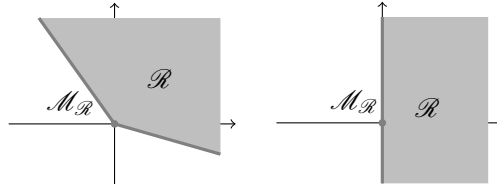
$$\mathcal{M}_{\mathcal{R}} := \{ f - \underline{P}_{\mathcal{R}}(f) : f \in \mathcal{G}(\Omega) \}. \quad (12)$$

The set of marginally desirable gambles  $\mathcal{M}_{\mathcal{R}}$  is completely determined by the lower prevision  $\underline{P}_{\mathcal{R}}$ . The converse is also true:

**Proposition 7.** *Let  $\mathcal{R}$  be a coherent set of desirable gambles. Then  $\underline{P}_{\mathcal{M}_{\mathcal{R}}} = \underline{P}_{\mathcal{R}}$  and*

$$\mathcal{M}_{\mathcal{R}} = \mathcal{M}_{\underline{P}_{\mathcal{R}}} := \{ f \in \mathcal{G}(\Omega) : \underline{P}_{\mathcal{R}}(f) = 0 \}. \quad (13)$$

The set of marginally desirable gambles  $\mathcal{M}_{\mathcal{R}}$  is the entire cone surface of  $\mathcal{R}$  and  $\mathcal{D}_{\mathcal{R}}$ , possibly including gambles that incur a partial (but not a sure) loss.



When  $\mathcal{R}$  is maximal,  $\mathcal{M}_{\mathcal{R}}$  constitutes a hyperplane.

<sup>7</sup>Compare this to the less stringent requirement for almost-desirability [27, Section 3.7.3]: if  $f \in \mathcal{D}_{\mathcal{R}}$  then  $\sup f \geq 0$  [avoiding sure loss].

<sup>8</sup>The proof for finite  $\Omega$ , given in [1], can be trivially extended to the infinite case. See also the proof of Proposition 24 in Appendix A.

**3.2. Updating sets of desirable gambles.** Consider a set of desirable gambles  $\mathcal{R}$  on  $\Omega$ . With a non-empty subset  $B$  of  $\Omega$ , we associate an *updated* set of desirable gambles on  $\Omega$ , as defined by Walley [28]:

$$\mathcal{R}|B := \{f \in \mathcal{G}(\Omega) : I_B f \in \mathcal{R}\}. \quad (14)$$

We find it more convenient to work with the following, slightly different but completely equivalent, version:

$$\mathcal{R}|B := \{f \in \mathcal{R} : I_B f = f\} = \mathcal{R} \cap \mathcal{G}(\Omega)|B, \quad (15)$$

which completely determines  $\mathcal{R}|B$ : for all  $f \in \mathcal{G}(\Omega)$ ,

$$f \in \mathcal{R}|B \Leftrightarrow I_B f \in \mathcal{R}|B. \quad (16)$$

In our version, updating corresponds to intersecting the convex cone  $\mathcal{R}$  with the linear subspace  $\mathcal{G}(\Omega)|B$ , which results in a convex cone  $\mathcal{R}|B$  of lower dimension. And since we can uniquely identify a gamble  $f = I_B f$  in  $\mathcal{G}(\Omega)|B$  with a gamble on  $B$ , namely its restriction  $f_B$  to  $B$ , and *vice versa*, we can also identify  $\mathcal{R}|B$  with a set of desirable gambles on  $B$ :

$$\mathcal{R}|B := \{f_B : f \in \mathcal{R}|B\} = \{f_B : f \in \mathcal{R}|B\} \subseteq \mathcal{G}(B). \quad (17)$$

**Proposition 8.** *If  $\mathcal{R}$  is a coherent set of desirable gambles on  $\Omega$ , then  $\mathcal{R}|B$  is coherent relative to  $\mathcal{G}(\Omega)|B$ , or equivalently,  $\mathcal{R}|B$  is a coherent set of desirable gambles on  $B$ .*

Our subject takes  $\mathcal{R}|B$  (or  $\mathcal{R}|B$ ) as his set of desirable gambles contingent on observing the event  $B$ .

#### 4. FINITE EXCHANGEABLE SEQUENCES

Now that we have become better versed in the theory of sets of desirable gambles, we are going to focus on the first main topic: reasoning about finite exchangeable sequences. We first show how they are related to count vectors (Section 4.1). Then we are ready to give a desirability-based definition of exchangeability (Section 4.2) and treat natural extension and updating under exchangeability (Sections 4.3 and 4.4). After presenting our Finite Representation Theorem (Section 4.5), we can show what natural extension and updating under exchangeability look like in terms of the count vector representation (Sections 4.6 and 4.7). Finally, we take a look at multinomial processes (Section 4.8), which will allow us to present a version of the Representation Theorem in terms of frequency vectors (Section 4.9).

Consider random variables  $X_1, \dots, X_N$  taking values in a non-empty finite set  $\mathcal{X}$ ,<sup>9</sup> where  $N \in \mathbb{N}_0$ , i.e., a positive (non-zero) integer. The possibility space is  $\Omega = \mathcal{X}^N$ .

**4.1. Count vectors.** We denote by  $x = (x_1, \dots, x_N)$  an arbitrary element of  $\mathcal{X}^N$ .  $\mathcal{P}_N$  is the set of all permutations  $\pi$  of the index set  $\{1, \dots, N\}$ . With any such permutation  $\pi$ , we associate a permutation of  $\mathcal{X}^N$ , also denoted by  $\pi$ , and defined by  $(\pi x)_k = x_{\pi(k)}$ , or in other words,  $\pi(x_1, \dots, x_N) = (x_{\pi(1)}, \dots, x_{\pi(N)})$ . Similarly, we lift  $\pi$  to a permutation  $\pi^t$  of  $\mathcal{G}(\mathcal{X}^N)$  by letting  $\pi^t f = f \circ \pi$ , so  $(\pi^t f)(x) = f(\pi x)$ .

The permutation invariant atoms  $[x] := \{\pi x : \pi \in \mathcal{P}_N\}$ ,  $x \in \mathcal{X}^N$  are the smallest permutation invariant subsets of  $\mathcal{X}^N$ . We introduce the *counting map*

$$T^N : \mathcal{X}^N \rightarrow \mathcal{N}^N : x \mapsto T^N(x), \quad (18)$$

where  $T^N(x)$  is the  $\mathcal{X}$ -tuple with components

$$T_z^N(x) := |\{k \in \{1, \dots, N\} : x_k = z\}| \quad \text{for all } z \in \mathcal{X}, \quad (19)$$

<sup>9</sup>A lot of functions and sets introduced below will depend on the set  $\mathcal{X}$ . We do not indicate this explicitly, to not overburden the notation and because we do not consider different sets of values in this paper.

and the set of possible *count vectors* is given by

$$\mathcal{N}^N := \left\{ m \in \mathbb{N}^{\mathcal{X}} : \sum_{x \in \mathcal{X}} m_x = N \right\}. \quad (20)$$

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  $[m]$ .

**Example 1** (Running example). To familiarise ourselves with some of the concepts introduced, we will use a running example, in which we assume a sample space  $\mathcal{X} := \{b, w\}$  (for *black* and *white* – the colours of marbles drawn from an urn containing a mixture of such marbles). Consider the situation  $N := 2$ , then

$$\mathcal{X}^N = \{b, w\}^2 = \{(b, b), (b, w), (w, b), (w, w)\} \quad \text{and} \quad \mathcal{N}^N = \{(2, 0), (1, 1), (0, 2)\}.$$

Their correspondence and the non-trivial permutations are implicitly given by

$$[2, 0] = \{(b, b)\}, \quad [1, 1] = \{(b, w), (w, b)\}, \quad [0, 2] = \{(w, w)\}. \quad \blacklozenge$$

**4.2. Defining exchangeability.** If a subject assesses that  $X_1, \dots, X_N$  are exchangeable, this means that for any gamble  $f$  and any permutation  $\pi$ , he finds exchanging  $\pi^t f$  for  $f$  weakly desirable, because he is indifferent between them [cf. 27, Section 4.1.1]. Let

$$\mathcal{D}_{\mathcal{P}_N} := \{f - \pi^t f : f \in \mathcal{G}(\mathcal{X}^N) \text{ and } \pi \in \mathcal{P}_N\}, \quad (21)$$

then we should have that  $\mathcal{D}_{\mathcal{P}_N} \subseteq \mathcal{D}_{\mathcal{R}}$ .<sup>10</sup> This is the basis for our definition of exchangeability.

What we would like to do now, is to provide alternative characterisations of exchangeability. These will be useful for the further development, and provide additional insight into what an assessment of exchangeability amounts to.

We begin by defining a special linear transformation  $\text{ex}^N$  of the linear space of gambles  $\mathcal{G}(\mathcal{X}^N)$ :

$$\text{ex}^N : \mathcal{G}(\mathcal{X}^N) \rightarrow \mathcal{G}(\mathcal{X}^N) : f \mapsto \text{ex}^N(f) := \frac{1}{N!} \sum_{\pi \in \mathcal{P}_N} \pi^t f. \quad (22)$$

The idea behind this linear transformation  $\text{ex}^N$  is that it renders a gamble  $f$  insensitive to permutation by replacing it with the uniform average  $\text{ex}^N(f)$  of all its permutations  $\pi^t f$ . Indeed, observe that for all gambles  $f$  and all permutations  $\pi$ :

$$\text{ex}^N(\pi^t f) = \text{ex}^N(f) \quad \text{and} \quad \pi^t(\text{ex}^N(f)) = \text{ex}^N(f). \quad (23)$$

So  $\text{ex}^N(f)$  is permutation invariant and therefore constant on the permutation invariant atoms  $[m]$ , and it assumes the same value for all gambles that can be related to each other through some permutation. But then, what is the value that  $\text{ex}^N(f)$  assumes on each such permutation invariant atom  $[m]$ ? It is not difficult to see that

$$\text{ex}^N = \sum_{m \in \mathcal{N}^N} \text{Hy}^N(\cdot | m) I_{[m]}, \quad (24)$$

or in other words, the gamble  $\text{ex}^N(f)$  assumes the constant value  $\text{Hy}^N(f|m)$  on  $[m]$ , where we let

$$\text{Hy}^N(f|m) := \frac{1}{|[m]|} \sum_{y \in [m]} f(y) \quad \text{and} \quad |[m]| = \binom{N}{m} := \frac{N!}{\prod_{z \in \mathcal{X}} m_z!}. \quad (25)$$

$\text{Hy}^N(\cdot | m)$  is the linear expectation operator associated with the uniform distribution on the invariant atom  $[m]$ . It characterises a (multivariate) *hyper-geometric distribution* [15, Section 39.2], associated with random sampling without replacement from an urn with  $N$  balls of types  $\mathcal{X}$ , whose composition is characterised by the count vector  $m$ .

<sup>10</sup>Note that the gambles in  $\mathcal{D}_{\mathcal{P}_N}$  cannot be assumed to be desirable, because  $\mathcal{D}_{\mathcal{P}_N}$  does not avoid non-positivity.



**Example 2.** To get some feeling for what this means, let us go back to our running example  $\mathcal{X} := \{b, w\}$  and let  $N := 4$  and  $m = (m_b, m_w) := (2, 2)$ . Then

$$[m] = [2, 2] = \{(w, w, b, b), (w, b, w, b), (w, b, b, w), (b, b, w, w), (b, w, w, b), (b, w, b, w)\}.$$

Consider the event  $A$  that amongst the first three observations, we see twice  $b$  and once  $w$ :

$$A = \{(b, b, w), (b, w, b), (w, b, b)\} \times \{b, w\}.$$

Then

$$\text{Hy}^4(I_A|2, 2) = \frac{1}{\binom{4}{2}} \sum_{y \in [2, 2]} I_A(y) = \frac{1}{6} \cdot 3 = \frac{1}{2}$$

is the probability of getting two black balls and one white when drawing three balls without replacement from an urn containing two black and two white balls.  $\blacklozenge$

So we see that the linear transformation  $\text{ex}^N$  is intricately linked with the  $N$ -variate hypergeometric distribution. If we also observe that  $\text{ex}^N \circ \text{ex}^N = \text{ex}^N$ , we see that  $\text{ex}^N$  is the linear projection operator of the linear space  $\mathcal{G}(\mathcal{X}^N)$  to the linear subspace

$$\mathcal{G}_{\mathcal{P}_N}(\mathcal{X}^N) := \{f \in \mathcal{G}(\mathcal{X}^N) : (\forall \pi \in \mathcal{P}_N) \pi^t f = f\} \quad (26)$$

of all permutation invariant gambles.

The linear transformation  $\text{ex}^N$  is also tightly connected with the previously defined set  $\mathcal{D}_{\mathcal{P}_N}$  of gambles  $f - \pi^t f$  that play a role in defining exchangeability. Indeed, if we look at the linear subspace  $\mathcal{D}_{\mathcal{U}_N}$  that is generated by such gambles  $f - \pi^t f$ , then it is not hard to see that

$$\mathcal{D}_{\mathcal{U}_N} := \text{span}(\mathcal{D}_{\mathcal{P}_N}) = \left\{ \sum_{k=1}^n \lambda_k f_k : n \geq 0, \lambda_k \in \mathbb{R}, f_k \in \mathcal{D}_{\mathcal{P}_N} \right\} \quad (27)$$

$$= \{f - \text{ex}^N(f) : f \in \mathcal{G}(\mathcal{X}^N)\} = \{f \in \mathcal{G}(\mathcal{X}^N) : \text{ex}^N(f) = 0\}, \quad (28)$$

where ‘ $\text{span}(\cdot)$ ’ denotes *linear span* of its argument set: the set of all linear combinations of elements from that set. The last equality tells us that the linear subspace  $\mathcal{D}_{\mathcal{U}_N}$  is the *kernel* of the linear projection operator  $\text{ex}^N$ : it contains precisely those gambles that are mapped to 0 by  $\text{ex}^N$ .

**Example 3.** Let us return for a moment to our running example  $\mathcal{X} := \{b, w\}$  and  $N := 2$ , then

$$\mathcal{D}_{\mathcal{P}_N} = \{f \in \mathcal{G}(\mathcal{X}^N) : f(b, b) = f(w, w) = 0 \text{ and } f(b, w) = -f(w, b)\}, \quad (29)$$

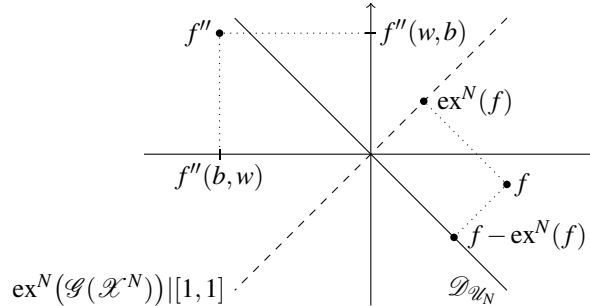
and  $\mathcal{D}_{\mathcal{U}_N} = \mathcal{D}_{\mathcal{P}_N}$ . Let  $f$  be some gamble on  $\mathcal{X}^N$  and  $f' := \text{ex}^N(f)$ , then

$$f'(b, b) = \text{Hy}^N(f|2, 0) = f(b, b),$$

$$f'(w, w) = \text{Hy}^N(f|0, 2) = f(w, w),$$

$$f'(b, w) = f'(w, b) = \text{Hy}^N(f|1, 1) = \frac{1}{2}(f(b, w) + f(w, b)).$$

The plane  $\mathcal{G}(\mathcal{X}^N)|[1, 1] = \{f \in \mathcal{G}(\mathcal{X}^N) : f(b, b) = f(w, w) = 0\}$ , which includes  $\mathcal{D}_{\mathcal{P}_N}$  and thus  $\mathcal{D}_{\mathcal{U}_N}$ , can be used for a graphical illustration:



We show the intersection of this plane and the range of the operator  $\text{ex}^N$  as well as its effect on a gamble  $f$  in  $\mathcal{G}(\mathcal{X}^N)|[1, 1]$ . The gamble  $f''$  is included to make it explicit which components are actually depicted.  $\blacklozenge$

**Definition 3** (Exchangeability). *A coherent set  $\mathcal{R}$  of desirable gambles on  $\mathcal{X}^N$  is called exchangeable if any (and hence all) of the following equivalent conditions is (are) satisfied:*

- (i) *all gambles in  $\mathcal{D}_{\mathcal{P}_N}$  are weakly desirable:  $\mathcal{D}_{\mathcal{P}_N} \subseteq \mathcal{D}_{\mathcal{R}}$ ;*
- (ii)  *$\mathcal{D}_{\mathcal{P}_N} + \mathcal{R} \subseteq \mathcal{R}$ ;*
- (iii) *all gambles in  $\mathcal{D}_{\mathcal{U}_N}$  are weakly desirable:  $\mathcal{D}_{\mathcal{U}_N} \subseteq \mathcal{D}_{\mathcal{R}}$ ;*
- (iv)  *$\mathcal{D}_{\mathcal{U}_N} + \mathcal{R} \subseteq \mathcal{R}$ ;*

*We call a coherent lower prevision  $\underline{P}$  on  $\mathcal{G}(\mathcal{X}^N)$  exchangeable if there is some exchangeable coherent set of desirable gambles  $\mathcal{R}$  such that  $\underline{P} = \underline{P}_{\mathcal{R}}$ .*

Because they are stated in terms of the kernel  $\mathcal{D}_{\mathcal{U}_N}$  of the projection operator  $\text{ex}^N$ , which we have seen is intricately linked with (multivariate) hypergeometric distributions, conditions (iii) and (iv) of this definition are quite closely related to the desirability version of a de Finetti-like representation theorem for finite exchangeable sequences in terms of sampling without replacement from an urn. They allow us to talk about exchangeability without explicitly invoking permutations. This is what we will address in Section 4.5.

**Example 4.** In our running example, if  $f \in \mathcal{G}(\mathcal{X}^N)$  is desirable, then because of the definition of exchangeability and Eq. (29), all gambles in the linear subspace

$$\{f' \in \mathcal{G}(\mathcal{X}^N) : f'(b, b) = f(b, b), f'(w, w) = f(w, w), \\ f'(b, w) + f'(w, b) = f(b, w) + f(w, b)\}$$

are also desirable. So under exchangeability transfers between  $(b, w)$  and  $(w, b)$ -components are irrelevant for desirability. This illustrates that, generally and geometrically speaking, any exchangeable set of desirable gambles  $\mathcal{R} \subseteq \mathcal{G}(\mathcal{X}^N)$  for any  $\mathcal{X}$  and any  $N > 1$  must be cylindrical along the directions in the linear subspace  $\mathcal{D}_{\mathcal{U}_N}$ .  $\blacklozenge$

A number of useful results follow from Definition 3:

**Proposition 9.** *Let  $\mathcal{R}$  be a coherent set of desirable gambles. If  $\mathcal{R}$  is exchangeable then it is also permutable:  $\pi^t f \in \mathcal{R}$  for all  $f \in \mathcal{R}$  and all  $\pi \in \mathcal{P}_N$ .*

We have seen above that the gambles  $f - \pi^t f$  in  $\mathcal{D}_{\mathcal{P}_N}$  span the kernel  $\mathcal{D}_{\mathcal{U}_N}$  of the linear operator  $\text{ex}^N$  that projects any gamble  $f$  on its symmetrised counterpart  $\text{ex}^N(f)$ . It should therefore not come as a surprise that for an exchangeable model, we can determine whether a gamble is desirable by looking at this symmetrised counterpart.

**Proposition 10.** *Let  $\mathcal{R}$  be a coherent and exchangeable set of desirable gambles. For all gambles  $f$  and  $f'$  on  $\mathcal{X}^N$ :*

- (i)  *$f \in \mathcal{R} \Leftrightarrow \text{ex}^N(f) \in \mathcal{R}$ ;*
- (ii) *If  $\text{ex}^N(f) = \text{ex}^N(f')$ , then  $f \in \mathcal{R} \Leftrightarrow f' \in \mathcal{R}$ .*

It follows from this last proposition and Eq. (28) that for any coherent and exchangeable set of desirable gambles  $\mathcal{R}$ :

$$\mathcal{R} \cap \mathcal{D}_{\mathcal{U}_N} = \emptyset. \quad (30)$$

We can use these ideas to derive a direct characterisation for the exchangeability of a lower prevision, without the intervention of sets of desirable gambles.

**Theorem 11.** *Let  $\underline{P}$  be a coherent lower prevision on  $\mathcal{G}(\mathcal{X}^N)$ . Then the following statements are equivalent.<sup>11</sup>*

- (i)  *$\underline{P}$  is exchangeable;*

<sup>11</sup>This also shows that the exchangeability of a lower prevision can also be expressed using marginally desirable gambles [see 23, Section 3.1.1].

- (ii)  $\underline{P}(f) = \overline{P}(f) = 0$  for all  $f \in \mathcal{D}_{\mathcal{F}_N}$ ;
- (iii)  $\underline{P}(f) = \overline{P}(f) = 0$  for all  $f \in \mathcal{D}_{\mathcal{U}_N}$ .

4.3. **Exchangeable natural extension.** Let us denote the set of all coherent and exchangeable sets of desirable gambles on  $\mathcal{X}^N$  by

$$\mathbb{D}_{\text{ex}}(\mathcal{X}^N) := \{\mathcal{R} \in \mathbb{D}(\mathcal{X}^N) : \mathcal{D}_{\mathcal{U}_N} + \mathcal{R} \subseteq \mathcal{R}\}. \quad (31)$$

This set is closed under arbitrary non-empty intersections. We shall see further on in Corollary 14 that it is also non-empty, and therefore has a smallest element.

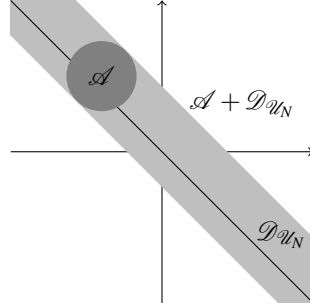
Suppose our subject has an assessment, or in other words, a set  $\mathcal{A}$  of gambles on  $\mathcal{X}^N$  that he finds desirable. Then we can ask if there is some coherent and exchangeable set of desirable gambles  $\mathcal{R}$  that includes  $\mathcal{A}$ . In other words, we want a set of desirable gambles  $\mathcal{R}$  to satisfy the requirements: (i)  $\mathcal{R}$  is coherent; (ii)  $\mathcal{A} \subseteq \mathcal{R}$ ; and (iii)  $\mathcal{D}_{\mathcal{U}_N} + \mathcal{R} \subseteq \mathcal{R}$ . The intersection  $\bigcap_{i \in I} \mathcal{R}_i$  of an arbitrary non-empty family of sets of desirable gambles  $\mathcal{R}_i$ ,  $i \in I$  that satisfy these requirements, will satisfy these requirements as well. This is the idea behind the following definition and results.

**Definition 4** (Avoiding non-positivity under exchangeability). *We say that a set  $\mathcal{A}$  of gambles on  $\mathcal{X}^N$  avoids non-positivity under exchangeability if  $[\mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A}] + \mathcal{D}_{\mathcal{U}_N}$  avoids non-positivity.*

**Proposition 12.** (i)  $\emptyset$  avoids non-positivity under exchangeability;

- (ii) A non-empty set of gambles  $\mathcal{A}$  on  $\mathcal{X}^N$  avoids non-positivity under exchangeability iff  $\mathcal{A} + \mathcal{D}_{\mathcal{U}_N}$  avoids non-positivity.

**Example 5.** For our running example, avoiding non-positivity under exchangeability is best illustrated graphically—again in the plane  $\mathcal{G}(\mathcal{X}^N)|[1, 1]$ —for the case that the given assessment  $\mathcal{A} \subset \mathcal{G}(\mathcal{X}^N)|[1, 1]$  avoids non-positivity, but not so under exchangeability:



(As a reminder:  $\mathcal{G}(\mathcal{X}^N)|[1, 1] = \{f \in \mathcal{G}(\mathcal{X}^N) : f(b, b) = f(w, w) = 0\}$ .)  $\blacklozenge$

**Theorem 13** (Exchangeable natural extension). *Consider a set  $\mathcal{A}$  of gambles on  $\mathcal{X}^N$ , and define its exchangeable natural extension  $\mathcal{E}_{\text{ex}}^N(\mathcal{A})$  by*

$$\mathcal{E}_{\text{ex}}^N(\mathcal{A}) := \bigcap \{\mathcal{R} \in \mathbb{D}_{\text{ex}}(\mathcal{X}^N) : \mathcal{A} \subseteq \mathcal{R}\}. \quad (32)$$

Then the following statements are equivalent:

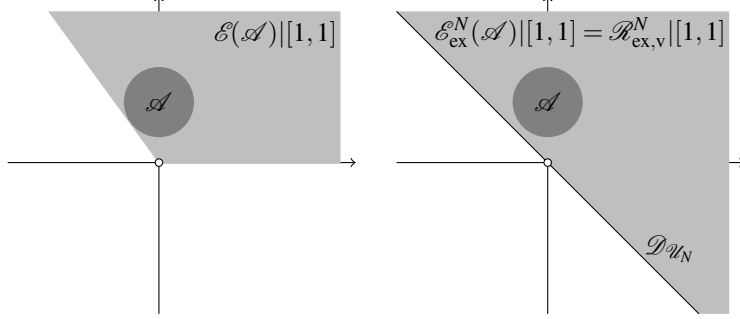
- (i)  $\mathcal{A}$  avoids non-positivity under exchangeability;
- (ii)  $\mathcal{A}$  is included in some coherent and exchangeable set of desirable gambles;
- (iii)  $\mathcal{E}_{\text{ex}}^N(\mathcal{A}) \neq \mathcal{G}(\mathcal{X}^N)$ ;
- (iv)  $\mathcal{E}_{\text{ex}}^N(\mathcal{A})$  is a coherent and exchangeable set of desirable gambles;
- (v)  $\mathcal{E}_{\text{ex}}^N(\mathcal{A})$  is the smallest coherent and exchangeable set of desirable gambles that includes  $\mathcal{A}$ .

When any (and hence all) of these equivalent statements hold, then

$$\mathcal{E}_{\text{ex}}^N(\mathcal{A}) = \text{posi}(\mathcal{D}_{\mathcal{U}_N} + [\mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A}]) \quad (33)$$

$$= \mathcal{D}_{\mathcal{U}_N} + \mathcal{E}(\mathcal{A}). \quad (34)$$

**Example 6.** Exchangeable natural extension is again best illustrated graphically. We use an assessment  $\mathcal{A} \subset \mathcal{G}(\mathcal{X}^N)|[1, 1]$  that avoids non-positivity under exchangeability and contrast natural extension with exchangeable natural extension:



Only the parts of an assessment that fall outside of  $\mathcal{R}_{\text{ex},v}^N$  have a nontrivial impact. This is something that cannot be illustrated in the context of drawings as these, but think about what the extensions would look like if the assessment would consist of the singleton  $\{f\}$  with  $f(b, b) = -f(w, w) = 1$  and  $f(b, w) = f(w, b) = 0$ : the natural extension  $\mathcal{E}(\{f\})$  consists of all gambles  $\mu\lambda f + (1-\mu)g$ , with  $\mu \in [0, 1]$ ,  $\lambda \in \mathbb{R}_0^+$ , and  $g \in \mathcal{G}_0^+(\mathcal{X}^N)$ , so the exchangeable natural extension  $\mathcal{E}_{\text{ex}}^N(\{f\})$  consists of all gambles  $\rho h + \mu\lambda f + (1-\mu)g$ , where additionally  $\rho \in \mathbb{R}$  and  $h(b, w) = -h(w, b) = 1$  and  $h(b, b) = h(w, w) = 0$ .  $\blacklozenge$

Eq. (34) shows that if we have an assessment  $\mathcal{A}$  with a finite description, it is possible, but not necessarily very efficient, to represent its exchangeable natural extension on a computer: besides the finite description of its extreme rays of  $\mathcal{E}(\mathcal{A})$ , we need to account for taking the Minkowski sum with  $\mathcal{D}_{\mathcal{M}_N}$ . We shall see further on in Theorem 19 that this extra complication can be circumvented by working with so-called count representations.

There is always a most conservative exchangeable belief model, which represents the effects of making only an assessment of exchangeability, and nothing more:

**Corollary 14.** *The set  $\mathbb{D}_{\text{ex}}(\mathcal{X}^N)$  is non-empty, and has a smallest element*

$$\mathcal{R}_{\text{ex},v}^N := \mathcal{E}_{\text{ex}}^N(\mathbf{0}) = \mathcal{D}_{\mathcal{M}_N} + \mathcal{G}_0^+(\mathcal{X}^N). \quad (35)$$

**4.4. Updating exchangeable models.** Consider an exchangeable and coherent set of desirable gambles  $\mathcal{R}$  on  $\mathcal{X}^N$ , and assume that we have observed the values  $\check{x} = (\check{x}_1, \check{x}_2, \dots, \check{x}_{\check{n}})$  of the first  $\check{n}$  variables  $X_1, \dots, X_{\check{n}}$ , and that we want to make inferences about the remaining  $\hat{n} := N - \check{n}$  variables. To do this, we simply update the set  $\mathcal{R}$  with the event  $C_{\check{x}} := \{\check{x}\} \times \mathcal{X}^{\hat{n}}$ , to obtain the set  $\mathcal{R}|C_{\check{x}}$ , also denoted as  $\mathcal{R}|\check{x} := \{f \in \mathcal{R} : f|_{C_{\check{x}}} = f\}$ . As we have seen in Section 3.2, this set can be identified with a coherent set of desirable gambles on  $\mathcal{X}^{\hat{n}}$ , which we denote by  $\mathcal{R}|\check{x}$ . With obvious notations:<sup>12</sup>

$$\mathcal{R}|\check{x} := \{f \in \mathcal{G}(\mathcal{X}^{\hat{n}}) : f|_{C_{\check{x}}} \in \mathcal{R}\}. \quad (36)$$

We already know that updating preserves coherence. We now see that this type of updating on an observed sample also preserves exchangeability.

**Proposition 15.** *Consider  $\check{x} \in \mathcal{X}^{\check{n}}$  and a coherent and exchangeable set of desirable gambles  $\mathcal{R}$  on  $\mathcal{X}^N$ . Then  $\mathcal{R}|\check{x}$  is a coherent and exchangeable set of desirable gambles on  $\mathcal{X}^{\hat{n}}$ .*

We also introduce another type of updating, where we observe a count vector  $\check{m} \in \mathcal{N}^{\check{n}}$ , and we update the set  $\mathcal{R}$  with the set  $C_{\check{m}} := [\check{m}] \times \mathcal{X}^{\hat{n}}$ , to obtain the set  $\mathcal{R}|C_{\check{m}}$ , also denoted

<sup>12</sup>Here and further on we silently use cylindrical extension on gambles, i.e., let them ‘depend’ on extra variables whose value does not influence the value they take.

as  $\mathcal{R}|\check{m} := \{f \in \mathcal{R} : fI_{C_{\check{m}}} = f\}$ . This set can be identified with a coherent set of desirable gambles on  $\mathcal{X}^{\check{n}}$ , which we also denote by  $\mathcal{R}|\check{m}$ . With obvious notations:

$$\mathcal{R}|\check{m} := \{f \in \mathcal{G}(\mathcal{X}^{\check{n}}) : fI_{C_{\check{m}}} \in \mathcal{R}\}. \quad (37)$$

Interestingly, the count vector  $\check{m}$  for an observed sample  $\check{x}$  is a *sufficient statistic* in that it extracts from  $\check{x}$  all the information that is needed to characterise the updated model:

**Proposition 16** (Sufficiency of observed count vectors). *Consider  $\check{x}, \check{y} \in \mathcal{X}^{\check{n}}$  and a coherent and exchangeable set of desirable gambles  $\mathcal{R}$  on  $\mathcal{X}^N$ . If  $\check{y} \in [\check{x}]$ , or in other words if  $T^{\check{n}}(\check{x}) = T^{\check{n}}(\check{y}) =: \check{m}$ , then  $\mathcal{R}|\check{x} = \mathcal{R}|\check{y} = \mathcal{R}|\check{m}$ .*

**4.5. Finite representation.** We can use the symmetry that an assessment of exchangeability generates to represent an exchangeable coherent set of desirable gambles in a much more economical, or condensed, fashion. This has already been made apparent in Proposition 10, where we saw that the desirability of any gamble  $f$  can be determined by looking at the desirability of its symmetrised counterpart  $\text{ex}^N(f)$ . We have seen that this projection  $\text{ex}^N(f)$  of  $f$  onto the linear subspace  $\mathcal{G}_{\mathcal{P}_N}(\mathcal{X}^N)$  of permutation invariant gambles assumes the constant value  $\text{Hy}^N(f|m)$  on the permutation invariant atoms  $[m]$ ,  $m \in \mathcal{N}^N$ .

Now, since a gamble is permutation invariant if and only if it is constant on these permutation invariant atoms, we can *identify* permutation invariant gambles on  $\mathcal{X}^N$  with gambles on  $\mathcal{N}^N$ . This identification is made more formal using the following linear isomorphism  $\text{Co}^N$  between the linear spaces  $\mathcal{G}(\mathcal{N}^N)$  and  $\mathcal{G}_{\mathcal{P}_N}(\mathcal{X}^N)$ :

$$\text{Co}^N : \mathcal{G}(\mathcal{N}^N) \rightarrow \mathcal{G}_{\mathcal{P}_N}(\mathcal{X}^N) : g \mapsto \text{Co}^N(g) := g \circ T^N, \quad (38)$$

so  $\text{Co}^N(g)$  is the permutation invariant gamble on  $\mathcal{X}^N$  that assumes the constant value  $g(m)$  on the invariant atom  $[m]$ .

Through the mediation of this identification  $\text{Co}^N$ , we can use the projection operator  $\text{ex}^N$  to turn a gamble  $f$  on  $\mathcal{X}^N$  into a gamble on  $\mathcal{N}^N$ , as follows:

$$\text{Hy}^N : \mathcal{G}(\mathcal{X}^N) \rightarrow \mathcal{G}(\mathcal{N}^N) : f \mapsto \text{Hy}^N(f) := \text{Hy}^N(f|\cdot), \quad (39)$$

so  $\text{Hy}^N(f)$  is the gamble on  $\mathcal{N}^N$  that assumes the value  $\text{Hy}^N(f|m)$  in the count vector  $m \in \mathcal{N}^N$ . By definition,  $\text{ex}^N(f) = \text{Co}^N(\text{Hy}^N(f))$  for all  $f \in \mathcal{G}(\mathcal{X}^N)$ , and similarly  $\text{Hy}^N(\text{Co}^N(g)) = g$ , for all  $g \in \mathcal{G}(\mathcal{N}^N)$ . Hence:

$$\text{ex}^N = \text{Co}^N \circ \text{Hy}^N \quad \text{and} \quad \text{Hy}^N \circ \text{Co}^N = \text{id}_{\mathcal{G}(\mathcal{N}^N)}. \quad (40)$$

Since  $\text{ex}^N$  is a projection operator, its restriction to  $\mathcal{G}_{\mathcal{P}_N}(\mathcal{X}^N)$  is the identity map, and therefore we infer from Eq. (40) that the restriction of  $\text{Hy}^N$  to  $\mathcal{G}_{\mathcal{P}_N}(\mathcal{X}^N)$  is the inverse of  $\text{Co}^N$ , and therefore also a linear isomorphism between  $\mathcal{G}_{\mathcal{P}_N}(\mathcal{X}^N)$  and  $\mathcal{G}(\mathcal{N}^N)$ .

If we invoke Eq. (23) we find that

$$\text{Hy}^N(\pi^t f) = \text{Hy}^N(f). \quad (41)$$

Also taking into account the linearity of  $\text{Hy}^N$  and Eq. (22), this leads to

$$\text{Hy}^N(\text{ex}^N(f)) = \text{Hy}^N(f). \quad (42)$$

The relationships between the three important linear maps  $\text{ex}^N$ ,  $\text{Hy}^N$  and  $\text{Co}^N$  we have introduced above are clarified by the commutative diagram in Fig. 1. (The bottom part of the diagram can be safely ignored for now.)

**Example 7.** In the context of our running example, we have the following: Take some gamble  $f$  on  $\mathcal{X}^N$  and let  $g := \text{Hy}^N(f)$ , then

$$g(2,0) = f(b,b), \quad g(0,2) = f(w,w), \quad g(1,1) = \frac{1}{2}(f(b,w) + f(w,b)).$$

Conversely, take some gamble  $g$  on  $\mathcal{N}^N$  and let  $f := \text{Co}^N(g)$ , then

$$f(b,b) = g(2,0), \quad f(w,w) = g(0,2), \quad f(b,w) = f(w,b) = g(1,1). \quad \blacklozenge$$

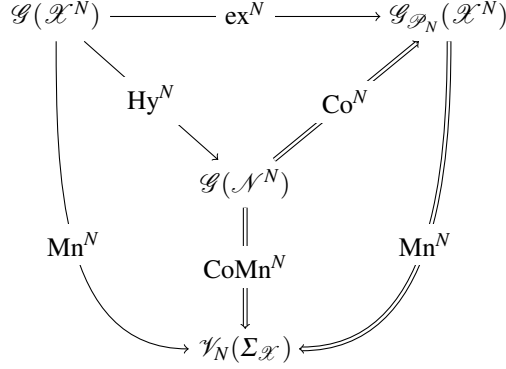


FIGURE 1. Single sequence length commutative diagram. Single arrows indicate linear monomorphisms (injective). Double arrows indicate linear isomorphisms (bijective).

For every gamble  $f$  on  $\mathcal{X}^N$ ,  $f = \text{ex}^N(f) + [f - \text{ex}^N(f)]$ , so it can be decomposed as a sum of a permutation invariant gamble  $\text{ex}^N(f)$  and an element  $f - \text{ex}^N(f)$  of the kernel  $\mathcal{D}_{\mathcal{M}_N}$  of the linear projection operator  $\text{ex}^N$ . Elements of this kernel are, by definition, irrelevant as far as desirability under exchangeability is concerned, so the only part of this decomposition that matters is the element  $\text{ex}^N(f)$  of  $\mathcal{G}_{\mathcal{D}_N}(\mathcal{X}^N)$ . Since we have seen that  $\text{Hy}^N$  acts as a linear isomorphism between the linear spaces  $\mathcal{G}_{\mathcal{D}_N}(\mathcal{X}^N)$  and  $\mathcal{G}(\mathcal{N}^N)$ , we now investigate whether we can use  $\text{Hy}^N$  to represent a coherent and exchangeable  $\mathcal{R}$  by some set of desirable *count gambles* on  $\mathcal{N}^N$ .

**Theorem 17** (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*

$$\mathcal{R} = (\text{Hy}^N)^{-1}(\mathcal{S}), \quad (43)$$

and in that case this  $\mathcal{S}$  is uniquely determined by

$$\mathcal{S} = \{g \in \mathcal{G}(\mathcal{N}^N) : \text{Co}^N(g) \in \mathcal{R}\} = \text{Hy}^N(\mathcal{R}). \quad (44)$$

This leads to the following representation result for lower previsions, formulated without the mediation of coherent sets of desirable gambles.

**Corollary 18.** *A lower prevision  $\underline{P}$  on  $\mathcal{G}(\mathcal{X}^N)$  is coherent and exchangeable iff there is some coherent lower prevision  $\underline{Q}$  on  $\mathcal{G}(\mathcal{N}^N)$  such that  $\underline{P} = \underline{Q} \circ \text{Hy}^N$ . In that case  $\underline{Q}$  is uniquely determined by  $\underline{Q} = \underline{P} \circ \text{Co}^N$ .*

We call the set  $\mathcal{S}$  and the lower prevision  $\underline{Q}$  the *count representations* of the exchangeable set  $\mathcal{R}$  and the exchangeable lower prevision  $\underline{P}$ , respectively. Our Finite Representation Theorem allows us to give an appealing geometrical interpretation to the notions of exchangeability and representation. The exchangeability of  $\mathcal{R}$  means that it is completely determined by its count representation  $\text{Hy}^N(\mathcal{R})$ , or what amounts to the same thing since  $\text{Co}^N$  is a linear isomorphism: by its projection  $\text{ex}^N(\mathcal{R})$  on the linear space of all permutation invariant gambles. This turns count vectors into useful sufficient statistics (compare with Proposition 16), because the dimension of  $\mathcal{G}(\mathcal{N}^N)$  is typically much smaller than that of  $\mathcal{G}(\mathcal{X}^N)$ . To give an easy example: when  $\mathcal{X}$  has two elements,  $\mathcal{G}(\mathcal{X}^N)$  has dimension  $2^N$ , whereas the dimension of  $\mathcal{G}(\mathcal{N}^N)$  is only  $N + 1$ .

**4.6. Exchangeable natural extension and representation.** The exchangeable natural extension is easy to calculate using natural extension in terms of count representations, and the following simple result therefore has important consequences for practical implementations of reasoning and inference under exchangeability.

**Theorem 19.** Let  $\mathcal{A}$  be a set of gambles on  $\mathcal{X}^N$ , then

- (i)  $\mathcal{A}$  avoids non-positivity under exchangeability iff  $\text{Hy}^N(\mathcal{A})$  avoids non-positivity.
- (ii)  $\text{Hy}^N(\mathcal{E}_{\text{ex}}^N(\mathcal{A})) = \mathcal{E}(\text{Hy}^N(\mathcal{A}))$ .

This result gives us an extra approach to calculating the exchangeable natural extension of an assessment. It reduces calculating the exchangeable natural extension to calculating a natural extension in the lower dimensional space of count gambles. The commutative diagram that corresponds to it is given in Fig. 2.

$$\begin{array}{ccc}
 \mathcal{G}(\mathcal{X}^N) & \xrightarrow{\quad \text{Hy}^N \quad} & \mathcal{G}(\mathcal{N}^N) \\
 \downarrow \mathcal{E}_{\text{ex}}^N & & \downarrow \mathcal{E} \\
 \mathcal{G}(\mathcal{X}^N) & \xrightarrow{\quad \text{Hy}^N \quad} & \mathcal{G}(\mathcal{N}^N)
 \end{array}$$

FIGURE 2. The relationship between exchangeable natural extension and count representation natural extension. The arrows indicate monomorphisms (injective).

**4.7. Updating and representation.** Suppose, as in Section 4.4, that we update a coherent and exchangeable set of desirable gambles  $\mathcal{R}$  after observing a sample  $\check{x}$  with count vector  $\check{m}$ . This leads to an updated coherent and exchangeable set of desirable gambles  $\mathcal{R} \upharpoonright \check{x} = \mathcal{R} \upharpoonright \check{m}$  on  $\mathcal{X}^{\hat{n}}$ . Here, we take a closer look at the corresponding set of desirable gambles on  $\mathcal{N}^{\hat{n}}$ , which we denote (symbolically) by  $\mathcal{S} \upharpoonright \check{m}$ . (But we do not want to suggest with this notation that this is in some way an updated set of gambles!) The Finite Representation Theorem 17 tells us that  $\mathcal{S} \upharpoonright \check{m} = \text{Hy}^{\hat{n}}(\mathcal{R} \upharpoonright \check{m})$ , but is there a direct way to infer the count representation  $\mathcal{S} \upharpoonright \check{m}$  of  $\mathcal{R} \upharpoonright \check{m}$  from the count representation  $\mathcal{S} = \text{Hy}^N(\mathcal{R})$  of  $\mathcal{R}$ ?

To show that there is, we need to introduce two new notions: the *likelihood function*

$$L_{\check{m}}: \mathcal{N}^{\hat{n}} \rightarrow \mathbb{R}: \hat{m} \mapsto L_{\check{m}}(\hat{m}) := \frac{|\check{m}| |\hat{m}|}{|\check{m} + \hat{m}|}, \quad (45)$$

associated with sampling without replacement, and the linear map  $+_{\check{m}}$  from the linear space  $\mathcal{G}(\mathcal{N}^{\hat{n}})$  to the linear space  $\mathcal{G}(\mathcal{N}^N)$  given by

$$+_{\check{m}}: \mathcal{G}(\mathcal{N}^{\hat{n}}) \rightarrow \mathcal{G}(\mathcal{N}^N): g \mapsto +_{\check{m}}g \quad (46)$$

where

$$+_{\check{m}}g(M) = \begin{cases} g(M - \check{m}) & \text{if } M \geq \check{m} \\ 0 & \text{otherwise.} \end{cases} \quad (47)$$

**Proposition 20.** Consider a coherent and exchangeable set of desirable gambles  $\mathcal{R}$  on  $\mathcal{X}^N$ , with count representation  $\mathcal{S}$ . Let  $\mathcal{S} \upharpoonright \check{m}$  be the count representation of the coherent and exchangeable set of desirable gambles  $\mathcal{R} \upharpoonright \check{m}$ , obtained after updating  $\mathcal{R}$  with a sample  $\check{x}$  with count vector  $\check{m}$ . Then

$$\mathcal{S} \upharpoonright \check{m} = \{g \in \mathcal{G}(\mathcal{N}^{\hat{n}}): +_{\check{m}}(L_{\check{m}}g) \in \mathcal{S}\}. \quad (48)$$

**Example 8.** In the context of our running example, where, recall,  $\mathcal{X} := \{b, w\}$  and  $N := 2$ , let  $\hat{n} := 1$  and  $\check{x} := w$  with count vector  $\check{m} = (1, 0)$ . Take  $g$  to be some gamble on  $\mathcal{N}^{\hat{n}}$  and let  $g' := +_{\check{m}}(L_{\check{m}}g)$ , then

$$g'(2, 0) = g(1, 0), \quad g'(0, 2) = 0, \quad g'(1, 1) = \frac{1}{2}g(0, 1),$$

So  $g \in \mathcal{S} \mid \check{m}$  if  $g' \in \mathcal{S}$ . Contrast this with updating  $\mathcal{S}$  with the information that one of the two observations is  $w$ , i.e., with conditioning event  $\{\check{m}\} + \mathcal{N}^{\hat{n}} = \{(2, 0), (1, 1)\}$ . In that case  $g \in \mathcal{S} \mid (\{\check{m}\} + \mathcal{N}^{\hat{n}})$  if  $g'' \in \mathcal{S}$ , where

$$g''(2, 0) = g(1, 0), \quad g''(0, 2) = 0, \quad g''(1, 1) = g(0, 1). \quad \blacklozenge$$

**4.8. Multinomial processes.** Next, we turn to a number of important ideas related to multinomial processes. They are at the same time useful for comparisons with the existing literature, and essential for our treatment of countable exchangeable sequences in Section 5.

Consider the  $\mathcal{X}$ -simplex

$$\Sigma_{\mathcal{X}} := \left\{ \theta \in \mathbb{R}^{\mathcal{X}} : \theta \geq 0 \text{ and } \sum_{x \in \mathcal{X}} \theta_x = 1 \right\}. \quad (49)$$

and, for  $N \in \mathbb{N}_0$ , the linear map  $\text{CoMn}^N$  from  $\mathcal{G}(\mathcal{N}^N)$  to  $\mathcal{G}(\Sigma_{\mathcal{X}})$  defined by

$$\text{CoMn}^N : \mathcal{G}(\mathcal{N}^N) \rightarrow \mathcal{G}(\Sigma_{\mathcal{X}}) : g \mapsto \text{CoMn}^N(g) = \text{CoMn}^N(g|\cdot), \quad (50)$$

where for all  $\theta \in \Sigma_{\mathcal{X}}$ ,

$$\text{CoMn}^N(g|\theta) := \sum_{m \in \mathcal{N}^N} g(m) B_m(\theta) \quad (51)$$

is the expectation associated with the *count multinomial* distribution with parameters  $N$  and  $\theta$ , and  $B_m$  is the multivariate *Bernstein (basis) polynomial of degree  $N$*  given by

$$B_m(\theta) := \binom{N}{m} \prod_{z \in \mathcal{X}} \theta_z^{m_z} = |[m]| \prod_{z \in \mathcal{X}} \theta_z^{m_z}. \quad (52)$$

$\text{CoMn}^N(\{m\}|\theta) = B_m(\theta)$  is the probability of observing a count vector  $m$  in a multinomial process where the possible outcomes  $z \in \mathcal{X}$  have probability  $\theta_z$ .

We also consider the related linear map  $\text{Mn}^N$  from  $\mathcal{G}(\mathcal{X}^N)$  to  $\mathcal{G}(\Sigma_{\mathcal{X}})$  defined by

$$\text{Mn}^N : \mathcal{G}(\mathcal{X}^N) \rightarrow \mathcal{G}(\Sigma_{\mathcal{X}}) : f \mapsto \text{Mn}^N(f) = \text{Mn}^N(f|\cdot), \quad (53)$$

where for all  $\theta \in \Sigma_{\mathcal{X}}$ ,

$$\text{Mn}^N(f|\theta) := \sum_{m \in \mathcal{N}^N} \text{Hy}^N(f|m) B_m(\theta) \quad (54)$$

is the expectation associated with the *multinomial* distribution with parameters  $N$  and  $\theta$ . We then have that

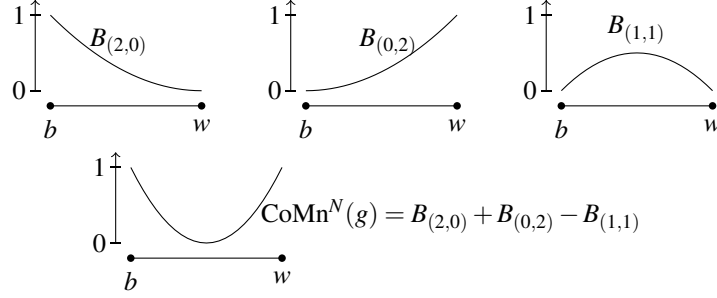
$$\text{CoMn}^N = \text{Mn}^N \circ \text{Co}^N \quad \text{and} \quad \text{Mn}^N = \text{CoMn}^N \circ \text{Hy}^N. \quad (55)$$

If we consider a sequence of observations  $x$  with count vector  $m$ , then  $\text{Mn}^N(\{x\}|\theta) = B_m(\theta)/|[m]| = \prod_{z \in \mathcal{X}} \theta_z^{m_z}$  is the probability of observing this sequence in a multinomial process where the possible outcomes  $z \in \mathcal{X}$  have probability  $\theta_z$ .

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  $N$ . 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  $\Sigma_{\mathcal{X}}$ . More details on Bernstein basis polynomials can be found in Appendix B.

**Example 9.** For our running example, the unit simplex  $\Sigma_{\mathcal{X}} = \Sigma_{\{b,w\}}$  is a line of unit length; frequency vectors  $(\theta_b, \theta_w)$  can be parametrised by  $\theta_b \in [0, 1]$ , as  $\theta_w = 1 - \theta_b$ . One of the line's extreme points corresponds to the frequency vector  $(1, 0)$ , 'b', the other to the frequency vector  $(0, 1)$ , 'w'.





The Bernstein basis polynomials for the case  $N = 2$  are given on the top row; on the bottom row, we give the polynomial  $\text{CoMn}^N(g)$  corresponding to the gamble  $g$  on  $\mathcal{N}^N$  defined by  $g(2,0) = g(0,2) = -g(1,1) = 1$ .  $\blacklozenge$

**4.9. Finite representation in terms of polynomials.** We see that the range of the linear maps  $\text{CoMn}^N$  and  $\text{Mn}^N$  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  $p = \text{CoMn}^N(b_p^N)$ ,  $\text{CoMn}^N$  is a *linear isomorphism* between the linear spaces  $\mathcal{G}(\mathcal{N}^N)$  and  $\mathcal{V}^N(\Sigma_{\mathcal{X}})$ . The relationships between the five important linear maps we have introduced so far are clarified by the commutative diagram in Fig. 1.

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*. Again, as explained above, the fundamental reason why this is possible, is that the Bernstein basis polynomials of degree  $N$  constitute a basis for the linear space of all polynomials of degree up to  $N$ , where a count gamble  $g$  plays the rôle of a coordinate representation for a polynomial  $p$  in this basis. The map  $\text{CoMn}^N$  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 5.

In order to lay the proper foundations for this work, we now prove a version of the finite representation theorem in terms of polynomial gambles of degree  $N$  on  $\Sigma_{\mathcal{X}}$ , rather than count gambles on  $\mathcal{N}^N$ .

**Definition 5** (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  $p, p_1, p_2 \in \mathcal{V}^N(\Sigma_{\mathcal{X}})$  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. This means that models in terms of sequences or count vectors are authoritative over those in terms of frequency vectors in the sense that polynomials are not directly behaviourally interpreted as gambles. This is related to the fact that not all possible frequency vectors can practically be observed.

**Example 10.** Any polynomial with a positive expansion in terms of Bernstein basis polynomials is positive. But the pair  $(g, \text{CoMn}^N(g))$  of Example 9 shows that a polynomial can be positive, while its Bernstein expansion is not. So the smallest set of polynomials Bernstein coherent at degree 2 is  $\text{posi}(\{B_{(2,0)}, B_{(0,2)}, B_{(1,1)}\})$ . Taking  $\text{CoMn}^N(g)$  to be desirable corresponds to the assessment that observing differing colors is less likely than observing identical ones.  $\blacklozenge$

Bernstein coherence at degree  $N$  is a special case of the general concept of coherence relative to  $(\mathcal{H}, \mathcal{C})$ , discussed in Section 2, where  $\mathcal{H} := \mathcal{V}^N(\Sigma_{\mathcal{X}})$  and  $\mathcal{C}$  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  $N$ :

$$\mathcal{C} := \{p \in \mathcal{V}^N(\Sigma_{\mathcal{X}}) : b_p^N \geq 0\}. \quad (56)$$

**Theorem 21** (Finite Representation). *A set of desirable gambles  $\mathcal{R}$  on  $\mathcal{X}^N$ , with count representation  $\mathcal{S} := \text{Hy}^N(\mathcal{R})$ , 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*

$$\mathcal{R} = (\text{Mn}^N)^{-1}(\mathcal{H}) \quad \text{or equivalently} \quad \mathcal{S} = (\text{CoMn}^N)^{-1}(\mathcal{H}), \quad (57)$$

and in that case this  $\mathcal{H}$  is uniquely determined by

$$\mathcal{H} = \text{Mn}^N(\mathcal{R}) = \text{CoMn}^N(\mathcal{S}). \quad (58)$$

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

## 5. COUNTABLE EXCHANGEABLE SEQUENCES

With the experience gained in investigating finite exchangeable sequences, we are now ready to tackle the problem of reasoning about countably infinite exchangeable sequences, our second main topic. The first step is to use the finite frequency representation results of Sections 4.8 and 4.9 to find a Representation Theorem for infinite exchangeable sequences (Section 5.1). We can then show what updating and natural extension look like in terms of this frequency representation (respectively Section 5.2 and Sections 5.4 and 5.5).

**5.1. 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)  $\mathcal{S}^n := \text{Hy}^n(\mathcal{R}^n)$  on  $\mathcal{X}^n$ , 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*, with  $n_1 \leq n_2$ :

$$\text{proj}_{n_2}^{n_1} : \mathcal{X}^{n_2} \rightarrow \mathcal{X}^{n_1} : (x_1, \dots, x_{n_2}) \mapsto \text{proj}_{n_2}^{n_1}(x_1, \dots, x_{n_2}) := (x_1, \dots, x_{n_1}). \quad (59)$$

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:

$$\text{ext}_{n_1}^{n_2} : \mathcal{G}(\mathcal{X}^{n_1}) \rightarrow \mathcal{G}(\mathcal{X}^{n_2}) : f \mapsto \text{ext}_{n_1}^{n_2}(f) = f \circ \text{proj}_{n_2}^{n_1}. \quad (60)$$

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:

$$(\forall n_1 \leq n_2) \text{ext}_{n_1}^{n_2}(\mathcal{R}^{n_1}) = \mathcal{R}^{n_2} \cap \text{ext}_{n_1}^{n_2}(\mathcal{G}(\mathcal{X}^{n_1})). \quad (61)$$

How can we translate this constraint in terms of the count representations  $\mathcal{S}^n$  or the frequency representations  $\mathcal{H}^n$ ? Using the Finite Representation Theorem 17, we see that

$f \in \mathcal{R}^{n_k} \Leftrightarrow \text{Hy}^{n_k}(f) \in \mathcal{S}^{n_k}$ . It follows from a few algebraic manipulations that for any gamble  $f$  on  $\mathcal{X}^{n_1}$  and all  $M \in \mathcal{N}^{n_2}$ :

$$\text{Hy}^{n_2}(\text{ext}_{n_1}^{n_2}(f)|M) = \sum_{m \in \mathcal{N}^{n_1}} \frac{|[M-m]| |[m]|}{|[M]|} \text{Hy}^{n_1}(f|m). \quad (62)$$

So 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  $\mathcal{G}(\mathcal{N}^{n_2})$  as follows:

$$\text{enl}_{n_1}^{n_2}: \mathcal{G}(\mathcal{N}^{n_1}) \rightarrow \mathcal{G}(\mathcal{N}^{n_2}): g \mapsto \text{enl}_{n_1}^{n_2}(g) := \sum_{m \in \mathcal{N}^{n_1}} \frac{|[\cdot - m]| |[m]|}{|[ \cdot ]|} g(m), \quad (63)$$

this can be summarised succinctly as:

$$\text{Hy}^{n_2} \circ \text{ext}_{n_1}^{n_2} = \text{enl}_{n_1}^{n_2} \circ \text{Hy}^{n_1}, \quad (64)$$

and we see that the time-consistency requirement (61) is then equivalent to [see Appendix A for a detailed proof]:

$$(\forall n_1 \leq n_2) \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1}) = \mathcal{S}^{n_2} \cap \text{enl}_{n_1}^{n_2}(\mathcal{G}(\mathcal{N}^{n_1})), \quad (65)$$

which is in turn equivalent to [see Appendix A for a detailed proof]:

$$(\forall n_1 \leq n_2) \mathcal{H}^{n_2} = \mathcal{H}^{n_1} \cap \mathcal{V}^{n_1}(\Sigma_{\mathcal{X}}). \quad (66)$$

We see that the time consistency condition can be most elegantly expressed in terms of the frequency representations.

**Example 11.** Let us illustrate the newly introduced operators in the context of our running example. Take  $n_1 := 1$ ,  $n_2 := 2$ , and  $f$  a gamble on  $\mathcal{X}^{n_1}$ ; let  $f' := \text{ext}_{n_1}^{n_2}(f)$ , then

$$f'(b, b) = f'(b, w) = f(b) \quad \text{and} \quad f'(w, w) = f'(w, b) = f(w).$$

Now take a gamble  $g$  on  $\mathcal{N}^{n_1}$ ; and let  $g' := \text{enl}_{n_1}^{n_2}(g)$ , then

$$g'(2, 0) = g(1, 0), \quad g'(0, 2) = g(0, 1), \quad g'(1, 1) = \frac{1}{2}(g(1, 0) + g(0, 1)). \quad \blacklozenge$$

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 Eq. (61).

The (count) multinomial expectations introduced in the previous section also satisfy a nice time consistency property. If we consider a gamble  $f_1$  on  $\mathcal{X}^{n_1}$ , then we can also consider it as a gamble  $\text{ext}_{n_1}^{n_2}(f_1)$  on  $\mathcal{X}^{n_2}$ , and of course both versions of this gamble should have the same multinomial expectation. This leads to the following identities:

$$\text{Mn}^{n_2} \circ \text{ext}_{n_1}^{n_2} = \text{Mn}^{n_1} \quad \text{and} \quad \text{CoMn}^{n_2} \circ \text{enl}_{n_1}^{n_2} = \text{CoMn}^{n_1}, \quad (67)$$

where the second identity follows from combining the first with Eqs. (64) and (55).

The relationships between three of the linear maps we encountered earlier and the maps related to time-consistency introduced here are clarified by the commutative diagrams in Fig. 3.

We can generalise the concept of Bernstein coherence given in Definition 5 to sets of polynomials of arbitrary degree:

**Definition 6** (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  $p, p_1, p_2 \in \mathcal{V}(\Sigma_{\mathcal{X}})$  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  $p_1 + p_2 \in \mathcal{H}$ .*

It is clear that we can replace B1 by the following requirement, because it is equivalent to it under B2–B4 [see Appendix A for a proof]:

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

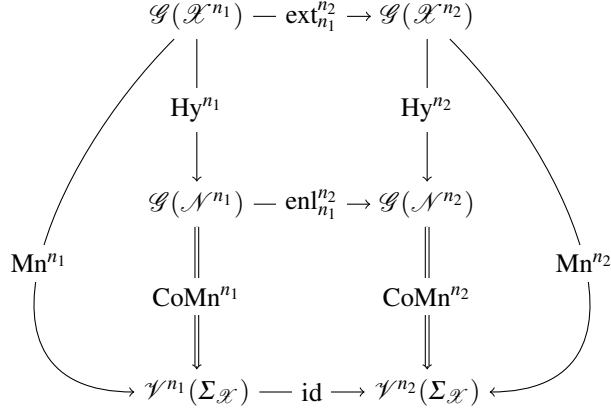


FIGURE 3. Different sequence length commutative diagram. Single arrows indicate linear monomorphisms (injective). Double arrows indicate linear isomorphisms (bijective).

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.

**Example 12.** The parabola  $\text{CoMn}^N(g)$  of Example 9 also shows that a polynomial can be positive, while no Bernstein expansion of any order is. This follows from the fact that all Bernstein basis polynomials are strictly positive on the interior of the unit simplex and that this parabola has a minimum of 0 within this interior.  $\blacklozenge$

Bernstein coherence is a special case of the general concept of coherence relative to  $(\mathcal{H}, \mathcal{C})$ , discussed in Section 2, where  $\mathcal{H} := \mathcal{V}(\Sigma_{\mathcal{X}})$  and  $\mathcal{C}$  is the convex cone of all polynomials with some non-negative Bernstein expansion:

$$\mathcal{C} := \mathcal{V}^+(\Sigma_{\mathcal{X}}) := \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : (\exists n \geq 0) b_p^n \geq 0\}. \quad (68)$$

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 22, which is a significant generalisation of de Finetti's representation result for countable sequences [9]. A similar result can also be proved for coherent lower previsions [8].

**Theorem 22 (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  $\mathcal{S}^n := \text{Hy}^n(\mathcal{R}^n)$  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  $n \in \mathbb{N}_0$ ,*

$$\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{H}) \quad \text{and} \quad \mathcal{R}^n = (\text{Mn}^n)^{-1}(\mathcal{H}), \quad (69)$$

and in that case this  $\mathcal{H}$  is uniquely given by

$$\mathcal{H} = \bigcup_{n \in \mathbb{N}_0} \mathcal{H}^n. \quad (70)$$

We call  $\mathcal{H}$  the frequency representation of the coherent, exchangeable and time-consistent family of sets of desirable gambles  $\mathcal{R}^n$ ,  $n \in \mathbb{N}_0$ .

**5.2. 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  $\mathcal{S}^n := \text{Hy}^n(\mathcal{R}^n)$  and associated 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, with associated count vector  $\check{m} := T^{\check{n}}(\check{x})$ , then we have seen in Section 4.7 that these models  $\mathcal{R}^n$  and  $\mathcal{S}^n$  (for  $n > \check{n}$ ) get updated to coherent and exchangeable models  $\mathcal{R}^{\hat{n}} \upharpoonright \check{m}$  with count representations  $\mathcal{S}^{\hat{n}} \upharpoonright \check{m} = \text{Hy}^{\hat{n}}(\mathcal{R}^{\hat{n}} \upharpoonright \check{m})$  for  $\hat{n} := n - \check{n} = 1, 2, \dots$ . It turns out that updating becomes especially easy in terms of the frequency representation.

**Theorem 23.** *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*

$$\mathcal{H} \upharpoonright \check{m} := \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : B_{\check{m}}p \in \mathcal{H}\}. \quad (71)$$

**5.3. Independence: iid sequences.** Theorem 23 can be used to find an easy and quite intriguing characterisation of a sequence of *independent and 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 beliefs about the remaining, unobserved ones. We infer from Theorem 23 that such will be the case iff the frequency representation  $\mathcal{H}$  of the sequence satisfies

$$(\forall \check{n} \in \mathbb{N}_0)(\forall \check{m} \in \mathcal{N}_{\mathcal{X}}^{\check{n}})\mathcal{H} \upharpoonright \check{m} = \mathcal{H}, \quad (72)$$

which is equivalent to

$$(\forall \check{n} \in \mathbb{N}_0)(\forall \check{m} \in \mathcal{N}_{\mathcal{X}}^{\check{n}})(\forall p \in \mathcal{V}(\Sigma_{\mathcal{X}}))(p \in \mathcal{H} \Leftrightarrow B_{\check{m}}p \in \mathcal{H}). \quad (73)$$

Any Bernstein coherent set of polynomials that satisfies one of the equivalent conditions (72) or (73) is an imprecise-probabilistic model for a (discrete-time) iid-process, or equivalently, a multinomial process, assuming values in a set  $\mathcal{X}$ .

Let us define  $e_z$  as the special count vector corresponding to a single observation of  $z \in \mathcal{X}$ : the  $z$ -component of  $e_z$  is one, and all other components are zero. Observe that  $B_{e_z}(\theta) = \theta_z$ . The precise-probabilistic iid-processes, or in other words, the multinomial processes, correspond to the *maximal* coherent sets of polynomials that satisfy the iid condition:

**Proposition 24.** *Consider any maximal element  $\mathcal{H}$  of  $\mathbb{D}_{\text{Be}}(\Sigma_{\mathcal{X}})$  that satisfies either of the equivalent conditions (72) or (73). Let  $\underline{P}_{\mathcal{H}}$  be the lower prevision defined on  $\mathcal{V}(\Sigma_{\mathcal{X}})$  in the usual way by letting  $\underline{P}_{\mathcal{H}}(p) := \sup\{\alpha : p - \alpha \in \mathcal{H}\}$  for all  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$ . Then  $\underline{P}_{\mathcal{H}}$  is a linear functional that dominates the min functional, and is completely determined by  $\underline{P}_{\mathcal{H}}(p) = p(\vartheta)$  for all  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$ , where  $\vartheta_z := \underline{P}_{\mathcal{H}}(B_{e_z})$  for all  $z \in \mathcal{X}$ . In addition, consider  $n \in \mathbb{N}$  and let  $\mathcal{S}^n := (\text{CoMn}^n)^{-1}(\mathcal{H})$  be the corresponding set of desirable count gambles on  $\mathcal{N}^n$ , with associated lower prevision  $\underline{P}_{\mathcal{S}^n}$ . Then  $\underline{P}_{\mathcal{S}^n}(g) = \underline{P}_{\mathcal{H}}(\text{CoMn}^n(g)) = \text{CoMn}^n(g|\vartheta) = \sum_{m \in \mathcal{N}^n} g(m)B_m(\vartheta)$  for all gambles  $g$  on  $\mathcal{N}^n$ , and in particular the probability of observing a count vector  $m \in \mathcal{N}^n$  is given by  $\underline{P}_{\mathcal{S}^n}(\{m\}) = B_m(\vartheta)$ .*

While it appears that such imprecise iid-processes are interesting, much more work needs to be done before we can get a complete picture of their structural properties and practical relevance. We leave this as a potential avenue for further research.

**5.4. 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  $\mathcal{H} := \mathcal{V}(\Sigma_{\mathcal{X}})$  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:

$$\mathcal{V}_0^+(\Sigma_{\mathcal{X}}) = \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : (\exists n \geq \deg(p))b_p^n > 0\}. \quad (74)$$

and by  $\mathcal{V}^-(\Sigma_{\mathcal{X}})$  the set of all polynomials on  $\Sigma_{\mathcal{X}}$  with some non-positive Bernstein expansion:

$$\mathcal{V}^-(\Sigma_{\mathcal{X}}) = \{p \in \mathcal{V}(\Sigma_{\mathcal{X}}) : (\exists n \geq \deg(p)) b_p^n \leq 0\}. \quad (75)$$

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.

$$\text{posi}(\mathcal{A}) \cap \mathcal{V}^-(\Sigma_{\mathcal{X}}) = \emptyset; \quad (76)$$

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 25** (Bernstein natural extension). *Consider a set of polynomials  $\mathcal{A} \subseteq \mathcal{V}(\Sigma_{\mathcal{X}})$ , and define its Bernstein natural extension*

$$\mathcal{E}_{\text{Be}}(\mathcal{A}) := \bigcap \{ \mathcal{H} \in \mathbb{D}_{\text{Be}}(\Sigma_{\mathcal{X}}) : \mathcal{A} \subseteq \mathcal{H} \}. \quad (77)$$

*Then the following statements are 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 that includes  $\mathcal{A}$ .

*When any (and hence all) of these equivalent statements hold, then*

$$\mathcal{E}_{\text{Be}}(\mathcal{A}) = \text{posi}(\mathcal{V}_0^+(\Sigma_{\mathcal{X}}) \cup \mathcal{A}). \quad (78)$$

**Example 13.** Recall that, for our running example, the unit simplex  $\Sigma_{\mathcal{X}} = \Sigma_{\{b,w\}}$  is a line of unit length parametrised by  $\theta_b \in [0, 1]$ , with  $\theta_w = 1 - \theta_b$ . Consider the polynomial  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$  such that  $p(\theta_b) = -1 + 3\theta_b - \theta_b^2$ . Because  $p(1) = 1$ , we have that  $\max p > 0$ , and thus via Eq. (121) that  $\max b_p^n > 0$  for all  $n \geq 2$ . So the assessment  $\{p\}$  avoids Bernstein non-positivity, because its Bernstein expansions of degree 2 and up are not non-positive.

**5.5. Exchangeable natural extension for infinite sequences.** To finish this discussion of exchangeability for infinite sequences of random variables, we take up the issue of inference, and extend the notion of exchangeable natural extension, discussed in Section 4.3, from finite to infinite sequences.

This extension is fairly straightforward. Suppose that for each  $i$  in the non-empty index set  $I$ , we consider a coherent, exchangeable and time-consistent family  $\mathcal{R}_i^n$ ,  $n \in \mathbb{N}_0$  of sets of desirable gambles. As we know from our Infinite Representation Theorem 22, each such family is represented by a Bernstein coherent set of polynomials on  $\Sigma_{\mathcal{X}}$ :

$$\mathcal{H}_i = \bigcup_{n \in \mathbb{N}_0} \text{Mn}^n(\mathcal{R}_i^n) \quad (79)$$

in the sense that, for all  $n \in \mathbb{N}_0$ ,

$$\mathcal{R}_i^n = (\text{Mn}^n)^{-1}(\mathcal{H}_i). \quad (80)$$

We know from the previous section that the intersection of a non-empty family of Bernstein coherent sets of polynomials is still Bernstein coherent. This implies that  $\mathcal{H} := \bigcap_{i \in I} \mathcal{H}_i$  is a Bernstein coherent set of polynomials such that, for all  $n \in \mathbb{N}_0$ ,

$$\mathcal{R}^n = \bigcap_{i \in I} \mathcal{R}_i^n = \bigcap_{i \in I} (\text{Mn}^n)^{-1}(\mathcal{H}_i) = (\text{Mn}^n)^{-1} \left( \bigcap_{i \in I} \mathcal{H}_i \right) = (\text{Mn}^n)^{-1}(\mathcal{H}), \quad (81)$$

implying that *the (element-wise) intersection  $\mathcal{R}^n$ ,  $n \in \mathbb{N}_0$  of the coherent, exchangeable and time-consistent families  $\mathcal{R}_i^n$ ,  $n \in \mathbb{N}_0$  is still a coherent, exchangeable and time-consistent family, whose frequency representation  $\mathcal{H}$  is the intersection of the frequency representations  $\mathcal{H}_i$ .*

Now suppose we have an assessment in the form 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  $n \in \mathbb{N}_0$ , which is equivalent to  $\text{Mn}^n(\mathcal{A}^n) \subseteq \text{Mn}^n(\mathcal{R}^n)$  for all  $n \in \mathbb{N}_0$ , which is in turn—because of Eq. (69)—equivalent to

$$\bigcup_{n \in \mathbb{N}_0} \text{Mn}^n(\mathcal{A}^n) \subseteq \bigcup_{n \in \mathbb{N}_0} \text{Mn}^n(\mathcal{R}^n) =: \mathcal{H}, \quad (82)$$

a condition formulated in terms of the frequency representation  $\mathcal{H}$  of the family  $\mathcal{R}^n$ ,  $n \in \mathbb{N}_0$ . The existence of this smallest family is implied by what we found in the previous paragraph. If we combine all this with the arguments in the previous section, we are led to the following theorem.

**Theorem 26.** *Suppose we have an assessment in the form of a set  $\mathcal{A}^n$  of desirable gambles on  $\mathcal{X}^n$  for each  $n \in \mathbb{N}_0$ , and consider the corresponding set of polynomials:*

$$\mathcal{A} := \bigcup_{n \in \mathbb{N}_0} \text{Mn}^n(\mathcal{A}^n). \quad (83)$$

*Then 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, and 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.*

## 6. EXTENDING FINITE EXCHANGEABLE SEQUENCES

Suppose we have  $n$  random variables  $X_1, \dots, X_n$ , that a subject judges to be exchangeable, and for which he has an assessment  $\mathcal{A}^n$  of desirable gambles on  $\mathcal{X}^n$ , with corresponding count representation  $\mathcal{S}^n = \text{Hy}^n(\mathcal{A}^n)$ . We here answer the question of when it is possible and how, if so, to extend such a sequence to a longer, finite or infinite sequence that is still exchangeable.

**6.1. Extension to a longer, finite exchangeable sequence.** In this section we ask: *Can the assessment  $\mathcal{A}^n$  be extended to a coherent exchangeable model for  $n+k$  variables? And if so, what is the most conservative such extended model?*

It is well-known [12] that when the subject's assessment is an exchangeable linear prevision, such an extension is not generally possible. In the much more general case that we are considering here, we now look at our Theorems 13 and 19 to provide us with an elegant answer: the extension problem considered here is a special case of the one studied in Section 4.3.

Indeed, since any gamble  $f$  on the first  $n$  variables  $X_1, \dots, X_n$  corresponds to the gamble  $\text{ext}_n^{n+k}(f)$  on the  $n+k$  variables  $X_1, \dots, X_n, \dots, X_{n+k}$ , we see that the assessment  $\mathcal{A}^n$  corresponds to an assessment

$$\mathcal{A}^{n+k} := \text{ext}_n^{n+k}(\mathcal{A}^n) = \left\{ \text{ext}_n^{n+k}(f) : f \in \mathcal{A}^n \right\} \quad (84)$$

of desirable gambles on  $\mathcal{X}^{n+k}$ . It is then clear from Theorem 13 that (i)  $\mathcal{A}^n$  can be extended to a coherent exchangeable model for  $n+k$  variables iff this  $\mathcal{A}^{n+k}$  avoids non-positivity under exchangeability; and if such is the case, that (ii) the smallest such coherent exchangeable extension is given by  $\mathcal{E}_{\text{ex}}^N(\mathcal{A}^{n+k})$ .

But we know from Theorem 19 that it is easier to express this in terms if the count representations. Since moreover, by Eq. (64),

$$\text{Hy}^{n+k}(\mathcal{A}^{n+k}) = \text{Hy}^{n+k}(\text{ext}_n^{n+k}(\mathcal{A}^n)) = \text{enl}_n^{n+k}(\text{Hy}^n(\mathcal{A}^n)) = \text{enl}_n^{n+k}(\mathcal{S}^n), \quad (85)$$

we are led to the following simple solution to the extension problem.

**Theorem 27.** Consider  $n$  and  $k$  in  $\mathbb{N}_0$ . An assessment  $\mathcal{A}^n$  of desirable gambles on  $\mathcal{X}^n$ , with corresponding count representation  $\mathcal{S}^n := \text{Hy}^n(\mathcal{A}^n)$ , can be extended to a coherent exchangeable model for  $n+k$  variables iff  $\text{enl}_n^{n+k}(\mathcal{S}^n)$  avoids non-positivity. In that case the most conservative such coherent exchangeable model has count representation  $\mathcal{E}(\text{enl}_n^{n+k}(\mathcal{S}^n))$ .

**Example 14.** In the context of our running example, take  $n = 2$  and consider the assessment  $\mathcal{A}^n := \{f\}$  where  $f$  is the gamble on  $\mathcal{X}^n$  given by

$$f(b, b) = f(w, w) := -3 \text{ and } f(b, w) = f(w, b) := 1,$$

whence

$$\mathcal{S}^n := \text{Hy}^n(\mathcal{A}^n) := \{g \in \mathcal{G}(\mathcal{N}^n) : g(2, 0) = g(0, 2) = -3 \text{ and } g(1, 1) = 1\}.$$

This singleton assessment avoids non-positivity under exchangeability and could be interpreted to express a strong belief that both colours will appear on the upcoming two draws, which could, e.g., be based on an observation of one black and one white marble being put in a seemingly empty urn. Now, let us see if this assessment can be extended to an exchangeable model for  $n+k := 3$  variables: Let  $g' := \text{enl}_n^{n+k}(g)$ , then

$$\begin{aligned} g'(3, 0) = g(2, 0) = -3, & \quad g'(2, 1) = \frac{1}{3}g(2, 0) + \frac{2}{3}g(1, 1) = -\frac{1}{3}, \\ g'(0, 3) = g(0, 2) = -3, & \quad g'(1, 2) = \frac{1}{3}g(0, 2) + \frac{2}{3}g(1, 1) = -\frac{1}{3}. \end{aligned}$$

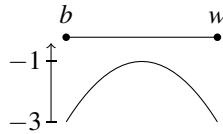
The gamble  $g'$  is clearly non-positive, so the assessment cannot be extended to a coherent exchangeable model. Learning that there are more marbles in the urn would force us to revise the initial assessment: making this assessment when there are (at least) three balls in the urn leads to a sure loss.  $\blacklozenge$

**6.2. Extension to an infinite exchangeable sequence.** Let us now extend the course of reasoning in the previous section to make it deal with infinite sequences of random variables. So in this section we ask: *Can the assessment  $\mathcal{A}^n$  be extended to a coherent exchangeable model for an infinite sequence of variables? And if so, what is the most conservative such extended model?*

Here, we look at Theorem 26 to provide us with an elegant answer: the present extension problem is a special case of that studied in Section 5.5. Indeed, the set of desirable gambles  $\mathcal{A}^n$  corresponds to an assessment of polynomials  $\text{Mn}^n(\mathcal{A}^n)$ , leading to the following simple solution to the extension problem.

**Theorem 28.** Consider  $n \in \mathbb{N}_0$ . An assessment  $\mathcal{A}^n$  of desirable gambles on  $\mathcal{X}^n$  can be extended to a coherent, exchangeable and time-consistent family iff  $\text{Mn}^n(\mathcal{A}^n)$  avoids Bernstein non-positivity. In that case the most conservative such family has frequency representation  $\mathcal{E}_{\text{Be}}(\text{Mn}^n(\mathcal{A}^n))$ .

**Example 15.** The singleton assessment  $\mathcal{A}^n$  of Example 14 can very quickly be seen to not be extendable to a coherent, exchangeable and time-consistent family, because the single polynomial in  $\text{Mn}^n(\mathcal{A}^n)$ —depicted below—is strictly negative, which by Proposition 37 assures us it incurs Bernstein non-positivity.



Had the parabola's top value been 0, this would not have been the case due to BP3.  $\blacklozenge$



## 7. CONCLUSIONS

We have shown that modelling a finite or infinite exchangeability assessment using sets of desirable gambles is not only possible, but also quite elegant. Our results indicate that, using sets of desirable gambles, it is conceptually easy to reason about exchangeable sequences.

Calculating the natural extension and updating are but simple geometrical operations: taking unions, sums and positive hulls and taking intersections, respectively. This approach has the added advantage that the exchangeability assessment is preserved under updating, also when the conditioning event has lower probability zero, which does not hold when using (lower) previsions (although this might be remedied by using full conditional measures, for which Cozman and Seidenfeld [2] give a good number of references).

Using our Finite Representation Theorem, reasoning about finite exchangeable sequences can be reduced to reasoning about count vectors or (polynomials of) frequency vectors. Working with these representations automatically guarantees that exchangeability is satisfied. The representation for the natural extension and for updated models can be derived directly from the representation of the original model, without having to go back to the (more complex) world of sequences.

Moreover, using our Infinite Representation Theorem, reasoning about infinite exchangeable sequences is reduced to reasoning about (polynomials of) frequency vectors. Doing this automatically guarantees that, next to the exchangeability of finite subsequences, time consistency of these subsequences is satisfied. Again, the representation for the natural extension and for updated models can be derived directly from the representation of the original model.

Additionally, using our results about representation and natural extension, we have shown when and how finite exchangeable sequences can be extended to longer, finite or infinite exchangeable sequences. However, we suspect there may be a more elegant characterisation of  $\mathcal{V}^-(\Sigma_{\mathcal{X}})$  than the one given above, which might make the characterisation in Theorem 28 more efficient to implement in terms of computer algorithms.

What are the advantages of our approach? It makes it easy for us to represent and reason with a finite number of expert assessments, and to see what its consequences are under exchangeability. Also, we have seen that there are simple geometrical representations and interpretations of coherence and exchangeability: due to the symmetry, the assessments can be represented in simpler, lower dimensional spaces, and there are linear maps effecting that representation.

The conceptual techniques employed in this paper are not restricted in use to a treatment of exchangeability. They could be applied to other structural assessments, e.g., invariance assessments, as long as this assessment allows us to identify a characterising set of weakly desirable gambles that is sufficiently well-behaved (cf. the first paragraph of Section 4.2). This idea was briefly taken up by one of us in another paper [4], but clearly merits further attention.

Thinking in even broader terms, we feel that using sets of desirable gambles can provide a refreshing and fruitful approach to many problems in uncertainty modelling, not only those related to structural assessments.

While writing this paper, we regularly wondered what Henry Kyburg would have thought about it. The topic surely has connections with his interests: exchangeability is an important basic assumption used in many models for statistical inference and our use of a model for uncertainty that is not just a precise probability, sets of desirable gambles. What we tried to do in this paper is in some sense clarify, in a very general setting, what the consequences are of an assessment of exchangeability. We know from his work that he thought it important for people to realise they are quite strong [18, p. 111ff., p. 122ff]. The preceding pages elaborately underline this point, and we agree it is an important one.

Although this paper sprouted from minds mildly seduced by subjectivist betting frameworks, nothing in it precludes using it objectively. We say this with a slightly mischievous

smile, mirroring a similar twinkle in Henry's eyes when we met him last, at a 2005 conference in Pittsburgh.

#### ACKNOWLEDGEMENTS

The authors wish to thank Teddy Seidenfeld for sharing some of the things he knows about Henry Kyburg and his work; they were both interesting and relevant. Erik Quaeghebeur was supported by a Fellowship of the Belgian American Educational Foundation and wishes to thank Carnegie Mellon University's Department of Philosophy for its hospitality.

#### REFERENCES

- [1] Inés Couso and Serafín Moral. Sets of desirable gambles and credal sets. In Thomas Augustin, Frank P. A. Coolen, Serafín Moral, and Matthias C. M. Troffaes, editors, *ISIPTA '09: Proceedings of the Sixth International Symposium on Imprecise Probabilities: Theories and Applications*, pages 99–108, Durham, United Kingdom, 2009. SIPTA. URL <http://www.sipta.org/isipta09/proceedings/063.html>.
- [2] Fabio Cozman and Teddy Seidenfeld. Independence for full conditional measures and their graphoid problems. In Benedikt Löwe, Eric Pacuit, and Jan-Willem Romeijn, editors, *Foundations of the Formal Sciences VI: Reasoning about Probabilities and Probabilistic Reasoning*, Studies in Logic. College Publications, 2009 (tentative). Conference held in Amsterdam, May 2-5, 2007.
- [3] Gert de Cooman. Belief models: an order-theoretic investigation. *Annals of Mathematics and Artificial Intelligence*, 45:5–34, 2005. doi: 10.1007/s10472-005-9006-x.
- [4] Gert de Cooman and Enrique Miranda. Symmetry of models versus models of symmetry. In W. L. Harper and G. R. Wheeler, editors, *Probability and Inference: Essays in Honor of Henry E. Kyburg, Jr.*, pages 67–149. King's College Publications, 2007.
- [5] Gert de Cooman and Erik Quaeghebeur. Exchangeability for sets of desirable gambles. In Thomas Augustin, Frank P. A. Coolen, Serafín Moral, and Matthias C. M. Troffaes, editors, *ISIPTA '09: Proceedings of the Sixth International Symposium on Imprecise Probabilities: Theories and Applications*, pages 159–168, Durham, United Kingdom, 2009. SIPTA. URL <http://hdl.handle.net/1854/LU-718913>.
- [6] Gert de Cooman and Erik Quaeghebeur. Infinite exchangeability for sets of desirable gambles. In Eyke Hüllermeier, Rudolf Kruse, and Frank Hoffmann, editors, *Communications in Computer and Information Science*, volume 80, pages 60–69. Springer Verlag Berlin, 2010. doi: 10.1007/978-3-642-14055-6\_7. URL <http://hdl.handle.net/1854/LU-984155>.
- [7] Gert de Cooman, Erik Quaeghebeur, and Enrique Miranda. Representing and assessing exchangeable lower previsions. In *Bulletin of the International Statistical Institute 56th Session – Proceedings*, number 1556, Lisboa, 2007. URL <http://hdl.handle.net/1854/8320>.
- [8] Gert de Cooman, Erik Quaeghebeur, and Enrique Miranda. Exchangeable lower previsions. *Bernoulli*, 15(3):721–735, 2009. doi: 10.3150/09-BEJ182. URL <http://hdl.handle.net/1854/LU-498518>.
- [9] Bruno de Finetti. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7:1–68, 1937. English translation in [19].
- [10] Bruno de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.
- [11] Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons, Chichester, 1974–1975. English translation of [10], two volumes.
- [12] Persi Diaconis and David A. Freedman. Finite exchangeable sequences. *The Annals of Probability*, 8:745–764, 1980.
- [13] David C. Heath and William D. Sudderth. De Finetti's theorem on exchangeable variables. *The American Statistician*, 30:188–189, 1976.

- [14] Edwin Hewitt and Leonard J. Savage. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80:470–501, 1955.
- [15] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan. *Discrete Multivariate Distributions*. Wiley Series in Probability and Statistics. John Wiley and Sons, New York, 1997.
- [16] Olav Kallenberg. *Foundations of Modern Probability*. Springer-Verlag, New York, second edition, 2002.
- [17] Olav Kallenberg. *Probabilistic Symmetries and Invariance Principles*. Springer, New York, 2005.
- [18] Henry E. Kyburg, Jr. *The Logical Foundations of Statistical Inference*, volume 65 of *Synthese Library*. D. Reidel Publishing Company, Dordrecht, Holland, 1974.
- [19] Henry E. Kyburg Jr. and Howard E. Smokler, editors. *Studies in Subjective Probability*. Wiley, New York, 1964. Second edition (with new material) 1980.
- [20] Serafín Moral. Epistemic irrelevance on sets of desirable gambles. In Gert de Cooman, Terrence L. Fine, and Teddy Seidenfeld, editors, *ISIPTA '01 – Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, pages 247–254. Shaker Publishing, Maastricht, 2000.
- [21] Serafín Moral. Epistemic irrelevance on sets of desirable gambles. *Annals of Mathematics and Artificial Intelligence*, 45:197–214, 2005.
- [22] Hartmut Prautzsch, Wolfgang Boehm, and Marco Paluszny. *Bézier and B-Spline Techniques*. Springer, Berlin, 2002.
- [23] Erik Quaeghebeur. *Learning from samples using coherent lower previsions*. PhD thesis, Ghent University, 2009. URL <http://hdl.handle.net/1854/LU-495650>.
- [24] Eric Schechter. *Handbook of Analysis and Its Foundations*. Academic Press, San Diego, CA, 1997.
- [25] Cedric A. B. Smith. Consistency in statistical inference and decision. *Journal of the Royal Statistical Society, Series A*, 23:1–37, 1961.
- [26] Wilfried Trump and Hartmut Prautzsch. Arbitrary degree elevation of Bézier representations. *Computer Aided Geometric Design*, 13:387–398, 1996.
- [27] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
- [28] Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 2000.
- [29] Peter M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Revised journal version: [32].
- [30] Peter M. Williams. Coherence, strict coherence and zero probabilities. In *Proceedings of the Fifth International Congress on Logic, Methodology and Philosophy of Science*, volume VI, pages 29–33. 1975.
- [31] Peter M. Williams. Indeterminate probabilities. In M. Przelecki, K. Szaniawski, and R. Wojcicki, editors, *Formal Methods in the Methodology of Empirical Sciences*, pages 229–246. Reidel, Dordrecht, 1976. Proceedings of a 1974 conference held in Warsaw.
- [32] Peter M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of [29].

#### APPENDIX A. PROOFS

**Proof of Theorem 1.** It follows from the fact that  $\mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega)$  is closed under arbitrary non-empty intersections, the definition of  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A})$ , and the fact that  $\mathcal{H}$  is not coherent relative to  $(\mathcal{H}, \mathcal{C})$  [because  $\mathcal{C} \subset \mathcal{H}$ ], that the last four statements are equivalent.

Next, we prove that (i)  $\Leftrightarrow$  (ii):

- $\Leftarrow$  Assume that  $\mathcal{A}$  is included in some set of desirable gambles  $\mathcal{R}$  that is coherent relative to  $(\mathcal{H}, \mathcal{C})$ . Since  $\mathcal{R} = \text{posi}(\mathcal{R})$ ,  $\mathcal{R}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$  by D5, and therefore so do all its subsets, including  $\mathcal{A}$ .
- $\Rightarrow$  Conversely, assume that  $\mathcal{A}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ . For notational convenience, let  $\mathcal{R}^* := \text{posi}(\mathcal{K}_{>0} \cup \mathcal{A})$ . It is clear that  $\mathcal{R}^*$  satisfies D2, D3 and D4. Consider any  $f \in \mathcal{R}^*$ , so there are  $n \geq 1$ , real  $\lambda_k > 0$ ,  $f_k \in \mathcal{K}_{>0} \cup \mathcal{A}$  such that  $f = \sum_{k=1}^n \lambda_k f_k$ . Let  $I := \{k \in \{1, \dots, n\} : f_k \succ 0\}$ , then  $f_\ell \in \mathcal{A}$  for all  $\ell \notin I$  and  $f = f_0 + \sum_{\ell \notin I} \lambda_\ell f_\ell$  with  $f_0 := \sum_{\ell \in I} \lambda_\ell f_\ell \succ 0$ . It therefore follows from the assumption that  $\sum_{\ell \notin I} \lambda_\ell f_\ell \not\leq 0$  and therefore *a fortiori*  $f \not\leq 0$ , so  $\mathcal{R}^*$  also satisfies D1 (or D5), and is therefore coherent relative to  $(\mathcal{H}, \mathcal{C})$ .

Finally, we prove that  $\mathcal{E}(\mathcal{A}) = \mathcal{R}^*$  whenever any (and hence all) of the equivalent statements (i)–(v) hold. Any coherent set of desirable gambles that includes  $\mathcal{A}$ , must also include  $\mathcal{R}^*$ , by the axioms D2, D3, and D4. Since we have proved above that  $\mathcal{R}^*$  also satisfies D1 and is therefore coherent relative to  $(\mathcal{H}, \mathcal{C})$ , it is the smallest set of desirable gambles that is coherent relative to  $(\mathcal{H}, \mathcal{C})$  and includes  $\mathcal{A}$ . Hence it is equal to  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A})$ , by (v).  $\square$

**Proof of Proposition 2.** We first prove sufficiency. Assume that Eq. (7) holds. Consider any  $\mathcal{R}' \in \mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega)$  such that  $\mathcal{R} \subseteq \mathcal{R}'$ , then we prove that also  $\mathcal{R}' \subseteq \mathcal{R}$ . To this effect, consider any  $f \in \mathcal{R}'$ , so  $-f \notin \mathcal{R}'$  by coherence, and therefore also  $-f \notin \mathcal{R}$ . Now invoke Eq. (7) to find that  $f \in \mathcal{R}$ .

Next, we turn to necessity. Assume that  $\mathcal{R}$  is maximal, consider any  $f \in \mathcal{K}_0$ , and assume that  $f \notin \mathcal{R}$ . We have to prove that  $-f \in \mathcal{R}$ . By Lemma 29, we get that  $\text{posi}(\mathcal{R} \cup \{-f\}) \in \mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega)$ , but since  $\mathcal{R} \subseteq \text{posi}(\mathcal{R} \cup \{-f\})$  and  $\mathcal{R}$  is maximal, we conclude that  $\mathcal{R} = \text{posi}(\mathcal{R} \cup \{-f\})$  and therefore indeed  $-f \in \mathcal{R}$ .  $\square$

**Lemma 29.** *Let  $\mathcal{H}$  be a linear subspace of  $\mathcal{G}(\Omega)$  and let  $\mathcal{C} \subset \mathcal{H}$  be a convex cone containing the zero gamble 0. Let  $\mathcal{R} \in \mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega)$  and let  $f$  be any non-zero gamble in  $\mathcal{H}$ . Then  $f \notin \mathcal{R}$  implies that  $\mathcal{R} \cup \{-f\}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ , and therefore  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{R} \cup \{-f\}) = \text{posi}(\mathcal{R} \cup \{-f\})$  is coherent relative to  $(\mathcal{H}, \mathcal{C})$ :*

$$(\forall f \in \mathcal{K}_0)(f \notin \mathcal{R} \Rightarrow \text{posi}(\mathcal{R} \cup \{-f\}) \in \mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega)). \quad (86)$$

**Proof of Lemma 29.** We give a proof by contradiction. Let  $f \in \mathcal{K}_0 \setminus \mathcal{R}$  and assume that  $\mathcal{R} \cup \{-f\}$  does not avoid non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ . This means that  $\text{posi}(\mathcal{R} \cup \{-f\}) \cap \mathcal{K}_{\leq 0} \neq \emptyset$ , and since  $\mathcal{R}$  does avoid non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ , this tells us that there are  $n \in \mathbb{N}_0$ ,  $f_1, \dots, f_n$  in  $\mathcal{R}$ ,  $\lambda$  in  $\mathbb{R}_0^+$ , and  $\lambda_1, \dots, \lambda_n$  in  $\mathbb{R}^+$  such that

$$\sum_{k=1}^n \lambda_k f_k + \lambda(-f) \leq 0 \text{ and therefore } f \succeq \sum_{k=1}^n \frac{\lambda_k}{\lambda} f_k. \quad (87)$$

Then obviously  $f \in \mathcal{R}$  since  $f \neq 0$ , a contradiction. The rest of the proof now follows from Theorem 1 and  $\mathcal{K}_{>0} \subseteq \mathcal{R}$ .  $\square$

**Proof of Theorem 3.** Sufficiency follows readily from Theorem 1.

For necessity, assume that  $\mathcal{A}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ , and consider the set  $\uparrow \mathcal{A} := \{\mathcal{R} \in \mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega) : \mathcal{A} \subseteq \mathcal{R}\}$ . This set is non-empty by Theorem 1, and partially ordered by set inclusion. We show that this poset has a maximal element, which is then automatically also a maximal element of  $\mathbb{D}_{(\mathcal{H}, \mathcal{C})}(\Omega)$ .

Consider any chain  $\mathbb{K} \subseteq \uparrow \mathcal{A}$ . We show that  $\bigcup \mathbb{K}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ . Consider arbitrary  $n \in \mathbb{N}_0$ ,  $f_1, \dots, f_n$  in  $\bigcup \mathbb{K}$ .  $f_k \in \bigcup \mathbb{K}$  means that there is some  $\mathcal{R}_k \in \mathbb{K}$  such that  $f_k \in \mathcal{R}_k$ , and therefore  $\{f_1, \dots, f_n\} \subseteq \bigcup_{k=1}^n \mathcal{R}_k =: \tilde{\mathcal{R}}$ . But  $\tilde{\mathcal{R}} \in \mathbb{K}$  because  $\mathbb{K}$  is a chain, and therefore  $\tilde{\mathcal{R}}$  is coherent relative to  $(\mathcal{H}, \mathcal{C})$ . This implies that  $\mathcal{K}_{\leq 0} \cap \tilde{\mathcal{R}} = \emptyset$ , and therefore *a fortiori*  $\mathcal{K}_{\leq 0} \cap \text{posi}(\{f_1, \dots, f_n\}) = \emptyset$ . So we find that  $\bigcup \mathbb{K}$  indeed avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ .

By Theorem 1,  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\bigcup \mathbb{K}) = \text{posi}(\bigcup \mathbb{K})$  is coherent relative to  $(\mathcal{H}, \mathcal{C})$  and includes  $\bigcup \mathbb{K}$ , so  $\mathcal{R} \subseteq \text{posi}(\bigcup \mathbb{K})$  for all  $\mathcal{R} \in \mathbb{K}$ . Because also  $\mathcal{A} \subseteq \text{posi}(\bigcup \mathbb{K})$ , we have just shown

that every chain  $\mathbb{K}$  in the poset  $\uparrow\mathcal{A}$  has an upper bound  $\text{posi}(\bigcup\mathbb{K})$  in  $\uparrow\mathcal{A}$ . By Zorn's Lemma,  $\uparrow\mathcal{A}$  has a maximal element.  $\square$

**Proof of Corollary 4.** We use the notation  $\mathbb{M}^* := \{\mathcal{R} \in \mathbb{M}_{(\mathcal{H}, \mathcal{C})}(\Omega) : \mathcal{A} \subseteq \mathcal{R}\}$  for the sake of brevity.

If  $\mathcal{A}$  does not avoid non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ , then  $\mathbb{M}^* = \emptyset$ , by Theorem 1, so  $\bigcap\mathbb{M}^* = \mathcal{H}$ . Again by Theorem 1, also  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A}) = \mathcal{H}$ .

Assume, therefore, that  $\mathcal{A}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ . Then  $\mathbb{M}^* \neq \emptyset$  by Theorem 3. Since  $\bigcap\mathbb{M}^*$  is coherent relative to  $(\mathcal{H}, \mathcal{C})$  and  $\mathcal{A} \subseteq \bigcap\mathbb{M}^*$ , we infer from Theorem 1 that  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A}) \subseteq \bigcap\mathbb{M}^*$ . Assume *ex absurdo* that  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A}) \subset \bigcap\mathbb{M}^*$ , so there is some (non-zero)  $f \in \bigcap\mathbb{M}^*$  such that  $f \notin \mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A})$ , and therefore  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A}) \cup \{-f\}$  avoids non-positivity relative to  $(\mathcal{H}, \mathcal{C})$ , by Lemma 29. By Theorem 3, there is some  $\mathcal{R}^*$  in  $\mathbb{M}_{(\mathcal{H}, \mathcal{C})}(\Omega)$  such that  $\mathcal{E}_{(\mathcal{H}, \mathcal{C})}(\mathcal{A}) \cup \{-f\} \subseteq \mathcal{R}^*$ . On the one hand, we infer that  $\mathcal{A} \subseteq \mathcal{R}^*$ , so  $\mathcal{R}^* \in \mathbb{M}^*$ , and therefore  $f \in \mathcal{R}^*$ . On the other hand, we infer that  $-f \in \mathcal{R}^*$ , which contradicts  $f \in \mathcal{R}^*$ , since  $\mathcal{R}^*$  is coherent.  $\square$

**Proof of Proposition 5.** The defining property of any gamble  $f$  in  $\mathcal{D}_{\mathcal{R}}$  is that  $f + f' \in \mathcal{R}$  for all gambles  $f'$  in  $\mathcal{R}$ .

- WD1. Let  $f < 0$ ; let  $f' = -f/2$  then  $f' > 0$  and therefore  $f' \in \mathcal{R}$ , by D1. But  $f'' = f + f' = f/2 < 0$  and thus, by D5,  $f'' \notin \mathcal{R}$ . Hence indeed  $f \notin \mathcal{D}_{\mathcal{R}}$ .
- WD2. Since clearly  $0 + \mathcal{R} = \mathcal{R}$ , we see that  $0 \in \mathcal{D}_{\mathcal{R}}$ . If  $f > 0$  then  $f \in \mathcal{R}$  by D2, and therefore  $f \in \mathcal{D}_{\mathcal{R}}$ , because  $\mathcal{R} \subseteq \mathcal{D}_{\mathcal{R}}$ .
- WD3. Consider  $f \in \mathcal{D}_{\mathcal{R}}$ . If  $\lambda = 0$  then  $\lambda f = 0 \in \mathcal{D}_{\mathcal{R}}$  by WD2. Assume therefore that  $\lambda > 0$ . Consider any  $f' \in \mathcal{R}$ . Then  $f'/\lambda \in \mathcal{R}$  by D3, so  $f + f'/\lambda \in \mathcal{R}$ , and therefore  $\lambda f + f' \in \mathcal{R}$ , again by D3. Hence indeed  $\lambda f \in \mathcal{D}_{\mathcal{R}}$ .
- WD4. Consider  $f_1, f_2 \in \mathcal{D}_{\mathcal{R}}$ , and any  $f' \in \mathcal{R}$ , so  $f'/2 \in \mathcal{R}$  by D3. Then  $f_1 + f'/2 \in \mathcal{R}$  and  $f_2 + f'/2 \in \mathcal{R}$ , and therefore  $f_1 + f_2 + f' \in \mathcal{R}$ , by D4. Hence indeed  $f_1 + f_2 \in \mathcal{D}_{\mathcal{R}}$ .  $\square$

**Proof of Theorem 6.** To prove that  $\underline{P}_{\mathcal{R}}$  is real-valued, we prove that  $\underline{P}_{\mathcal{R}}(f)$  is bounded for all gambles  $f \in \mathcal{G}(\Omega)$ —which are bounded and real-valued by definition. It follows from D5 that if  $f - \mu \in \mathcal{R}$ , then  $f \not\leq \mu$ , so  $\sup f > \mu$ , whence  $\underline{P}_{\mathcal{R}}(f) \leq \sup f < +\infty$ . It follows from D2 that  $f - \mu \in \mathcal{R}$  if  $f - \mu > 0$ ; let  $\mu$  be any real number such that  $\mu < \inf f$ , then  $f - \mu > f - \inf f \geq 0$ , so  $f - \mu \in \mathcal{R}$ , whence  $\underline{P}_{\mathcal{R}}(f) \geq \inf f > \mu > -\infty$ .

To prove the equality of  $\underline{P}_{\mathcal{R}}$  and  $\underline{P}_{\mathcal{D}_{\mathcal{R}}}$ , consider any gamble  $f \in \mathcal{G}(\Omega)$ . Since  $\mathcal{R} \subseteq \mathcal{D}_{\mathcal{R}}$ , we immediately get that

$$\{\mu \in \mathbb{R} : f - \mu \in \mathcal{R}\} \subseteq \{\mu \in \mathbb{R} : f - \mu \in \mathcal{D}_{\mathcal{R}}\} \quad (88)$$

and therefore  $\underline{P}_{\mathcal{R}}(f) \leq \underline{P}_{\mathcal{D}_{\mathcal{R}}}(f)$ . Conversely, consider any  $\alpha > 0$ , then  $\alpha \in \mathcal{R}$  by coherence [D2], and therefore

$$\{\mu \in \mathbb{R} : f - \mu \in \mathcal{D}_{\mathcal{R}}\} \subseteq \{\mu : f - \mu + \alpha \in \mathcal{R}\} \quad (89)$$

$$= \alpha + \{\mu : f - \mu \in \mathcal{R}\}, \quad (90)$$

whence  $\underline{P}_{\mathcal{D}_{\mathcal{R}}}(f) \leq \alpha + \underline{P}_{\mathcal{R}}(f)$ . Since this holds for all  $\alpha > 0$ , we also have  $\underline{P}_{\mathcal{D}_{\mathcal{R}}}(f) \leq \underline{P}_{\mathcal{R}}(f)$ .

Next, consider any  $f \in \mathcal{D}_{\mathcal{R}}$ . Because  $f = f - 0$  this tells us that  $\underline{P}_{\mathcal{R}}(f) = \underline{P}_{\mathcal{D}_{\mathcal{R}}}(f) \geq 0$ .

The rest of the proof is now standard, see for instance [28, Section 6].  $\square$

**Proof of Proposition 7.** Since it follows from Theorem 6 that  $\underline{P}_{\mathcal{R}}(f - \underline{P}_{\mathcal{R}}(f)) = \underline{P}_{\mathcal{R}}(f) - \underline{P}_{\mathcal{R}}(f) = 0$  for all gambles  $f$ , it follows that  $\mathcal{M}_{\mathcal{R}} \subseteq \{f \in \mathcal{G}(\Omega) : \underline{P}_{\mathcal{R}}(f) = 0\}$ . For the converse inequality, assume that  $\underline{P}_{\mathcal{R}}(f) = 0$  holds; then  $f = f - \underline{P}_{\mathcal{R}}(f) \in \mathcal{M}_{\mathcal{R}}$ .

This also means that  $\underline{P}_{\mathcal{R}}(g) = 0$  iff  $g \in \mathcal{M}_{\mathcal{R}}$ , so for every gamble  $f$  we can write:

$$\underline{P}_{\mathcal{M}_{\mathcal{R}}}(f) = \sup \{\mu \in \mathbb{R} : f - \mu \in \mathcal{M}_{\mathcal{R}}\} \quad (91)$$

$$= \sup \{\mu \in \mathbb{R} : \underline{P}_{\mathcal{R}}(f - \mu) = 0\} \quad (92)$$

$$= \sup \{ \mu \in \mathbb{R} : \mu = \underline{P}_{\mathcal{R}}(f) \} = \underline{P}_{\mathcal{R}}(f), \quad (93)$$

which proves the equality of  $\underline{P}_{\mathcal{M}_{\mathcal{R}}}$  and  $\underline{P}_{\mathcal{R}}$ .  $\square$

**Proof of Proposition 8.** We need to prove that the appropriate versions of D1–D4 hold for  $\mathcal{R}|B$ , with  $\mathcal{K} = \mathcal{G}(\Omega)|B$  and  $\mathcal{C} = \mathcal{G}(\Omega)|B \cap \mathcal{G}^+(\Omega)$ . For D1, consider  $f \in \mathcal{G}(\Omega)|B$  and assume that  $f = 0$ . Then by coherence  $f \notin \mathcal{R}$  and hence  $f \notin \mathcal{R}|B$ . For D2, consider  $f \in \mathcal{G}(\Omega)|B$  and assume that  $f > 0$ . Then by coherence  $f \in \mathcal{R}$  and hence  $f \in \mathcal{R}|B$ . The proof for D3 is similar to the one for D4. For D4, consider  $f_1, f_2 \in \mathcal{R}|B$ , then on the one hand  $f_1, f_2 \in \mathcal{R}$  and therefore  $f_1 + f_2 \in \mathcal{R}$  by coherence; and on the other hand  $f_1, f_2 \in \mathcal{G}(\Omega)|B$  and therefore  $f_1 + f_2 = I_B f_1 + I_B f_2 = I_B(f_1 + f_2)$ , so  $f_1 + f_2 \in \mathcal{G}(\Omega)|B$  and hence  $f_1 + f_2 \in \mathcal{R}|B$ .  $\square$

**Proof of the equivalences in Definition 3.** That (i) $\Leftrightarrow$ (ii) and (iii) $\Leftrightarrow$ (iv) is an immediate consequence of the definition of weak desirability. We show that (i) $\Leftrightarrow$ (iii). For the ‘ $\Rightarrow$ ’ part, observe that  $f - \text{ex}^N(f) = \frac{1}{N!} \sum_{\pi \in \mathcal{D}_N} [f - \pi^t f] \in \mathcal{D}_{\mathcal{R}}$ , since  $\mathcal{D}_{\mathcal{R}}$  is a convex cone by Proposition 5. The ‘ $\Leftarrow$ ’ part follows from  $\mathcal{D}_{\mathcal{R}} \subseteq \mathcal{D}_{\mathcal{M}_N}$ , i.e., from Eq. (27).  $\square$

**Proof of Proposition 9.** Consider  $f \in \mathcal{R}$ . Since  $\pi^t f - f = (-f) - \pi^t(-f) \in \mathcal{D}_{\mathcal{M}_N}$ , we see that  $\pi^t f = f + \pi^t f - f \in \mathcal{R} + \mathcal{D}_{\mathcal{M}_N} \subseteq \mathcal{R}$ , using the exchangeability condition of Definition 3(ii).  $\square$

**Proof of Proposition 10.** The first statement is a consequence of the second, with  $f' = \text{ex}^N(f)$ , because then  $\text{ex}^N(f') = \text{ex}^N(\text{ex}^N(f)) = \text{ex}^N(f)$ . For the second statement, consider arbitrary gambles  $f$  and  $f'$  on  $\mathcal{X}^N$  such that  $\text{ex}^N(f) = \text{ex}^N(f')$ , and assume that  $f \in \mathcal{R}$ . We prove that then also  $f' \in \mathcal{R}$ . Since  $\text{ex}^N(f) - f = (-f) - \text{ex}^N(-f) \in \mathcal{D}_{\mathcal{R}}$  and  $f' - \text{ex}^N(f') \in \mathcal{D}_{\mathcal{R}}$ , we see that  $f' - f \in \mathcal{D}_{\mathcal{R}}$  by WD4, and therefore  $f' = f + f' - f \in \mathcal{R} + \mathcal{D}_{\mathcal{R}} \subseteq \mathcal{R}$ .  $\square$

**Proof of Theorem 11.** We give a circular proof. We first show that (ii) holds if  $\underline{P}$  is exchangeable, i.e., if there is some coherent and exchangeable  $\mathcal{R}$  such that  $\underline{P} = \underline{P}_{\mathcal{R}}$ . We already know from Theorem 6 that  $\underline{P} = \underline{P}_{\mathcal{R}}$  satisfies P1–P3, because  $\mathcal{R}$  is coherent. Consider any  $f \in \mathcal{D}_{\mathcal{M}_N}$ . Since  $\mathcal{D}_{\mathcal{M}_N} \subseteq \mathcal{D}_{\mathcal{R}}$ , it also follows from Theorem 6 that  $\underline{P}_{\mathcal{R}}(f) \geq 0$  and similarly  $-\overline{P}_{\mathcal{R}}(f) = \underline{P}_{\mathcal{R}}(-f) \geq 0$  because also  $-f \in \mathcal{D}_{\mathcal{M}_N}$ . Hence indeed  $0 \leq \underline{P}_{\mathcal{R}}(f) \leq \overline{P}_{\mathcal{R}}(f) \leq 0$ , where the second inequality is a consequence of P1 and P2.

That (ii) implies (iii) follows the super-additivity of  $\underline{P}$  and the sub-additivity of  $\overline{P}$ .

Finally, we show that (iii) implies that  $\underline{P}$  is exchangeable. The standard argument in [28, Section 6] tells us that  $\mathcal{R}' := \{f \in \mathcal{G}(\mathcal{X}^N) : f > 0 \text{ or } \underline{P}(f) > 0\}$  is a coherent set of desirable gambles such that  $\underline{P}_{\mathcal{R}'} = \underline{P}$ . Now consider the set  $\mathcal{R} := \mathcal{R}' + \mathcal{D}_{\mathcal{M}_N}$ . We show that this  $\mathcal{R}$  is a coherent and exchangeable set of desirable gambles, and that  $\underline{P}_{\mathcal{R}} = \underline{P}$ . It is clear from its definition that  $\mathcal{R}$  satisfies D2, D3 and D4, so let us assume *ex absurdo* that  $0 \in \mathcal{R}$ , meaning that there is some  $f \in \mathcal{R}'$  such that  $f' := -f \in \mathcal{D}_{\mathcal{M}_N}$ . There are two possibilities. Either  $f > 0$ , so  $f' < 0$ , which contradicts Lemma 30. Or  $\underline{P}(f) > 0$ . But it follows from (iii) and the coherence of the lower prevision  $\underline{P}$  that  $0 = \underline{P}(f + f') = \underline{P}(f) > 0$ , a contradiction too. So  $\mathcal{R}$  satisfies D1 as well, and is therefore coherent. It is obvious that  $\mathcal{R}$  is exchangeable:  $\mathcal{R} + \mathcal{D}_{\mathcal{M}_N} = \mathcal{R}' + \mathcal{D}_{\mathcal{M}_N} + \mathcal{D}_{\mathcal{M}_N} = \mathcal{R}' + \mathcal{D}_{\mathcal{M}_N} = \mathcal{R}$ . The proof is complete if we can show that  $\underline{P} = \underline{P}_{\mathcal{R}}$ . Fix any gamble  $f$ . Observe that  $f - \alpha \in \mathcal{R}$  iff there are  $f' \in \mathcal{R}'$  and  $f'' \in \mathcal{D}_{\mathcal{M}_N}$  such that  $f - \alpha = f' + f''$ . But then it follows from the coherence of  $\underline{P}$  and the assumption that  $\underline{P}(f) = \alpha + \underline{P}(f' + f'') = \alpha + \underline{P}(f') \geq \alpha$ , and therefore  $\underline{P}_{\mathcal{R}}(f) \leq \underline{P}(f) = \underline{P}_{\mathcal{R}'}(f)$ . For the converse inequality, we infer from  $0 \in \mathcal{D}_{\mathcal{M}_N}$  that  $\mathcal{R}' \subseteq \mathcal{R}$ , and therefore  $\underline{P}_{\mathcal{R}'} \leq \underline{P}_{\mathcal{R}}$ .  $\square$

**Lemma 30.** For all  $f$  in  $\mathcal{D}_{\mathcal{M}_N}$ ,  $f \not\prec 0$ .

*Proof.* First of all, observe that for any gamble  $f'$  on  $\mathcal{X}^N$ , if  $f' > 0$  then also  $\text{ex}^N(f') > 0$ . Now consider  $f \in \mathcal{D}_{\mathcal{M}_N}$  and assume *ex absurdo* that  $f < 0$ . Then  $-f > 0$  and therefore

$-\text{ex}^N(f) = \text{ex}^N(-f) > 0$ , whence  $\text{ex}^N(f) < 0$ . But since  $f \in \mathcal{D}_{\mathcal{M}_N}$  we also have that  $\text{ex}^N(f) = 0$ , a contradiction.  $\square$

**Proof of Proposition 12.** For the first statement, we have to prove that  $\mathcal{G}_0^+(\mathcal{X}^N) + \mathcal{D}_{\mathcal{M}_N}$  avoids non-positivity. Consider any  $f' \in \mathcal{D}_{\mathcal{M}_N}$  and any  $f'' \in \mathcal{G}_0^+(\mathcal{X}^N)$ , then we have to prove that  $f := f' + f'' \not\leq 0$ . There are two possibilities. Either  $f' = 0$  and then  $f = f'' > 0$ . Or  $f' \neq 0$ , and then Lemma 30 tells us that  $f' \not\leq 0$  and therefore *a fortiori*  $f \not\leq 0$ .

For the second statement, it clearly suffices to prove the ‘if’ part. Assume therefore that  $\mathcal{A} + \mathcal{D}_{\mathcal{M}_N}$  avoids non-positivity. Consider any  $f$  in  $\text{posi}([\mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A}] + \mathcal{D}_{\mathcal{M}_N})$ , so there are  $n \geq 1$ ,  $\lambda_k \in \mathbb{R}_0^+$ ,  $f' \in \mathcal{D}_{\mathcal{M}_N}$ ,  $f_k \in \mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A}$  such that  $f = f' + \sum_{k=1}^n \lambda_k f_k$ . Let  $I := \{k \in \{1, \dots, n\} : f_k > 0\}$  then  $f_\ell \in \mathcal{A}$  for all  $\ell \notin I$ , and  $f = f_0 + f' + \sum_{\ell \notin I} \lambda_\ell f_\ell$  with  $f_0 > 0$ . By assumption  $f' + \sum_{\ell \notin I} \lambda_\ell f_\ell \not\leq 0$ , and therefore *a fortiori*  $f \not\leq 0$ .  $\square$

**Proof of Theorem 13.** It is immediately clear from the fact that  $\mathbb{D}_{\text{ex}}(\mathcal{X}^N)$  is closed under arbitrary non-empty intersections, the definition of  $\mathcal{E}_{\text{ex}}^N(\mathcal{A})$ , and the fact that  $\mathcal{G}(\mathcal{X}^N)$  is not a coherent set of desirable gambles, that the last four statements are equivalent.

Next, we prove that (i)  $\Leftrightarrow$  (ii).

$\Leftarrow$  Assume that  $\mathcal{A}$ , and therefore also  $\mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A}$ , is included in some coherent and exchangeable set of desirable gambles  $\mathcal{R}$ . By exchangeability, we know  $[\mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A}] + \mathcal{D}_{\mathcal{M}_N} \subseteq \mathcal{R} + \mathcal{D}_{\mathcal{M}_N} \subseteq \mathcal{R}$ . Since  $\text{posi}(\mathcal{R}) = \mathcal{R}$  avoids non-positivity, so does any of its subsets, and therefore in particular  $[\mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A}] + \mathcal{D}_{\mathcal{M}_N}$ . This means that  $\mathcal{A}$  indeed avoids non-positivity under exchangeability.

$\Rightarrow$  Conversely, assume that  $\mathcal{A}$  avoids non-positivity under exchangeability. For the sake of convenience, denote the set on the right-hand side of Eq. (33) by  $\mathcal{R}^*$ . It is clear that  $\mathcal{R}^*$  satisfies D2, D3 and D4. Consider any  $f \in \mathcal{R}^*$ , then  $f \not\leq 0$ , precisely because  $\mathcal{A}$  avoids non-positivity under exchangeability. Hence  $\mathcal{R}^*$  also satisfies D1, and is therefore coherent. The exchangeability of  $\mathcal{R}^*$  immediately follows from the fact that  $\mathcal{D}_{\mathcal{M}_N} + \mathcal{E}(\mathcal{A}) + \mathcal{D}_{\mathcal{M}_N} = \mathcal{D}_{\mathcal{M}_N} + \mathcal{E}(\mathcal{A})$ .

Finally, we prove Eqs. (33) and (34) whenever any (and hence all) of the equivalent statements (i)–(v) holds. Eq. (34) follows from Eq. (33) and Theorem 1, since  $\mathcal{D}_{\mathcal{M}_N}$  is a convex cone. Let us prove that  $\mathcal{E}_{\text{ex}}^N(\mathcal{A}) = \mathcal{R}^*$ . It is clear that any coherent and exchangeable set of desirable gambles that includes  $\mathcal{A}$ , must also include  $\mathcal{R}^*$ , by the axioms D2, D3, and D4. Since we have just proved above that  $\mathcal{R}^*$  is coherent and exchangeable, it is the smallest coherent and exchangeable set of desirable gambles that includes  $\mathcal{A}$ , and for this reason it is equal to  $\mathcal{E}_{\text{ex}}^N(\mathcal{A})$ , by (v).  $\square$

**Proof of Corollary 14.** This is an immediate consequence of Proposition 12(i) and Theorem 13.  $\square$

**Proof of Proposition 15.** The coherence of  $\mathcal{R} \upharpoonright \check{x}$  is guaranteed by Proposition 8. We show that  $\mathcal{R} \upharpoonright \check{x}$  is exchangeable. Consider arbitrary  $f \in \mathcal{G}(\mathcal{X}^{\check{n}})$ ,  $\hat{\pi} \in \mathcal{P}_{\check{n}}$  and  $f_1 \in \mathcal{R} \upharpoonright \check{x}$ . Then we must show that  $f_1 + f - \hat{\pi}^t f \in \mathcal{R} \upharpoonright \check{x}$ , or in other words that  $I_{C_{\check{x}}}[f_1 + f - \hat{\pi}^t f] \in \mathcal{R}$ . But since  $f_1 \in \mathcal{R} \upharpoonright \check{x}$ , we know that  $I_{C_{\check{x}}} f_1 \in \mathcal{R}$ . And if we consider the permutation  $\pi \in \mathcal{P}_N$  defined by

$$\pi(k) := \begin{cases} k & 1 \leq k \leq \check{n} \\ \check{n} + \hat{\pi}(k - \check{n}) & \check{n} + 1 \leq k \leq N, \end{cases} \quad (94)$$

then clearly  $I_{C_{\check{x}}} \hat{\pi}^t f = \pi^t(I_{C_{\check{x}}} f)$  and therefore  $I_{C_{\check{x}}}[f_1 + f - \hat{\pi}^t f] = I_{C_{\check{x}}} f_1 + I_{C_{\check{x}}} f - \pi^t(I_{C_{\check{x}}} f)$  and this gamble belongs to  $\mathcal{R}$  because  $\mathcal{R}$  is exchangeable.  $\square$

**Proof of Proposition 16.** Consider  $\check{\pi} \in \mathcal{P}_{\check{n}}$  and any gamble  $f$  on  $\mathcal{X}^{\check{n}}$ . Assume that  $I_{C_{\check{x}}} f \in \mathcal{R}$ .

We first prove that  $I_{C_{\check{x}\check{x}}} f \in \mathcal{R}$ . Consider the permutation  $\pi \in \mathcal{P}_N$  defined by

$$\pi(k) := \begin{cases} \check{\pi}^{-1}(k) & 1 \leq k \leq \check{n} \\ k & \check{n} + 1 \leq k \leq N, \end{cases} \quad (95)$$

then clearly  $\pi^t(I_{C_{\check{x}}}f) = (I_{C_{\check{x}}}f) \circ \pi = (I_{C_{\check{x}}} \circ \check{\pi}^{-1})f = I_{C_{\check{\pi}\check{x}}}f$ , so it follows from Proposition 9 that indeed  $I_{C_{\check{\pi}\check{x}}}f \in \mathcal{R}$ . This already implies that  $\mathcal{R}] \check{x} = \mathcal{R}] \check{\pi}\check{x}$ , and therefore also that  $\mathcal{R}] \check{x} = \mathcal{R}] \check{y}$ .

Since  $\mathcal{R}$  is coherent, it also follows from  $I_{C_{\check{x}}}f \in \mathcal{R}$  and the reasoning above that  $I_{C_{\check{m}}}f = \sum_{\check{y} \in [\check{m}]} I_{C_{\check{y}}}f \in \mathcal{R}$ , whence  $\mathcal{R}] \check{x} \subseteq \mathcal{R}] \check{m}$ . To prove the converse inequality, assume that  $I_{C_{\check{m}}}f \in \mathcal{R}$ . We know that  $[\check{m}] = \{\check{\pi}\check{x} : \check{\pi} \in \mathcal{P}_{\check{n}}\}$ , and therefore for any  $\check{y} \in [\check{m}]$  we can pick a  $\check{\pi}_{\check{y}} \in \mathcal{P}_{\check{n}}$  such that  $\check{\pi}_{\check{y}}\check{x} = \check{y}$ . With this  $\check{\pi}_{\check{y}}$  we construct a permutation  $\pi_{\check{y}} \in \mathcal{P}_N$  in the manner described above, which satisfies  $\pi_{\check{y}}^t(I_{C_{\check{x}}}f) = I_{C_{\check{y}}}f$ . But then the exchangeability and coherence of  $\mathcal{R}$  tell us that

$$I_{C_{\check{m}}}f + \sum_{\check{y} \in [\check{m}]} [(I_{C_{\check{x}}}f) - \pi_{\check{y}}^t(I_{C_{\check{x}}}f)] = I_{C_{\check{m}}}f + f \sum_{\check{y} \in [\check{m}]} [I_{C_{\check{x}}} - I_{C_{\check{y}}}] = |[\check{m}]|fI_{C_{\check{x}}} \quad (96)$$

belongs to  $\mathcal{R}$ , whence also  $I_{C_{\check{x}}}f \in \mathcal{R}$ , by coherence.  $\square$

**Proof of Theorem 17.** We begin with the sufficiency part. Assume that there is some coherent set  $\mathcal{S}$  of desirable gambles on  $\mathcal{N}^N$  such that  $\mathcal{R} = (\text{Hy}^N)^{-1}(\mathcal{S})$ . We show that  $\mathcal{R}$  is coherent and exchangeable, and that  $\mathcal{S} = \text{Hy}^N(\mathcal{R})$ .

We first show that  $\mathcal{R}$  is coherent. For D1, consider  $f \in \mathcal{G}(\mathcal{X}^N)$  with  $f = 0$ . Then obviously also  $\text{Hy}^N(f) = 0$  and therefore  $\text{Hy}^N(f) \notin \mathcal{S}$ . Hence  $f \notin \mathcal{R}$ . For D2, let  $f > 0$ . Then obviously also  $\text{Hy}^N(f) > 0$ , and therefore  $\text{Hy}^N(f) \in \mathcal{S}$ . Hence  $f \in \mathcal{R}$ . The proof for D3 is similar to the one for D4. For D4, let  $f_1, f_2 \in \mathcal{R}$ . Then  $g_1 := \text{Hy}^N(f_1) \in \mathcal{S}$  and  $g_2 := \text{Hy}^N(f_2) \in \mathcal{S}$ . This implies that  $\text{Hy}^N(f_1 + f_2) = g_1 + g_2 \in \mathcal{S}$ , so again  $f_1 + f_2 \in \mathcal{R}$ .

To show that  $\mathcal{R}$  is exchangeable, consider any  $f \in \mathcal{R}$  and  $f' \in \mathcal{D}_{\mathcal{N}}$ . We have to show that  $f + f' \in \mathcal{R}$ . It is clear that  $\text{Hy}^N(f + f') = \text{Hy}^N(f) + 0 = \text{Hy}^N(f) \in \mathcal{S}$ . Hence  $f + f' \in (\text{Hy}^N)^{-1}(\mathcal{S})$ , so indeed  $f + f' \in \mathcal{R}$ .

We show that  $\mathcal{S} = \text{Hy}^N(\mathcal{R})$ . Consider any gamble  $g \in \mathcal{G}(\mathcal{N}^N)$ , then using Eq. (40),  $\text{Hy}^N(\text{Co}^N(g)) = g$ . Since by assumption  $\mathcal{R} = (\text{Hy}^N)^{-1}(\mathcal{S})$ , we see that

$$g \in \mathcal{S} \Leftrightarrow \text{Hy}^N(\text{Co}^N(g)) \in \mathcal{S} \Leftrightarrow \text{Co}^N(g) \in \mathcal{R}. \quad (97)$$

This shows that  $\mathcal{S} = \{g \in \mathcal{G}(\mathcal{N}^N) : \text{Co}^N(g) \in \mathcal{R}\}$ . We show that also  $\mathcal{S} = \text{Hy}^N(\mathcal{R})$ . Let  $g \in \mathcal{S}$ , then we have just proved that  $\text{Co}^N(g) \in \mathcal{R}$ , and therefore, using Eq. (40),  $g = \text{Hy}^N(\text{Co}^N(g)) \in \text{Hy}^N(\mathcal{R})$ . Conversely, let  $g \in \text{Hy}^N(\mathcal{R})$ . Then there is some  $f \in \mathcal{R}$  such that  $g = \text{Hy}^N(f)$  and therefore  $\text{Co}^N(g) = \text{Co}^N(\text{Hy}^N(f)) = \text{ex}^N(f)$ , where the last equality follows from Eq. (40). Now Proposition 10 tells us that  $\text{ex}^N(f) \in \mathcal{R}$ , because  $f \in \mathcal{R}$  and  $\mathcal{R}$  is exchangeable. Hence  $\text{Co}^N(g) \in \mathcal{R}$  and therefore  $g \in \mathcal{S}$ .

Next, we turn to the necessity part. Suppose that  $\mathcal{R}$  is coherent and exchangeable. It suffices to prove that  $\mathcal{S} := \text{Hy}^N(\mathcal{R})$  is a coherent set of desirable gambles on  $\mathcal{N}^N$ , and that Eq. (43) is satisfied for this choice of  $\mathcal{S}$ .

We begin with the coherence of  $\text{Hy}^N(\mathcal{R})$ . For D1, consider  $g \in \mathcal{G}(\mathcal{N}^N)$  with  $g = 0$ . Assume *ex absurdo* that  $g \in \text{Hy}^N(\mathcal{R})$ , meaning that there is some  $f \in \mathcal{R}$  such that  $0 = g = \text{Hy}^N(f)$ , or in other words  $f \in \mathcal{D}_{\mathcal{N}}$ . This is impossible, due to Eq. (30). For D2, let  $g > 0$ . Then obviously also  $f := \text{Co}^N(g) > 0$ . Therefore  $f \in \mathcal{R}$  and, because of Eq. (40),  $g = \text{Hy}^N(\text{Co}^N(g)) = \text{Hy}^N(f) \in \text{Hy}^N(\mathcal{R})$ . The proof for D3 is similar to the one for D4. For D4, let  $g_1, g_2 \in \text{Hy}^N(\mathcal{R})$ , so there are  $f_1, f_2 \in \mathcal{R}$  such that  $g_1 = \text{Hy}^N(f_1)$  and  $g_2 = \text{Hy}^N(f_2)$ . Then by coherence of  $\mathcal{R}$ ,  $f_1 + f_2 \in \mathcal{R}$ , and therefore, by linearity of  $\text{Hy}^N$ ,

$$g_1 + g_2 = \text{Hy}^N(f_1) + \text{Hy}^N(f_2) = \text{Hy}^N(f_1 + f_2) \in \text{Hy}^N(\mathcal{R}). \quad (98)$$

Finally, we show that  $\mathcal{R} = (\text{Hy}^N)^{-1}(\text{Hy}^N(\mathcal{R}))$ . Consider  $f \in \mathcal{R}$ , then  $\text{Hy}^N(f) \in \text{Hy}^N(\mathcal{R})$  and therefore  $f \in (\text{Hy}^N)^{-1}(\text{Hy}^N(\mathcal{R}))$ . Conversely, consider a gamble  $f$  in  $(\text{Hy}^N)^{-1}(\text{Hy}^N(\mathcal{R}))$ . Then  $g := \text{Hy}^N(f) \in \text{Hy}^N(\mathcal{R})$ , so we infer that there is some  $f' \in \mathcal{R}$  such that  $g = \text{Hy}^N(f) = \text{Hy}^N(f')$ . Hence  $\text{Hy}^N(f - f') = 0$ , so  $f - f' \in \mathcal{D}_{\mathcal{N}}$  by Eqs. (40) and (28), and therefore  $f = f' + f - f' \in \mathcal{R} + \mathcal{D}_{\mathcal{N}}$ . This implies that  $f \in \mathcal{R}$ , since  $\mathcal{R}$  is exchangeable.  $\square$



**Proof of Corollary 18.** This result can be easily proved as an immediate consequence of Theorem 17 and Eq. (10). As an illustration, we give a more direct proof of the necessity part, based on Theorem 11. This theorem, together with Eq. (40), tells us that for any gamble  $f$  on  $\mathcal{X}^N$ ,  $\underline{P}(f) = \underline{P}(\text{ex}^N(f)) = \underline{P}(\text{Co}^N(\text{Hy}^N(f))) = \underline{Q}(\text{Hy}^N(f))$ .  $\square$

**Proof of Theorem 19.** We begin with the second statement. Recall that  $\mathcal{E}_{\text{ex}}^N(\mathcal{A}) = \mathcal{D}_{\mathcal{U}_N} + \mathcal{E}(\mathcal{A})$  from Theorem 13. Since  $\text{Hy}^N$  is a linear operator, it commutes with the posi operator, and therefore:

$$\begin{aligned} \text{Hy}^N(\mathcal{E}_{\text{ex}}^N(\mathcal{A})) &= \text{Hy}^N(\mathcal{D}_{\mathcal{U}_N}) + \text{Hy}^N(\mathcal{E}(\mathcal{A})) \\ &= \text{Hy}^N(\mathcal{E}(\mathcal{A})) \\ &= \text{posi}(\text{Hy}^N(\mathcal{G}_0^+(\mathcal{X}^N) \cup \mathcal{A})) \\ &= \text{posi}(\text{Hy}^N(\mathcal{G}_0^+(\mathcal{X}^N)) \cup \text{Hy}^N(\mathcal{A})) \\ &= \text{posi}(\mathcal{G}_0^+(\mathcal{X}^N) \cup \text{Hy}^N(\mathcal{A})) \\ &= \mathcal{E}(\text{Hy}^N(\mathcal{A})), \end{aligned}$$

where the second equality follows from  $\text{Hy}^N(\mathcal{D}_{\mathcal{U}_N}) = \{0\}$ , the third from Theorem 13, and the last from Theorem 1. The first statement is an immediate consequence of the second and Theorems 1, 13 and 17.  $\square$

**Proof of Proposition 20.** Recall that  $g \in \mathcal{S} \upharpoonright \check{m}$  iff there is some  $f \in \mathcal{G}(\mathcal{X}^{\check{m}})$  such that at the same time  $g = \text{Hy}^{\check{m}}(f)$  and  $I_{C_{[\check{m}]}}f \in \mathcal{R}$ , or in other words  $\text{Hy}^N(I_{C_{[\check{m}]}}f) \in \mathcal{S}$ . We therefore consider  $M \in \mathcal{N}^N$  and observe that

$$\text{Hy}^N(I_{C_{[\check{m}]}}f|M) = \frac{1}{|[M]|} \sum_{x \in [M]} (I_{C_{[\check{m}]}}f)(x) = \frac{1}{|[M]|} \sum_{\substack{\hat{x} \in [\check{m}], \hat{x} \in \mathcal{X}^{\check{m}} \\ (\hat{x}, \hat{x}) \in [M]}} f(\hat{x}), \quad (99)$$

so this value is zero unless  $M \geq \check{m}$ . In that case we can write  $M = \check{m} + \hat{m}$ , where  $\hat{m} := M - \check{m}$  is a count vector in  $\mathcal{N}^{\hat{m}}$ ; so we find that

$$\text{Hy}^N(I_{C_{[\check{m}]}}f|\check{m} + \hat{m}) = \frac{1}{|[\check{m} + \hat{m}]|} \sum_{\hat{x} \in [\hat{m}], \hat{x} \in [\hat{m}]} f(\hat{x}) = \frac{|[\check{m}]| |[\hat{m}]|}{|[\check{m} + \hat{m}]|} \text{Hy}^{\hat{m}}(f|\hat{m}). \quad (100)$$

Hence indeed  $g \in \mathcal{S} \upharpoonright \check{m}$  iff  ${}_{+\check{m}}(L_{\check{m}}g) \in \mathcal{S}$ .  $\square$

**Proof of Theorem 21.** It clearly suffices to give the proof in terms of count gambles. Because we have seen that  $\text{CoMn}^N$  is a linear isomorphism between the linear spaces  $\mathcal{G}(\mathcal{N}^N)$  and  $\mathcal{V}^N(\Sigma_{\mathcal{X}})$ , it is clear that  $\mathcal{S} = (\text{CoMn}^N)^{-1}(\mathcal{H})$  iff  $\mathcal{H} = \text{CoMn}^N(\mathcal{S})$ .

Suppose that  $\mathcal{S}$  is coherent, then we have to prove that  $\mathcal{H} = \text{CoMn}^N(\mathcal{S})$  is Bernstein coherent at degree  $N$ . Since  $\text{CoMn}^N$  is a linear isomorphism, it is clear that  $\mathcal{H}$  satisfies  $\text{B}_N1$ ,  $\text{B}_N3$  and  $\text{B}_N4$ , because  $\mathcal{S}$  satisfies  $\text{D1}$ ,  $\text{D3}$  and  $\text{D4}$ . To show that  $\mathcal{H}$  satisfies  $\text{B}_N2$ , consider  $p$  such that  $b_p^N > 0$  and therefore  $b_p^N \in \mathcal{S}$  by  $\text{D2}$ . Hence indeed  $p = \text{CoMn}^N(b_p^N) \in \text{CoMn}^N(\mathcal{S}) = \mathcal{H}$ .

Suppose that  $\mathcal{H}$  is Bernstein coherent at degree  $N$ , then we have to prove that  $\mathcal{S} = (\text{CoMn}^N)^{-1}(\mathcal{H})$  is coherent. Since  $(\text{CoMn}^N)^{-1}$  is a linear isomorphism, it is clear that  $\mathcal{S}$  satisfies  $\text{D1}$ ,  $\text{D3}$  and  $\text{D4}$ , because  $\mathcal{H}$  satisfies  $\text{B}_N1$ ,  $\text{B}_N3$  and  $\text{B}_N4$ . To show that  $\mathcal{S}$  satisfies  $\text{D2}$ , consider  $g > 0$ . Then  $p = \text{CoMn}^N(g)$  is such that  $b_p^N = g > 0$  and therefore  $p \in \mathcal{H}$  by  $\text{B}_N2$ . Hence indeed  $g = (\text{CoMn}^N)^{-1}(p) \in (\text{CoMn}^N)^{-1}(\mathcal{H}) = \mathcal{S}$ .  $\square$

**Proof of the equivalence of Eqs. (61) and (65).** We begin by proving that Eq. (61) implies Eq. (65). Consider any  $n_1 \leq n_2$ .

- $\subseteq$  Consider any  $g_2 \in \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1})$ , so there is some  $g_1 \in \mathcal{S}^{n_1}$  such that  $g_2 = \text{enl}_{n_1}^{n_2}(g_1)$ . Then it remains to prove that  $g_2 \in \mathcal{S}^{n_2}$ . But  $g_1 \in \mathcal{S}^{n_1}$  means that there is some  $f_1 \in$

$\mathcal{R}^{n_1}$  such that  $g_1 = \text{Hy}^{n_1}(f_1)$ . It then follows from Eq. (61) that  $f_2 := \text{ext}_{n_1}^{n_2}(f_1) \in \mathcal{R}^{n_2}$ , and therefore  $\text{Hy}^{n_2}(f_2) \in \mathcal{S}^{n_2}$ . But Eq. (64) tells us that

$$\text{Hy}^{n_2}(f_2) = \text{Hy}^{n_2}(\text{ext}_{n_1}^{n_2}(f_1)) = \text{enl}_{n_1}^{n_2}(\text{Hy}^{n_1}(f_1)) = \text{enl}_{n_1}^{n_2}(g_1) = g_2. \quad (101)$$

⊇ Consider any  $g_2 \in \mathcal{S}^{n_2} \cap \text{enl}_{n_1}^{n_2}(\mathcal{G}(\mathcal{N}^{n_1}))$ . We have to show that  $g_2 \in \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1})$ . On the one hand,  $g_2 \in \mathcal{S}^{n_2}$  implies that there is some  $f_2 \in \mathcal{R}^{n_2}$  such that  $g_2 = \text{Hy}^{n_2}(f_2)$ . On the other hand,  $g_2 \in \text{enl}_{n_1}^{n_2}(\mathcal{G}(\mathcal{N}^{n_1}))$  means that there is some gamble  $g_1$  on  $\mathcal{N}^{n_1}$  such that  $g_2 = \text{enl}_{n_1}^{n_2}(g_1)$ , and therefore also some gamble  $f_1$  on  $\mathcal{X}^{n_1}$  such that  $g_1 = \text{Hy}^{n_1}(f_1)$  and therefore

$$g_2 = \text{enl}_{n_1}^{n_2}(g_1) = \text{enl}_{n_1}^{n_2}(\text{Hy}^{n_1}(f_1)) = \text{Hy}^{n_2}(\text{ext}_{n_1}^{n_2}(f_1)), \quad (102)$$

if we also consider Eq. (64). Hence  $\text{Hy}^{n_2}(f_2) = \text{Hy}^{n_2}(\text{ext}_{n_1}^{n_2}(f_1))$ , and therefore also  $\text{ex}^{n_2}(f_2) = \text{ex}^{n_2}(\text{ext}_{n_1}^{n_2}(f_1))$ , by Eq. (40). Since  $f_2 \in \mathcal{R}^{n_2}$  we conclude from Proposition 10 that also  $\text{ext}_{n_1}^{n_2}(f_1) \in \mathcal{R}^{n_2}$ . Now we invoke Eq. (61) to find that  $\text{ext}_{n_1}^{n_2}(f_1) \in \text{ext}_{n_1}^{n_2}(\mathcal{R}^{n_1})$ , and therefore  $f_1 \in \mathcal{R}^{n_1}$ . But this implies that  $g_1 \in \mathcal{S}^{n_1}$  and consequently  $g_2 \in \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1})$ .

Next, we prove that Eq. (65) implies Eq. (61). Consider any  $n_1 \leq n_2$ .

⊆ Consider any  $f_2 \in \text{ext}_{n_1}^{n_2}(\mathcal{R}^{n_1})$ , so there is some  $f_1 \in \mathcal{R}^{n_1}$  such that  $f_2 = \text{ext}_{n_1}^{n_2}(f_1)$ . Then  $\text{Hy}^{n_1}(f_1) \in \mathcal{S}^{n_1}$ , and therefore Eq. (64) tells us that

$$\text{Hy}^{n_2}(f_2) = \text{Hy}^{n_2}(\text{ext}_{n_1}^{n_2}(f_1)) = \text{enl}_{n_1}^{n_2}(\text{Hy}^{n_1}(f_1)) \in \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1}). \quad (103)$$

We then deduce from Eq. (65) that  $\text{Hy}^{n_2}(f_2) \in \mathcal{S}^{n_2}$ , whence indeed  $f_2 \in \mathcal{R}^{n_2}$ .

⊇ Consider any  $f_2 \in \mathcal{R}^{n_2} \cap \text{ext}_{n_1}^{n_2}(\mathcal{G}(\mathcal{X}^{n_1}))$ . Then  $\text{Hy}^{n_2}(f_2) \in \mathcal{S}^{n_2}$  and there is some gamble  $f_1$  on  $\mathcal{X}^{n_1}$  such that  $f_2 = \text{ext}_{n_1}^{n_2}(f_1)$ . So we deduce from Eq. (64) that

$$\text{Hy}^{n_2}(f_2) = \text{Hy}^{n_2}(\text{ext}_{n_1}^{n_2}(f_1)) = \text{enl}_{n_1}^{n_2}(\text{Hy}^{n_1}(f_1)) \in \text{enl}_{n_1}^{n_2}(\mathcal{G}(\mathcal{N}^{n_1})) \quad (104)$$

as well. Therefore Eq. (65) tells us that  $\text{Hy}^{n_2}(f_2) \in \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1})$ , so there is some  $g_1 \in \mathcal{S}^{n_1}$  such that  $\text{Hy}^{n_2}(f_2) = \text{enl}_{n_1}^{n_2}(g_1)$ . Hence  $\text{enl}_{n_1}^{n_2}(\text{Hy}^{n_1}(f_1)) = \text{enl}_{n_1}^{n_2}(g_1)$ , and we infer from Lemma 31 that therefore  $\text{Hy}^{n_1}(f_1) = g_1$ , whence  $f_1 \in \mathcal{R}^{n_1}$ . This implies that indeed  $f_2 = \text{ext}_{n_1}^{n_2}(f_1) \in \text{ext}_{n_1}^{n_2}(\mathcal{R}^{n_1})$ .

This completes the proof.  $\square$

**Lemma 31.** *Consider any  $n_1 \leq n_2$  in  $\mathbb{N}_0$ . Then the extension map  $\text{enl}_{n_1}^{n_2}$  is one-to-one.*

*Proof.* Consider any gambles  $g_1$  and  $g_2$  on  $\mathcal{N}^{n_1}$  and assume that  $\text{enl}_{n_1}^{n_2}(g_1) = \text{enl}_{n_1}^{n_2}(g_2) =: g$ . Then we must prove that  $g_1 = g_2$ . Consider the polynomial  $p := \text{CoMn}^{n_2}(g)$ , then we infer from Eq. (67) that

$$\text{CoMn}^{n_1}(g_1) = \text{CoMn}^{n_2}(\text{enl}_{n_1}^{n_2}(g_1)) = p = \text{CoMn}^{n_2}(\text{enl}_{n_1}^{n_2}(g_2)) = \text{CoMn}^{n_1}(g_2), \quad (105)$$

which means that, with the notations of Appendix B,  $b_p^{n_1} = g_1 = g_2$  is the unique decomposition of the polynomial  $p$  in terms of the Bernstein basis polynomials of degree  $n_1$ .  $\square$

**Proof of the equivalence of Eqs. (65) and (66).** As a first step, we prove that Eq. (65) implies Eq. (66). Consider any  $n_1 \leq n_2$ .

⊆ Choose any  $p \in \mathcal{H}^{n_1}$ , then we know from Theorem 21 and the discussion in Appendix B that there is a unique  $g_1 := b_p^{n_1}$  in  $\mathcal{S}^{n_1}$  such that  $p = \text{CoMn}^{n_1}(g_1)$ . If we let  $g_2 := \text{enl}_{n_1}^{n_2}(g_1)$  then we infer from Eq. (67) that  $p = \text{CoMn}^{n_2}(g_2)$  as well. Since we infer from Eq. (65) that  $g_2 \in \mathcal{S}^{n_2}$ , we see that indeed  $p \in \text{CoMn}^{n_2}(\mathcal{S}^{n_2}) = \mathcal{H}^{n_2}$ .

⊇ Choose any  $p \in \mathcal{H}^{n_2} \cap \mathcal{V}^{n_1}(\Sigma_{\mathcal{X}})$ . Since  $p \in \mathcal{H}^{n_2}$  we infer from Theorem 21 and the discussion in Appendix B that there is a unique  $g_2 := b_p^{n_2}$  in  $\mathcal{S}^{n_2}$  such that  $p = \text{CoMn}^{n_2}(g_2)$ . On the other hand, since  $p$  is a polynomial of degree at most  $n_1$ , we know from the discussion in Appendix B that there is a unique Bernstein expansion  $g_1 := b_p^{n_1}$  in  $\mathcal{G}(\mathcal{N}^{n_1})$  such that  $p = \text{CoMn}^{n_1}(g_1)$ . The relation between the unique Bernstein expansions  $g_1$  and  $g_2$  is given by Zhou's formula:  $g_2 =$

$\text{enl}_{n_1}^{n_2}(g_1)$ . Hence  $g_2 \in \text{enl}_{n_1}^{n_2}(\mathcal{G}(\mathcal{N}^{n_1}))$  as well, and we infer from Eq. (65) that there is some  $g_3 \in \mathcal{S}^{n_1}$  such that  $g_2 = \text{enl}_{n_1}^{n_2}(g_3)$ . But since we have shown before that  $\text{enl}_{n_1}^{n_2}$  is one-to-one [Lemma 31], we infer that  $g_1 = g_3$  and therefore  $g_1 \in \mathcal{S}^{n_1}$ , whence indeed  $p \in \text{CoMn}^{n_1}(\mathcal{S}^{n_1}) = \mathcal{H}^{n_1}$ .

Next, we prove that Eq. (66) implies Eq. (65). Consider any  $n_1 \leq n_2$ .

- $\subseteq$  Choose any  $g_2 \in \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1})$ . Then there is some  $g_1 \in \mathcal{S}^{n_1}$  such that  $g_2 = \text{enl}_{n_1}^{n_2}(g_1)$ . Let  $p = \text{CoMn}^{n_1}(g_1)$ , then we infer from Theorem 21 that  $p \in \mathcal{H}^{n_1}$ . But Eq. (67) also tells us that  $p = \text{CoMn}^{n_1}(g_1) = \text{CoMn}^{n_1}(\text{enl}_{n_1}^{n_2}(g_1)) = \text{CoMn}^{n_2}(g_2)$ , and since also  $p \in \mathcal{H}^{n_2}$  by Eq. (66), we see that indeed  $g_2 \in (\text{CoMn}^{n_2})^{-1}(\mathcal{H}^{n_2}) = \mathcal{S}^{n_2}$ .
- $\supseteq$  Choose any  $g_2 \in \mathcal{S}^{n_2} \cap \text{enl}_{n_1}^{n_2}(\mathcal{G}(\mathcal{N}^{n_1}))$ . Let  $p = \text{CoMn}^{n_2}(g_2)$  then it follows from Theorem 21 that  $p \in \mathcal{H}^{n_2}$ . But we also know that there is some  $g_1 \in \mathcal{G}(\mathcal{N}^{n_1})$  such that  $g_2 = \text{enl}_{n_1}^{n_2}(g_1)$  and therefore  $p = \text{CoMn}^{n_2}(\text{enl}_{n_1}^{n_2}(g_1)) = \text{CoMn}^{n_1}(g_1)$ , by Eq. (67). So  $p$  is a polynomial of degree at most  $n_1$ , and we then infer from Eq. (66) that  $p \in \mathcal{H}^{n_1}$ , whence  $g_1 \in (\text{CoMn}^{n_1})^{-1}(\mathcal{H}^{n_1}) = \mathcal{S}^{n_1}$ , and therefore indeed  $g_2 \in \text{enl}_{n_1}^{n_2}(\mathcal{S}^{n_1})$ .

This completes the proof.  $\square$

**Proof that B1 is equivalent to B5 under B2–B4.** It is clear that B5 implies B1, because if a polynomial is zero, then so are all its Bernstein expansions. The proof is therefore complete if we can show that B5 follows from B1–B4. Consider a polynomial  $p$  for which there is some  $n \geq \deg(p)$  such that  $b_p^n \leq 0$ , and assume *ex absurdo* that  $p \in \mathcal{H}$ . Then clearly  $p \neq 0$  by B1, and therefore  $b_p^n < 0$ . But then  $b_{-p}^n = -b_p^n > 0$ , so  $-p \in \mathcal{H}$  by B2, and then  $0 = p + (-p) \in \mathcal{H}$  by B4, a contradiction.  $\square$

**Lemma 32.** Consider a subset  $\mathcal{F}$  of  $\mathcal{V}(\Sigma_{\mathcal{X}})$ , and define the sets  $\mathcal{F}^n := \mathcal{F} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})$  for all  $n \in \mathbb{N}$ . Then:

- (i)  $\mathcal{F}^{n_1} = \mathcal{F}^{n_2} \cap \mathcal{V}^{n_1}(\Sigma_{\mathcal{X}})$  for all  $0 \leq n_1 \leq n_2$ ;
- (ii) For all  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$ , if  $n \geq \deg(p)$  then  $p \in \mathcal{F} \Leftrightarrow p \in \mathcal{F}^n$ ;
- (iii) For all  $k \geq 0$ ,  $\mathcal{F} = \bigcup_{n \in \mathbb{N}} \mathcal{F}^n = \bigcup_{n \geq k} \mathcal{F}^n$ ;
- (iv)  $\mathcal{F}$  is Bernstein coherent iff  $\mathcal{F}^n$  is Bernstein coherent at degree  $n$  for all  $n \in \mathbb{N}_0$ .

**Proof of Lemma 32.** The proof of the first two statements is trivial.

We turn to the proof of (iii). Since  $\mathcal{F}^n \subseteq \mathcal{F}$  for all  $n \geq 1$ , we see at once that  $\bigcup_{n \in \mathbb{N}_0} \mathcal{F}^n \subseteq \mathcal{F}$ . To prove the converse inequality, consider any  $p \in \mathcal{F}$ . With  $m = \deg(p)$  we infer from (ii) that  $p \in \mathcal{F}^m$  and therefore  $p \in \bigcup_{n \in \mathbb{N}} \mathcal{F}^n$ . The second equality now follows at once from (i).

On to the proof of (iv).

- $\Rightarrow$  Assume first of all that  $\mathcal{F}$  is Bernstein coherent, and consider any  $n \in \mathbb{N}_0$ . Then we have to prove that  $\mathcal{F}^n$  is Bernstein coherent at degree  $n$ . It is obvious that  $\mathcal{F}^n$  satisfies  $B_{n,1}$ ,  $B_{n,3}$  and  $B_{n,4}$  because  $\mathcal{F}$  satisfies B1, B3 and B4. To prove that  $\mathcal{F}^n$  satisfies  $B_{n,2}$ , consider  $p \in \mathcal{V}^n(\Sigma_{\mathcal{X}})$  with  $b_p^n > 0$ . Since clearly  $n \geq \deg(p)$ , we infer from B2 that  $p \in \mathcal{F}$  and therefore indeed  $p \in \mathcal{F} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}}) = \mathcal{F}^n$ .
- $\Rightarrow$  Finally, assume that  $\mathcal{F}^n$  is Bernstein coherent at degree  $n$  for all  $n \in \mathbb{N}_0$ . Then we have to prove that  $\mathcal{F}$  is Bernstein coherent. It follows readily from (iii) that  $\mathcal{F}$  satisfies B1, B3 and B4. To prove that  $\mathcal{F}$  satisfies B2, consider any polynomial  $p$  and assume that  $b_p^n > 0$  for some  $n \geq \deg(p)$ . Then clearly  $p \in \mathcal{V}^n(\Sigma_{\mathcal{X}})$  and therefore  $p \in \mathcal{F}^n$ , by  $B_{n,2}$ . Hence indeed  $p \in \mathcal{F}$ .  $\square$

**Lemma 33.** Consider a time-consistent, coherent and exchangeable family  $\mathcal{R}^n$ ,  $n \in \mathbb{N}_0$  of sets of desirable gambles on  $\mathcal{X}^n$ , and the associated count representations  $\mathcal{S}^n := \text{Hy}^n(\mathcal{R}^n)$  on  $\mathcal{N}^n$  and frequency representations  $\mathcal{H}^n := \text{CoMn}^n(\mathcal{S}^n) = \text{Mn}^n(\mathcal{R}^n)$  on  $\mathcal{V}^n(\Sigma_{\mathcal{X}})$ . Let  $\mathcal{H} := \bigcup_{n \in \mathbb{N}_0} \mathcal{H}^n$ . Then the sequence  $\mathcal{H}^n$  is non-decreasing, and  $\mathcal{H}^n = \mathcal{H} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})$ .

*Proof.* Because the family  $\mathcal{R}^n$ ,  $n \in \mathbb{N}_0$  is time-consistent, the sets  $\mathcal{H}^n$  satisfy the time-consistency property (66). This already implies that the sequence  $\mathcal{H}^n$  is non-decreasing.

We now show that  $\mathcal{H}^n = \mathcal{H} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})$ . Indeed:

$$\begin{aligned} \mathcal{H} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}}) &= \bigcup_{k \geq 1} \mathcal{H}^k \cap \mathcal{V}^n(\Sigma_{\mathcal{X}}) = \left( \bigcup_{1 \leq k \leq n} \mathcal{H}^k \cap \mathcal{V}^n(\Sigma_{\mathcal{X}}) \right) \cup \left( \bigcup_{k > n} \mathcal{H}^k \cap \mathcal{V}^n(\Sigma_{\mathcal{X}}) \right) \\ &= \left( \bigcup_{1 \leq k \leq n} \mathcal{H}^k \right) \cup \left( \bigcup_{k > n} \mathcal{H}^k \right) = \mathcal{H}^n \cup \mathcal{H}^n = \mathcal{H}^n, \end{aligned} \quad (106)$$

where the third and fourth equalities follow from the time-consistency condition (66).  $\square$

**Proof of Theorem 22.** It clearly suffices to give the proof in terms of the count representations.

First of all, consider a Bernstein coherent  $\mathcal{F} \subseteq \mathcal{V}(\Sigma_{\mathcal{X}})$ , then we have to prove that the  $\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{F})$ ,  $n \in \mathbb{N}_0$  are coherent and satisfy the time-consistency condition (65). Let  $\mathcal{F}^n := \mathcal{F} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})$  then clearly

$$\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{F}) = (\text{CoMn}^n)^{-1}(\mathcal{F} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})) = (\text{CoMn}^n)^{-1}(\mathcal{F}^n). \quad (107)$$

We then infer from Lemma 32(iv)&(i) that  $\mathcal{F}^n$  is Bernstein coherent at degree  $n$ , and that the  $\mathcal{S}^n$ ,  $n \in \mathbb{N}_0$  satisfy the time-consistency condition (66). Hence the  $\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{F}^n)$  satisfy the time consistency condition (65), and we infer from the Finite Representation Theorem 21 that all  $\mathcal{S}^n$  are coherent.

Conversely, suppose that we have a family of coherent  $\mathcal{S}^n$  that satisfy the time-consistency condition (65). Let  $\mathcal{H}^n = \text{CoMn}^n(\mathcal{S}^n)$  then we know that  $\mathcal{H}^n$  is Bernstein coherent at degree  $n$  [by Theorem 21] and that the  $\mathcal{H}^n$  satisfy the time-consistency condition (66). Let  $\mathcal{H} := \bigcup_{n \in \mathbb{N}_0} \mathcal{H}^n$ . Then it follows from Lemma 33 that  $\mathcal{H}^n = \mathcal{H} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})$ , and from Lemma 32(iv) that  $\mathcal{H}$  is Bernstein coherent. Moreover, since  $\mathcal{H}^n = \text{CoMn}^n(\mathcal{S}^n)$  and  $\text{CoMn}^n$  is a linear isomorphism,

$$\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{H}^n) = (\text{CoMn}^n)^{-1}(\mathcal{H} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})) = (\text{CoMn}^n)^{-1}(\mathcal{H}). \quad (108)$$

To prove unicity, consider any  $\mathcal{F} \subseteq \mathcal{V}(\Sigma_{\mathcal{X}})$  such that  $\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{F})$  and let  $\mathcal{F}^n := \mathcal{F} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})$ . Then

$$\mathcal{S}^n = (\text{CoMn}^n)^{-1}(\mathcal{F}) = (\text{CoMn}^n)^{-1}(\mathcal{F} \cap \mathcal{V}^n(\Sigma_{\mathcal{X}})) = (\text{CoMn}^n)^{-1}(\mathcal{F}^n) \quad (109)$$

and therefore  $\mathcal{F}^n = \text{CoMn}^n(\mathcal{S}^n) = \mathcal{H}^n$ . We then infer from Lemma 32(iii) that  $\mathcal{F} = \bigcup_{n \in \mathbb{N}_0} \mathcal{F}^n = \bigcup_{n \in \mathbb{N}_0} \mathcal{H}^n = \mathcal{H}$ .  $\square$

**Proof of Theorem 23.** We already know that the models in the updated family  $\mathcal{R}^{\hat{n}} \upharpoonright \check{m}$ ,  $\hat{n} \in \mathbb{N}_0$  are coherent and exchangeable, by Propositions 15 and 16. To show that this family has a frequency representation, it suffices, by the Infinite Representation Theorem 22, to show that it is time-consistent (satisfies Eq. (61)). Consider any  $\hat{r} \leq \hat{s}$  in  $\mathbb{N}_0$ , then we have to show that

$$\text{ext}_{\hat{r}}^{\hat{s}}(\mathcal{R}^{\hat{r}} \upharpoonright \check{m}) = \mathcal{R}^{\hat{s}} \upharpoonright \check{m} \cap \text{ext}_{\hat{r}}^{\hat{s}}(\mathcal{G}(\mathcal{X}^{\hat{r}})). \quad (110)$$

- $\subseteq$  Let  $f' \in \text{ext}_{\hat{r}}^{\hat{s}}(\mathcal{R}^{\hat{r}} \upharpoonright \check{m})$ , so there is some  $f \in \mathcal{R}^{\hat{r}} \upharpoonright \check{m}$  such that  $f' = \text{ext}_{\hat{r}}^{\hat{s}}(f)$ . Now  $f \in \mathcal{R}^{\hat{r}} \upharpoonright \check{m}$  means that  $f I_{C_{\check{m}}} \in \mathcal{R}^{\hat{n}+\hat{r}}$ . Since the family  $\mathcal{R}^n$ ,  $n \in \mathbb{N}_0$  is by assumption time-consistent, we infer that  $\text{ext}_{\hat{r}}^{\hat{s}}(f) I_{C_{\check{m}}} = \text{ext}_{\hat{n}+\hat{r}}^{\hat{n}+\hat{s}}(f I_{C_{\check{m}}}) \in \mathcal{R}^{\hat{n}+\hat{s}}$ , and therefore  $f' = \text{ext}_{\hat{r}}^{\hat{s}}(f) \in \mathcal{R}^{\hat{s}} \upharpoonright \check{m}$ .
- $\supseteq$  To prove the converse inequality, let  $f' \in \mathcal{R}^{\hat{s}} \upharpoonright \check{m} \cap \text{ext}_{\hat{r}}^{\hat{s}}(\mathcal{G}(\mathcal{X}^{\hat{r}}))$ .  $f' \in \mathcal{R}^{\hat{s}} \upharpoonright \check{m}$  means that  $f' I_{C_{\check{m}}} \in \mathcal{R}^{\hat{n}+\hat{s}}$ . On the other hand,  $f' \in \text{ext}_{\hat{r}}^{\hat{s}}(\mathcal{G}(\mathcal{X}^{\hat{r}}))$  means that there is some  $f \in \mathcal{G}(\mathcal{X}^{\hat{r}})$  such that  $f' = \text{ext}_{\hat{r}}^{\hat{s}}(f)$ , and therefore  $\text{ext}_{\hat{r}}^{\hat{s}}(f) I_{C_{\check{m}}} \in \mathcal{R}^{\hat{n}+\hat{s}}$ . So we infer from the time-consistency of the family  $\mathcal{R}^n$ ,  $n \in \mathbb{N}_0$  that  $\text{ext}_{\hat{r}}^{\hat{s}}(f) I_{C_{\check{m}}} \in \text{ext}_{\hat{n}+\hat{r}}^{\hat{n}+\hat{s}}(\mathcal{R}^{\hat{n}+\hat{r}})$ . This means that there is some  $f'' \in \mathcal{R}^{\hat{n}+\hat{r}}$  such that  $\text{ext}_{\hat{r}}^{\hat{s}}(f) I_{C_{\check{m}}} = \text{ext}_{\hat{n}+\hat{r}}^{\hat{n}+\hat{s}}(f'')$ , which clearly implies that  $f'' = f I_{C_{\check{m}}}$ , and therefore indeed  $f \in \mathcal{R}^{\hat{r}} \upharpoonright \check{m}$ .

The only thing that remains to be proved is Eq. (71). We already know from Theorem 22 that  $\mathcal{H} \upharpoonright \check{m} = \bigcup_{\hat{n} \in \mathbb{N}_0} \text{CoMn}^{\hat{n}}(\mathcal{S}^{\hat{n}} \upharpoonright \check{m})$ . This triggers a concatenation of equivalences:

$$\begin{aligned}
p \in \mathcal{H} \upharpoonright \check{m} &\Leftrightarrow (\exists \hat{n} \in \mathbb{N}_0) (\exists g \in \mathcal{G}(\mathcal{N}^{\hat{n}})) (\exists g' \in \mathcal{S}^{\check{n}+\hat{n}}) (g' = +_{\check{m}}(L_{\check{m}}g) \text{ and } p = \text{CoMn}^{\hat{n}}(g)) \\
&\Leftrightarrow (\exists \hat{n} \in \mathbb{N}_0) (\exists g' \in \mathcal{S}^{\check{n}+\hat{n}}) \text{CoMn}^{\check{n}+\hat{n}}(g') = B_{\check{m}}p \\
&\Leftrightarrow B_{\check{m}}p \in \bigcup_{\hat{n} \in \mathbb{N}_0} \text{CoMn}^{\check{n}+\hat{n}}(\mathcal{S}^{\check{n}+\hat{n}}) \\
&\Leftrightarrow B_{\check{m}}p \in \mathcal{H}, \tag{111}
\end{aligned}$$

where the first equivalence follows from Eq. (48) and the second from Lemma 34. For the last equivalence, consider Lemma 33 and the fact that  $B_{\check{m}}p$  is a polynomial of degree at least  $\check{n}$ .  $\square$

**Lemma 34.** Consider  $\check{n}, \hat{n} \in \mathbb{N}_0$ , and  $\check{m} \in \mathcal{N}^{\check{n}}$ . For all gambles  $g$  on  $\mathcal{N}^{\hat{n}}$  and  $g'$  on  $\mathcal{N}^{\check{n}+\hat{n}}$ :

$$\text{CoMn}^{\check{n}+\hat{n}}(g') = B_{\check{m}} \text{CoMn}^{\hat{n}}(g) \Leftrightarrow g' = +_{\check{m}}(L_{\check{m}}g). \tag{112}$$

*Proof.* We find that

$$\begin{aligned}
\text{CoMn}^{\check{n}+\hat{n}}(+_{\check{m}}(L_{\check{m}}g)) &= \sum_{M \in \mathcal{N}^{\check{n}+\hat{n}}} +_{\check{m}}(L_{\check{m}}g)(M) B_M \\
&= \sum_{\hat{m} \in \mathcal{N}^{\hat{n}}} L_{\check{m}}(\hat{m}) g(\hat{m}) B_{\check{m}+\hat{m}} \\
&= \sum_{\hat{m} \in \mathcal{N}^{\hat{n}}} g(\hat{m}) B_{\check{m}} B_{\hat{m}} = B_{\check{m}} \sum_{\hat{m} \in \mathcal{N}^{\hat{n}}} g(\hat{m}) B_{\hat{m}} = B_{\check{m}} \text{CoMn}^{\hat{n}}(g), \tag{113}
\end{aligned}$$

where the second equality follows from Eq. (47), and the third from Eqs. (45) and (52). The first and last equalities go back to Eq. (51).

Conversely, consider any  $g'$  in  $\mathcal{G}(\mathcal{N}^{\check{n}+\hat{n}})$  such that  $\text{CoMn}^{\check{n}+\hat{n}}(g') = B_{\check{m}} \text{CoMn}^{\hat{n}}(g)$ . Since  $B_{\check{m}} \text{CoMn}^{\hat{n}}(g)$  is a polynomial of degree at most  $\check{n} + \hat{n}$ , we know from the discussion in Appendix B that there is one and only one such  $g'$ , as it represents the coefficients of the unique expansion of the polynomial  $B_{\check{m}} \text{CoMn}^{\hat{n}}(g)$  in the multivariate Bernstein basis of degree  $\check{n} + \hat{n}$ . Since we have seen in the first part of the proof that  $\text{CoMn}^{\check{n}+\hat{n}}(+_{\check{m}}(L_{\check{m}}g)) = B_{\check{m}} \text{CoMn}^{\hat{n}}(g)$ , it follows that  $g' = +_{\check{m}}(L_{\check{m}}g)$ .  $\square$

**Proof of Proposition 24.** That  $\underline{P}_{\mathcal{H}}$  is a linear functional that dominates the min functional follows from Lemma 35. We now show that  $\underline{P}_{\mathcal{H}}(p) = p(\vartheta)$  for all  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$ , where  $\vartheta_z := \underline{P}_{\mathcal{H}}(B_{e_z})$  for all  $z \in \mathcal{X}$ . Consider any  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$  and  $n \geq \deg(p)$ , then we know that  $p = \sum_{m \in \mathcal{N}^n} b_p^n(m) B_m$ , and therefore  $\underline{P}_{\mathcal{H}}(p) = \sum_{m \in \mathcal{N}^n} b_p^n(m) \underline{P}_{\mathcal{H}}(B_m)$ , using the linearity of  $\underline{P}_{\mathcal{H}}$  [Lemma 35]. To find out what  $\underline{P}_{\mathcal{H}}(B_m)$  is, observe that we can write  $B_m$  as a product of simpler Bernstein basis polynomials:  $B_m = \binom{n}{m} \prod_{z \in \mathcal{X}} B_{e_z}^{m_z}$ , and therefore Lemmas 35 and 36 tell us that  $\underline{P}_{\mathcal{H}}(B_m) = \binom{n}{m} \prod_{z \in \mathcal{X}} \underline{P}_{\mathcal{H}}(B_{e_z})^{m_z} = \binom{n}{m} \prod_{z \in \mathcal{X}} \vartheta^{m_z} = B_m(\vartheta)$ . Hence indeed  $\underline{P}_{\mathcal{H}}(p) = \sum_{m \in \mathcal{N}^n} b_p^n(m) B_m(\vartheta) = p(\vartheta)$ .

To complete the proof, consider any gamble  $g$  on  $\mathcal{N}^n$ . Clearly,

$$\begin{aligned}
\underline{P}_{\mathcal{S}^n}(g) &= \sup \{ \alpha : g - \alpha \in \mathcal{S}^n \} = \sup \{ \alpha : g - \alpha \in (\text{CoMn}^n)^{-1}(\mathcal{H}) \} \\
&= \sup \{ \alpha : \text{CoMn}^n(g - \alpha) \in \mathcal{H} \} = \sup \{ \alpha : \text{CoMn}^n(g) - \alpha \in \mathcal{H} \} \\
&= \underline{P}_{\mathcal{H}}(\text{CoMn}^n(g)).
\end{aligned}$$

The rest of the proof is now immediate.  $\square$

**Lemma 35.** Consider any maximal element  $\mathcal{H}$  of  $\mathbb{D}_{\text{Be}}(\Sigma_{\mathcal{X}})$  that satisfies either of the equivalent conditions (72) or (73). Then  $\underline{P}_{\mathcal{H}}$  is a linear functional that dominates the min functional.

*Proof.* It follows from the Bernstein coherence of  $\mathcal{H}$  that  $\underline{P}_{\mathcal{H}}$  is super-additive [use B4] and positively homogeneous [use B3]. For any polynomial  $p$ :

$$\begin{aligned}\underline{P}_{\mathcal{H}}(p) &= \sup\{\alpha: p - \alpha \in \mathcal{H}\} = \inf\{\beta: p - \beta \notin \mathcal{H}\} \\ &= \inf\{\beta: \beta - p \in \mathcal{H}\} = \overline{P}_{\mathcal{H}}(p);\end{aligned}$$

the second equality follows from the fact that  $\{\alpha: p - \alpha \in \mathcal{H}\}$  is a down-set [use B2 and B4], and the third equality follows from the maximality of  $\mathcal{H}$  and Proposition 2. This shows that  $\underline{P}_{\mathcal{H}}$  is self-conjugate, which together with the super-additivity and positive homogeneity readily implies that  $\underline{P}_{\mathcal{H}}$  is additive and homogeneous, and therefore a linear functional. To show that  $\underline{P}_{\mathcal{H}}$  dominates min, consider any polynomial  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$  and any  $n \geq \deg(p)$ . Then there are two possibilities. If  $p$  is a constant, then  $p = \min b_p^n$  and therefore  $p - \alpha \in \mathcal{H} \Leftrightarrow \alpha < \min b_p^n$ , so  $\underline{P}_{\mathcal{H}}(p) = \min b_p^n$ . If  $p$  is not constant, then we infer from Eq. (117) in Appendix B that  $p - \min b_p^n > 0$  and therefore  $p - \min b_p^n \in \mathcal{H}$ , by B2. Hence  $\underline{P}_{\mathcal{H}}(p) \geq \min b_p^n$ . So we infer that this inequality holds for all  $p$  and all  $n \geq \deg(p)$ , whence indeed  $\underline{P}_{\mathcal{H}}(p) \geq \sup_{n \geq \deg(p)} \min b_p^n = \min p$ , where the equality follows from Proposition 37 in Appendix B.  $\square$

**Lemma 36.** *Consider any maximal element  $\mathcal{H}$  of  $\mathbb{D}_{\text{Be}}(\Sigma_{\mathcal{X}})$  that satisfies either of the equivalent conditions (72) or (73). Then  $\underline{P}_{\mathcal{H}}(B_m p) = \underline{P}_{\mathcal{H}}(B_m) \underline{P}_{\mathcal{H}}(p)$  for all  $p \in \mathcal{V}(\Sigma_{\mathcal{X}})$  and all count vectors  $m$ .*

*Proof.* Observe that for all real  $\alpha$  and  $\beta$ :

$$B_m p - \alpha = B_m(p - \beta) + (\beta B_m - \alpha). \quad (114)$$

First, consider any  $\alpha < \underline{P}_{\mathcal{H}}(B_m p)$  and  $\beta > \underline{P}_{\mathcal{H}}(p)$ . Then  $B_m p - \alpha \in \mathcal{H}$  and  $p - \beta \notin \mathcal{H}$ . If we take into account the maximality of  $\mathcal{H}$  and Proposition 2, the latter leads to  $\beta - p \in \mathcal{H}$ , and therefore  $B_m(\beta - p) \in \mathcal{H}$ , using condition (73). But then Eq. (114) and B4 lead to the conclusion that  $\beta B_m - \alpha \in \mathcal{H}$ . Hence  $\underline{P}_{\mathcal{H}}(\beta B_m - \alpha) \geq 0$ , whence  $\beta \underline{P}_{\mathcal{H}}(B_m) \geq \alpha$ , using the linearity of  $\underline{P}_{\mathcal{H}}$  [see Lemma 35]. Since this inequality holds for all  $\alpha < \underline{P}_{\mathcal{H}}(B_m p)$  and  $\beta > \underline{P}_{\mathcal{H}}(p)$ , we infer that  $\underline{P}_{\mathcal{H}}(p) \underline{P}_{\mathcal{H}}(B_m) \geq \underline{P}_{\mathcal{H}}(B_m p)$ .

To prove the converse inequality, consider any  $\alpha > \underline{P}_{\mathcal{H}}(B_m p)$  and  $\beta < \underline{P}_{\mathcal{H}}(p)$ . Then  $B_m p - \alpha \notin \mathcal{H}$  and  $p - \beta \in \mathcal{H}$ . If we take into account the maximality of  $\mathcal{H}$  and Proposition 2, the former leads to  $\alpha - B_m p \in \mathcal{H}$ , and the latter to  $B_m(p - \beta) \in \mathcal{H}$ , using condition (73). But then Eq. (114) and B4 lead to the conclusion that  $\alpha - \beta B_m \in \mathcal{H}$ . Hence  $\underline{P}_{\mathcal{H}}(\alpha - \beta B_m) \geq 0$ , whence  $\beta \underline{P}_{\mathcal{H}}(B_m) \leq \alpha$ , using the linearity of  $\underline{P}_{\mathcal{H}}$  [see Lemma 35]. Since this inequality holds for all  $\alpha > \underline{P}_{\mathcal{H}}(B_m p)$  and  $\beta < \underline{P}_{\mathcal{H}}(p)$ , we infer that  $\underline{P}_{\mathcal{H}}(p) \underline{P}_{\mathcal{H}}(B_m) \leq \underline{P}_{\mathcal{H}}(B_m p)$ .  $\square$

**Proof of Theorem 25.** This is an instance of Theorem 1 with linear space  $\mathcal{K} := \mathcal{V}(\Sigma_{\mathcal{X}})$  and cone  $\mathcal{C} := \mathcal{V}^+(\Sigma_{\mathcal{X}})$ .  $\square$

**Proof of Proposition 37.** Eq. (122) follows from the fact that the  $b_p^n$  converge uniformly to the polynomial  $p$  as  $n \rightarrow \infty$ ; see for instance Trump and Prautzsch [26]. Alternatively, it can be shown [see 22, Section 11.9] that for  $n \geq r$  and  $M \in \mathcal{N}^n$ :

$$b_p^n(M) = \sum_{m \in \mathcal{N}^r} b_p^r(m) B_m \left(\frac{M}{n}\right) + O\left(\frac{1}{n}\right) = p\left(\frac{M}{n}\right) + O\left(\frac{1}{n}\right). \quad (115)$$

Hence  $\min b_p^n \geq \min p + O\left(\frac{1}{n}\right)$  for any  $n \geq r$ , and as a consequence  $\lim_{n \rightarrow \infty, n \geq r} \min b_p^n \geq \min p$ . If we now use Equation (121), we see that  $\lim_{n \rightarrow \infty, n \geq r} \min b_p^n = \min p$ . The proof of the other equality is analogous.  $\square$

## APPENDIX B. MULTIVARIATE BERNSTEIN BASIS POLYNOMIALS

With any  $n \geq 0$  and  $m \in \mathcal{N}^n$  there corresponds a Bernstein (basis) polynomial of degree  $n$  on  $\Sigma_{\mathcal{X}}$ , given by  $B_m(\theta) = \frac{|[m]|}{|[M]|} \prod_{x \in \mathcal{X}} \theta_x^{m_x}$ ,  $\theta \in \Sigma_{\mathcal{X}}$ . These polynomials have a number of very interesting properties, see for instance Prautzsch et al. [22, Chapters 10 and 11], which we list here:

BP1. The set  $\{B_m : m \in \mathcal{N}^n\}$  of all Bernstein basis polynomials of fixed degree  $n$  is linearly independent: if  $\sum_{m \in \mathcal{N}^n} \lambda_m B_m = 0$ , then  $\lambda_m = 0$  for all  $m$  in  $\mathcal{N}^n$ .

BP2. The set  $\{B_m : m \in \mathcal{N}^n\}$  of all Bernstein basis polynomials of fixed degree  $n$  forms a partition of unity:  $\sum_{m \in \mathcal{N}^n} B_m = 1$ .

BP3. All Bernstein basis polynomials are non-negative, and strictly positive in the interior of  $\Sigma_{\mathcal{X}}$ .

BP4. The set  $\{B_m : m \in \mathcal{N}^n\}$  of all Bernstein basis polynomials of fixed degree  $n$  forms a basis for the linear space of all polynomials whose degree is at most  $n$ .

Property BP4 follows from BP1 and BP2. It follows from BP4 that:

BP5. Any polynomial  $p$  of degree  $r$  has a unique expansion in terms of the Bernstein basis polynomials of fixed degree  $n \geq r$ ,

or in other words, there is a unique gamble  $b_p^n$  on  $\mathcal{N}^n$  such that

$$p = \sum_{m \in \mathcal{N}^n} b_p^n(m) B_m = \text{CoMn}^n(b_p^n). \quad (116)$$

This tells us [also use BP2 and B3] that each  $p(\theta)$  is a convex combination of the Bernstein coefficients  $b_p^n(m)$ ,  $m \in \mathcal{N}^n$  whence for all  $\theta \in \Sigma_{\mathcal{X}}$

$$\min b_p^n \leq \min p \leq p(\theta) \leq \max p \leq \max b_p^n. \quad (117)$$

It follows from a combination of BP2 and BP4 that for all  $k \geq 0$  and all  $M$  in  $\mathcal{N}^{n+k}$ ,

$$b_p^{n+k}(M) = \sum_{m \in \mathcal{N}^n} \frac{|[m]| |[M-m]|}{|[M]|} b_p^n(m), \quad (118)$$

or in other words

$$b_p^{n+k} = \text{enl}_n^{n+k}(b_p^n). \quad (119)$$

This is *Zhou's formula* [see 22, Section 11.9]. Hence [let  $p = 1$  and use BP2] we find that for all  $k \geq 0$  and all  $M$  in  $\mathcal{N}^{n+k}$ ,

$$\sum_{m \in \mathcal{N}^n} \frac{|[m]| |[M-m]|}{|[M]|} = 1. \quad (120)$$

The expressions (118) and (120) also imply that each  $b_p^{n+k}(M)$  is a convex combination of the  $b_p^n(m)$ , and therefore  $\min b_p^{n+k} \geq \min b_p^n$  and  $\max b_p^{n+k} \leq \max b_p^n$ . Combined with the inequalities in (117), this leads to:

$$[\min p, \max p] \subseteq [\min b_p^{n+k}, \max b_p^{n+k}] \subseteq [\min b_p^n, \max b_p^n] \quad (121)$$

for all  $n \geq m$  and  $k \geq 0$ . This means that the non-decreasing sequence  $\min b_p^n$  converges to some real number not greater than  $\min p$ , and, similarly, the non-increasing sequence  $\max b_p^n$  converges to some real number not smaller than  $\max p$ . The following proposition strengthens this.

**Proposition 37.** *For any polynomial  $p$  on  $\Sigma_{\mathcal{X}}$  of degree up to  $r$ ,*

$$\lim_{\substack{n \rightarrow \infty \\ n \geq r}} [\min b_p^n, \max b_p^n] = [\min p, \max p] = p(\Sigma_{\mathcal{X}}). \quad (122)$$

GHENT UNIVERSITY, SYSTEMS RESEARCH GROUP, TECHNOLOGIEPARK-ZWIJNAARDE 914, 9052  
ZWIJNAARDE, BELGIUM

*E-mail address:* Gert.deCooman@UGent.be

GHENT UNIVERSITY, SYSTEMS RESEARCH GROUP, TECHNOLOGIEPARK-ZWIJNAARDE 914, 9052  
ZWIJNAARDE, BELGIUM

*E-mail address:* Erik.Quaeghebur@UGent.be