ARTICLE IN PRESS
FSS:6674

Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems ••• (••••) •••–•••

FUZZY sets and systems

www.elsevier.com/locate/fss

Representation theorems for partially exchangeable random variables

Jasper De Bock [a,*], Arthur Van Camp [a], Márcio A. Diniz [b], Gert de Cooman [a]

[a] *Ghent University, SYSTeMS Research Group, Technologiepark-Zwijnaarde 914, 9052 Zwijnaarde, Belgium*
[b] *Federal University of S. Carlos, Dept. of Statistics, Rod. Washington Luis, km 235, S. Carlos, Brazil*

Received 10 February 2014; received in revised form 27 August 2014; accepted 31 October 2014

Abstract

We provide representation theorems for both finite and countable sequences of finite-valued random variables that are considered to be partially exchangeable. In their most general form, our results are presented in terms of sets of desirable gambles, a very general framework for modelling uncertainty. Its key advantages are that it allows for imprecision, is more expressive than almost every other imprecise-probabilistic framework and makes conditioning on events with (lower) probability zero non-problematic. We translate our results to more conventional, although less general frameworks as well: lower previsions, linear previsions and probability measures. The usual, precise-probabilistic representation theorems for partially exchangeable random variables are obtained as special cases.

Keywords: Partial exchangeability; Sets of desirable gambles; Lower previsions; de Finetti's representation theorem; Indifferent gambles

1. Introduction

The objective of this paper is to model finite as well as countable sequences of finite-valued random variables that are considered to be partially exchangeable. We derive representation theorems for such variables, in the style of de Finetti, but within a more general framework: that of imprecise probabilities. The usual precise-probabilistic representation theorems are recovered as special cases. Since partial exchangeability has never before been discussed within such a general framework, we mainly focus on theoretical aspects, discussing both mathematical and philosophical issues in a rather high level of detail. The practical advantages of using our imprecise-probabilistic notion of partial exchangeability in an applied setting are briefly discussed in our conclusions. However, actual applications to statistical problems are left for future work.

* Corresponding author.
E-mail addresses: jasper.debock@ugent.be (J. De Bock), arthur.vancamp@ugent.be (A. Van Camp), marcio.diniz@ugent.be (M.A. Diniz), gert.decooman@ugent.be (G. de Cooman).

http://dx.doi.org/10.1016/j.fss.2014.10.027
0165-0114/

Loosely speaking, making a judgement of partial exchangeability means that the order in which certain groups of variables are observed is deemed irrelevant. De Finetti introduced this as a generalisation of (regular) exchangeability, useful in situations where a judgement of complete symmetry between all variables is unrealistic [1]. He proposed the following example: two people are tossing coins, or the same person tosses the same coin under two different conditions of temperature, atmospheric pressure, and so on. In this case, it makes sense to judge the tosses made by the same person, or under the same conditions, exchangeable. More generally, it may be possible to divide the experiments into g types that are considered exchangeable only with the other experiments of the same type. In that case, the corresponding variables are called g-fold partially exchangeable [2], which is the kind of partial exchangeability we consider in this paper.[1] For $g = 1$, it reduces to regular exchangeability.

The most general framework in which we will study this concept of partial exchangeability is that of sets of desirable gambles; however, we translate our results to other, less expressive frameworks as well: lower previsions, linear previsions and probability measures. The central idea within the theory of sets of desirable gambles is to model a subject's beliefs by considering the set of gambles—bets—that he finds desirable, in the sense that he prefers them over the status quo—no bet at all. Based on the ideas of de Finetti [8], the main concepts behind this theory were originally introduced by Smith [9] and Williams [10]. Instead of considering the two-sided bets of de Finetti, they used one-sided bets, thereby allowing them to move from linear subspaces of bets to general cones. Later, Walley [11] further developed the theory and gave it its present name. For recent work on sets of desirable gambles, see for example Refs. [12–18].

Although sets of desirable gambles are not as well known as other probabilistic models, they have clear advantages. First of all, they allow for imprecisely specified probabilities. Loosely speaking: lower and upper probabilities; if both coincide, we obtain the usual case. As such, sets of desirable gambles can be used to model imprecision, indecision and partial or complete ignorance, all of which cannot be adequately dealt with using classical probability theory [11, Chapter 5]. Secondly, within the theory of imprecise probabilities, sets of desirable gambles are one of the most expressive models available: lower and upper previsions, lower and upper probabilities, belief functions, possibility measures and necessity measures can all be regarded as special cases [19]. Thirdly, and related to the second advantage, conditioning a set of desirable gambles is non-problematic, even if the conditioning event has (lower) probability zero [16,19]. As we will argue in the conclusions, these advantages are particularly relevant to the present subject: partial exchangeability.

The idea of using sets of desirable gambles to model a structural assessment of symmetry, such as partial exchangeability, is not new. In Ref. [20], one of the authors conducted a general study on how to model symmetry assessments through sets of desirable gambles and in Ref. [15], the particular case of regular exchangeability was covered in detail. Similar studies have been conducted for lower previsions as well; see for example Refs. [20,21] and [11, Section 9.5]. Our main contribution consists in applying these ideas to the more involved case of partial exchangeability, thereby generalising previous results on regular exchangeability, mainly those in Refs. [15,21]. Besides this generalisation to partial exchangeability, there are some other notable differences with this previous work as well.[2] Although—for the case of regular exchangeability—they lead to mathematically equivalent results, we consider our approach to be more elegant, as well as more intuitive. Of course, we leave this to the reader to decide.

This paper is organised as follows. We start in Section 2 with an introduction to sets of desirable gambles and lower previsions, and we relate the latter to expectation operators and probability measures. This section also includes a fresh look at the concept of indifferent gambles. In Section 3, we introduce our notation for (multiple) sequences of random variables, and we explain how to model a subject's beliefs about these variables by means of the tools discussed in Section 2. Then, in Section 4, we provide a definition of partial exchangeability in terms of assessments of indifference, and we present finite representation theorems, which are stated in terms of count vectors and an operator related to the multivariate hypergeometric distribution. We also introduce polynomial gambles, explain how these are related to multivariate Bernstein polynomials, and translate our finite representation theorems to this framework. In Section 5, we move from finite to countable partial exchangeability and provide countable versions of our representation theorems. Here, the representation is expressed in terms of polynomial gambles, which are related to the random variables that we are modelling by means of an operator that is connected to the multinomial distribution. We comment on the

[1] There are other types of partial exchangeability; see for example Refs. [3–7].

[2] We define (partial) exchangeability in terms of indifferent gambles (Section 2.3) rather than weakly desirable ones. Also, we model countable sequences by means of gambles of finite structure (Section 3.2) instead of using a time-consistent family of finite models.

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

differences with more conventional (often measure-theoretic) representation theorems, and discuss the advantages of our approach. Section 6 closes with some remarks and perspectives for future research. We discuss the relevance of our results from a more applied point of view, and explain how our representation theorems could serve as a first step in developing an imprecise-probabilistic notion of predictive inference under partial exchangeability. In order to make our main argumentation as readable as possible, we have moved all technical proofs to Appendix A, which also contains a number of supplementary lemmas.

2. Sets of desirable gambles and related concepts

Consider a random variable X—for example, the outcome of some experiment—that assumes values in some non-empty possibility space Ω.[3] In the present section, we discuss a number of different but closely related frameworks that can be used to model a subject's uncertainty associated with the value of X—the outcome of the experiment. We start with sets of desirable and sets of indifferent gambles—which constitute the most general framework considered here—and go on to discuss derived concepts such as coherent lower previsions, linear previsions and (finitely additive) probability measures.

2.1. Basic nomenclature: gambles

A *gamble* f is a bounded real-valued function on Ω. It is interpreted as an uncertain reward: if the value of X turns out to be ω, the gamble f results in a—positive or negative—payoff $f(\omega)$, expressed in some predetermined linear utility scale. We denote the set of all gambles on Ω as $\mathcal{G}(\Omega)$. It is a linear space under pointwise addition of gambles and pointwise multiplication of gambles with real numbers.

For any two f_1 and f_2 in $\mathcal{G}(\Omega)$, we write '$f_1 \geq f_2$' if $(\forall\omega \in \Omega)\, f_1(\omega) \geq f_2(\omega)$ and '$f_1 > f_2$' if $f_1 \geq f_2$ and $f_1 \neq f_2$. As an example: for any $f \in \mathcal{G}(\Omega)$, we write '$f \geq 0$' if f is non-negative and '$f > 0$' if, additionally, $f(\omega) > 0$ for at least one $\omega \in \Omega$. Subsets of $\mathcal{G}(\Omega)$ are denoted by using predicates as subscripts; e.g., $\mathcal{G}(\Omega)_{\geq 0} := \{f \in \mathcal{G}(\Omega)\colon f \geq 0\}$ is the set of all non-negative gambles on Ω. We refer to subsets of linear subspaces $\mathcal{K}$ of $\mathcal{G}(\Omega)$ in a similar way; e.g., $\mathcal{K}_{>0} := \{f \in \mathcal{K}\colon f > 0\}$. Finally, for any subsets $\mathcal{A}$, $\mathcal{A}_1$ and $\mathcal{A}_2$ of $\mathcal{G}(\Omega)$, $\operatorname{span}(\mathcal{A})$ is the set of all finite linear combinations of gambles in $\mathcal{A}$ and the (Minkowski) sum of $\mathcal{A}_1$ and $\mathcal{A}_2$ is defined as $\mathcal{A}_1 + \mathcal{A}_2 := \{f_1 + f_2\colon f_1 \in \mathcal{A}_1, f_2 \in \mathcal{A}_2\}$.

2.2. Sets of desirable gambles

As a basic tool to model a subject's beliefs about the value of X, we consider a set $\mathcal{D}$ of gambles that he finds desirable. A subject is said to find a gamble $f \in \mathcal{G}(\Omega)$ *desirable* if he prefers it to the zero gamble—the status quo. By this we mean that he is willing to engage in a transaction where (i) the actual value $\omega \in \Omega$ of X is determined and (ii) he receives the payoff $f(\omega)$. Even stronger, he prefers this transaction to the status quo—no transaction at all. A set of desirable gambles is considered to represent a rational subject's beliefs if it is *coherent*.

Definition 1 *(Coherence for sets of desirable gambles).* Consider any linear subspace $\mathcal{K}$ of $\mathcal{G}(\Omega)$. A set of desirable gambles $\mathcal{D} \subseteq \mathcal{K}$ is called *coherent relative to* $\mathcal{K}$ if

D1. $0 \notin \mathcal{D}$,
D2. $\mathcal{K}_{>0} \subseteq \mathcal{D}$, [desiring partial gain]
D3. $\lambda f \in \mathcal{D}$ for all $f \in \mathcal{D}$ and all $\lambda \in \mathbb{R}_{>0}$, [positive scaling]
D4. $f_1 + f_2 \in \mathcal{D}$ for all $f_1, f_2 \in \mathcal{D}$ [combination]

and, consequently [use D1, D2 and D4]

D5. $\mathcal{K}_{\leq 0} \cap \mathcal{D} = \emptyset$.

[3] By random, we mean that its value is possibly uncertain or unknown. Unlike what is commonly done in statistics, the "values" of which we speak are not required to be numerical, nor do we impose some kind of measurability condition.

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

Observe that we define coherence *relative* to some linear subspace $\mathcal{K}$ of $\mathcal{G}(\Omega)$,[4] which allows us to focus on linear subspaces of $\mathcal{G}(\Omega)$ that are of particular interest, as we will do in, for example, Section 3.2. For $\mathcal{K} = \mathcal{G}(\Omega)$, our definition coincides with the more standard one; see for example Refs. [17,19]. In this case, we drop the words "relative to $\mathcal{K}$", which we also do when $\mathcal{K}$ is clear from the context.

2.3. Sets of indifferent gambles

In addition to a subject's set $\mathcal{D}$ of desirable gambles—the ones he prefers to the zero gamble—we can also consider the gambles that he considers to be *equivalent* to the zero gamble. We call these gambles *indifferent* and collect them in a set $\mathcal{I}$ of indifferent gambles.[5] Any reasonable set of indifferent gambles should satisfy at least the following four rationality criteria, each of which follows intuitively from the interpretation of $\mathcal{I}$:

I1. $0 \in \mathcal{I}$,[6]
I2. $\lambda f \in \mathcal{I}$ for all $f \in \mathcal{I}$ and all $\lambda \in \mathbb{R}$, [scaling]
I3. $f_1 + f_2 \in \mathcal{I}$ for all $f_1, f_2 \in \mathcal{I}$, [combination]
I4. $f \notin \mathcal{I}$ if $f > 0$ or $f < 0$.

Together, requirements I1–I3 are equivalent to imposing that $\mathcal{I}$ should be a linear subspace of $\mathcal{G}(\Omega)$.

The interaction between indifferent and desirable gambles is subject to rationality criteria as well. The most important such criterion is what we call *compatibility* of $\mathcal{D}$ and $\mathcal{I}$.

Definition 2 *(Compatibility of $\mathcal{D}$ and $\mathcal{I}$).* Consider any linear subspace $\mathcal{K}$ of $\mathcal{G}(\Omega)$, any set of desirable gambles $\mathcal{D} \subseteq \mathcal{K}$ that is coherent relative to $\mathcal{K}$, and any set of indifferent gambles $\mathcal{I} \subseteq \mathcal{K}$ that satisfies I1–I4. Then $\mathcal{D}$ is said to be compatible with $\mathcal{I}$ if

ID1. $f_1 + f_2 \in \mathcal{D}$ for all $f_1 \in \mathcal{D}$ and all $f_2 \in \mathcal{I}$. [compatibility]

Simply put, adding an indifferent gamble to a desirable one should result in a desirable gamble. Alternatively, this can be formulated as $\mathcal{D} + \mathcal{I} \subseteq \mathcal{D}$. Since, due to I1, we also have that $\mathcal{D} \subseteq \mathcal{D} + \mathcal{I}$, compatibility is equivalent to $\mathcal{D} + \mathcal{I} = \mathcal{D}$. Hence, the indifferent gambles—similar to the zero gamble—are neutral elements with respect to Minkowski addition. We invite the interested reader to check that, by combining ID1 with D1 and I2, one can derive the following additional rationality criterion as well:

ID2. $\mathcal{D} \cap \mathcal{I} = \emptyset$.

This means that a rational subject should not consider a gamble as both desirable—preferred to the zero gamble—and indifferent—equivalent to the zero gamble—at the same time.

2.4. Coherent lower previsions

When presented with a subject's set $\mathcal{D}$ of desirable gambles, we can use it to derive his supremum buying price $\underline{P}_{\mathcal{D}}(f)$—his *lower prevision*—and infimum selling price $\overline{P}_{\mathcal{D}}(f)$—his *upper prevision*—for any gamble f on Ω [11]:

$$\underline{P}_{\mathcal{D}}(f) = \sup\{\mu \in \mathbb{R}: f - \mu \in \mathcal{D}\} \quad \text{and} \quad \overline{P}_{\mathcal{D}}(f) = \inf\{\mu \in \mathbb{R}: \mu - f \in \mathcal{D}\}. \tag{1}$$

Instead of deriving them from a set of desirable gambles, these lower and upper previsions can also be given directly, in which case we drop the subscript $\mathcal{D}$ and simply write $\underline{P}(f)$ and $\overline{P}(f)$. In any case, due to their interpretation as a supremum buying price and infimum selling price, $\underline{P}(f)$ and $\overline{P}(f)$ should be related trough conjugacy: $\overline{P}(f) =$

[4] This definition is a special case of the one given in Ref. [15, Definition 1], where the authors define *coherence relative to* $(\mathcal{K}, \mathcal{C})$. There, $\mathcal{C} \subset \mathcal{K}$ is a convex cone that defines an order on all gambles. In this paper, our definition amounts to the special case for which $\mathcal{C} = \mathcal{K}_{\geq 0}$. Similar definitions have been suggested by Walley as well; see Ref. [11].

[5] The concept of indifferent gambles, as we introduce it in the present section, is a simplified version of an even more expressive framework, tailored towards our specific needs. For more information, we refer to Ref. [22] for a general overview of indifference and its connection to other types of assessments, and to Ref. [23, Chapter 1] for a study about reasoning with indifference assessments.

[6] At first sight, it might seem as if this criterion is redundant because it is trivially implied by I2. However, this is only true if $\mathcal{I} \neq \emptyset$.

$-\underline{P}(-f)$.[7] Given this connection, one can restrict attention to either one of them; in our case, we choose to focus on lower previsions.

Definition 3 *(Coherence for lower previsions).* A real-valued functional $\underline{P}$, defined on some linear subspace $\mathcal{K}$ of $\mathcal{G}(\Omega)$,[8] is called *coherent* if

P1. $\underline{P}(f) \geq \inf f$ for all $f \in \mathcal{K}$, [accepting sure gain]
P2. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ for all $f \in \mathcal{K}$ and all $\lambda \in \mathbb{R}_{>0}$, [positive homogeneity]
P3. $\underline{P}(f_1 + f_2) \geq \underline{P}(f_1) + \underline{P}(f_2)$ for all $f_1, f_2 \in \mathcal{K}$. [super-additivity]

Again, we drop the reference to $\mathcal{K}$ whenever $\mathcal{K}$ is either clear from the context or equal to $\mathcal{G}(\Omega)$. The following properties are direct consequences of coherence [11, Section 2.6.1], whenever the gambles involved are elements of $\mathcal{K}$:

P4. $\inf f \leq \underline{P}(f) \leq \overline{P}(f) \leq \sup f$,
P5. $\underline{P}(f + \mu) = \underline{P}(f) + \mu$ and $\underline{P}(\mu) = \mu$ for any constant gamble $\mu \in \mathbb{R}$,
P6. If $\sup |f_n - f| \to 0$ as $n \to \infty$, then $\underline{P}(f_n) \to \underline{P}(f)$.

Coherence of a lower prevision is closely related to the corresponding notion for sets of desirable gambles: if $\mathcal{D}$ is coherent, $\underline{P}_{\mathcal{D}}$ will be as well and, for every coherent $\underline{P}$, there is some coherent $\mathcal{D}$ such that $\underline{P} = \underline{P}_{\mathcal{D}}$.

Proposition 1. *Consider any linear subspace $\mathcal{K}$ of $\mathcal{G}(\Omega)$ that includes all constant gambles. A lower prevision $\underline{P}$ on $\mathcal{K}$ is coherent if and only if there is some set of desirable gambles $\mathcal{D} \subseteq \mathcal{K}$ that is coherent relative to $\mathcal{K}$ and for which $\underline{P} = \underline{P}_{\mathcal{D}}$.*[9]

Compatibility of a coherent lower prevision with a set of indifferent gambles is defined as follows.

Definition 4 *(Compatibility of $\underline{P}$ and $\mathcal{I}$).* Consider any linear subspace $\mathcal{K}$ of $\mathcal{G}(\Omega)$, any coherent lower prevision $\underline{P}$ on $\mathcal{K}$, and any set of indifferent gambles $\mathcal{I} \subseteq \mathcal{K}$ that satisfies I1–I4. Then $\underline{P}$ is said to be compatible with $\mathcal{I}$ if any (and hence all) of the following statements holds:

IP1. $\underline{P}(f) = \overline{P}(f) = 0$ for all $f \in \mathcal{I}$; [compatibility]
IP2. $\underline{P}(f) \geq 0$ for all $f \in \mathcal{I}$;
IP3. $\underline{P}(f + g) = \underline{P}(g)$ for all $f \in \mathcal{I}$ and $g \in \mathcal{K}$.

The proofs of these equivalences are trivial, for instance, for the first two conditions, use conjugacy, P4 and I2.

The intuition behind Definition 4 is that, if a rational subject considers a gamble equivalent to the zero gamble, then his supremum buying price and infimum selling price for this gamble should both be zero. For $f = 0$, IP1 holds trivially because of P4. Alternatively, as follows from our next result, Definition 4 can be derived from Definition 2 as well.

Proposition 2. *Consider any linear subspace $\mathcal{K}$ of $\mathcal{G}(\Omega)$ that includes all constant gambles and any set of indifferent gambles $\mathcal{I} \subseteq \mathcal{K}$ that satisfies I1–I4. Then a lower prevision $\underline{P}$ on $\mathcal{K}$ is coherent and compatible with $\mathcal{I}$ if and only if there is a set of desirable gambles $\mathcal{D} \subseteq \mathcal{K}$ that is coherent relative to $\mathcal{K}$ and compatible with $\mathcal{I}$, and for which $\underline{P} = \underline{P}_{\mathcal{D}}$.*

2.5. Linear previsions

If for some gamble f on Ω, a subject's lower and upper prevision coincide, then $P(f) := \underline{P}(f) = \overline{P}(f)$ is his *fair price* for that gamble: he is willing to buy f for any price strictly lower than $P(f)$ and sell it for any price strictly

[7] Conjugacy does not hinge on the subjective interpretation that is adopted in this paper. See Section 2.5 for an alternative interpretation, which naturally leads to conjugacy as well.

[8] If $\mathcal{K}$ is not a linear subspace, coherence can still be defined [11]. However, this comes at the expense of a more elaborate definition. If $\mathcal{K}$ is a linear subspace—which will always be the case in the present paper, this general definition can be shown to reduce to requirements P1–P3.

[9] Similar results, which use slightly different notions of desirability, can be found in, amongst others, Ref. [11, Section 3.8.1]. For completeness, and due to these small differences, we choose to provide a proof of our own.

higher than $P(f)$. Following de Finetti, we refer to $P(f)$ as the subject's *prevision* for f [8]. If this happens for all $f \in \mathcal{K}$, with $\mathcal{K}$ some linear subspace of $\mathcal{G}(\Omega)$, one obtains a real-valued operator P on $\mathcal{K}$ that is self-conjugate: $P(f) = -P(-f)$ for all $f \in \mathcal{K}$. If P is furthermore coherent (as a lower prevision), then this operator will be a real linear functional, and it is then called a *linear prevision* on $\mathcal{K}$ [11, Section 2.8]. We use $\mathbb{P}(\mathcal{K})$ to denote the set of all linear previsions on $\mathcal{K}$.

Interestingly, linear previsions are tightly connected to finitely additive probabilities. Every linear prevision is the expectation operator—or, in case $\mathcal{K} \neq \mathcal{G}(\Omega)$, its restriction to $\mathcal{K}$—of a finitely, but not necessarily countably additive probability measure [11, Section 3.2]. For every event E—some subset of Ω, its probability $P(E)$ is given by the prevision of the indicator $\mathbb{I}_E$, which is a gamble on Ω that assumes the value 1 on E and 0 elsewhere. If $\mathcal{K} = \mathcal{G}(\Omega)$, this probability measure is furthermore unique.[10]

Linear previsions—and consequently, finitely additive probability measures—are closely related to coherent lower previsions, and not only because the former are a special case of the latter: every coherent lower prevision is a lower envelope of linear ones.

Theorem 3 *(Lower envelope theorem). (See [11, Section 3.3.3].) A lower prevision $\underline{P}$ on $\mathcal{K}$, with $\mathcal{K}$ a linear subspace of $\mathcal{G}(\Omega)$, is coherent if and only if it is the lower envelope of the (convex set of) linear previsions that dominate it:*[11]

$$\underline{P}(f) = \min\big\{P(f) : P \in \mathcal{M}(\underline{P})\big\} \quad \text{for all } f \text{ in } \mathcal{K}, \tag{2}$$

with

$$\mathcal{M}(\underline{P}) := \big\{P \in \mathbb{P}(\mathcal{K}) : (\forall f \in \mathcal{K})\, P(f) \geq \underline{P}(f)\big\}. \tag{3}$$

Furthermore, since coherence is preserved under taking lower envelopes [11, Section 2.6.3(b)], we have in particular that the lower envelope of *any* set of linear previsions will be a coherent lower prevision. Hence, we conclude that a lower prevision is coherent if and only if it is the lower envelope of a set of linear previsions or, equivalently, the lower envelope of a set of expectation operators.

Due to this result, the reader can, if wanted, interpret the coherent lower previsions in this paper as lower envelopes of expectation operators. This so-called *sensitivity analysis interpretation* serves as an alternative to the behavioural interpretation in terms of supremum buying prices [11, Section 2.10.4], which we introduced in Section 2.4. Let us illustrate the difference between these two interpretations by applying them to the concept of conjugacy: $\overline{P}(f) = -\underline{P}(-f)$. Under the behavioural interpretation, and as we saw in Section 2.4, conjugacy is motivated by the fact that the supremum buying price of $-f$ should be equal to minus the infimum selling price of f. Under the sensitivity analysis interpretation, conjugacy is motivated by the fact that it is equivalent to adopting the following expression for the upper prevision:

$$\overline{P}(f) = -\underline{P}(-f) = \max\big\{P(f) : P \in \mathcal{M}(\underline{P})\big\} \quad \text{for all } f \text{ in } \mathcal{K}.$$

In other words, the conjugate of the lower envelope of a set of linear previsions is its upper envelope. As such, under the sensitivity analysis interpretation, we can interpret upper previsions as upper envelopes of expectation operators.

Although we prefer, and mainly focus on the behavioural interpretation—in part, due to its close connection with sets of desirable gambles, which we consider as more fundamental—we want to stress that all of our results on coherent lower previsions, and in particular our representation theorems, are valid regardless of their interpretation.

3. Modelling countable sequences of variables

Sets of desirable gambles and coherent lower previsions are versatile tools. By choosing the possibility space Ω and domain $\mathcal{K}$ appropriately, it is possible to cover a diverse range of settings. In the present section, we show how they can be used to model countable sequences of variables. The fact that partial exchangeability deals with multiple

[10] As illustrated in Section 5.1, uniqueness can sometimes be obtained for cases where $\mathcal{K} \neq \mathcal{G}(\Omega)$ as well, by restricting the domain of the probability measure.

[11] Technically, in order for it to be a lower envelope, the minimum in Eq. (2) should be an infimum. However, in this case, Walley has shown that the infimum is always attained and, therefore, it can be replaced by a minimum [11, Section 3.3.3].

such sequences at the same time, gives rise to some notational challenges. Most papers on partial exchangeability tend to avoid these, by using as little notation as possible. However, given the theoretical nature of this paper, the generality of our imprecise-probabilistic framework, and the high level of formalism that is required for the proofs, we have chosen not to follow this tradition. Instead, we introduce a highly detailed notation.

3.1. Countable sequences of variables: notational conventions

Consider $g \in \mathbb{N}$ countable sequences $X_{i1}, \dots, X_{ij}, \dots$ of random variables that assume values in a finite, non-empty possibility space $\mathcal{X}_i$. The first index i is used to refer to the particular sequence and takes values in $G := \{1, \dots, g\}$. Variables within the same sequence—with the same i—are said to belong to the same *group* or, alternatively, to be of the same *type*. The second index j refers to the different variables within a single sequence—a group—and takes values in $\mathbb{N}$. For every $i \in G$ and $j \in \mathbb{N}$, the random variable X_{ij} assumes values in its possibility space $\mathcal{X}_{ij} = \mathcal{X}_i$. Generic elements of $\mathcal{X}_{ij}$ are denoted as x_{ij}. If X_{ij} assumes some value $x_{ij} \in \mathcal{X}_i$, we write $X_{ij} = x_{ij}$. In practice, $\mathcal{X}_i$ is often taken to be the same set $\mathcal{X}$ for all $i \in G$. However, there is no *a priori* reason, in principle, to restrict the framework to this case.

Running example. Consider a clinical trial in which some new drug is tested on both men and women. We use $i = 1$ and $i = 2$ as indices to refer to the experiments on men and women respectively. For a single test person, the experiment may result in just one of two states: "cured" or "not cured". Therefore, we have $g = 2$, $G = \{1, 2\}$ and $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$, denoting "cured" by 1. Since both state spaces are identical, we simply say that $\mathcal{X} = \{0, 1\}$. The random variable X_{23} represents the outcome of the experiment on the third woman and may take any value in $\mathcal{X}_{23} = \mathcal{X} = \{0, 1\}$. After conducting the experiment, one could find that $X_{23} = 1$, meaning that the third woman was "cured". □

For any $i \in G$ and $J_i \subseteq \mathbb{N}$, we denote by $X_{i(J_i)}$ the vector that has the random variables X_{ij}, $j \in J_i$, as its elements, ordered with respect to the indices j, in increasing order. This vector $X_{i(J_i)}$ can be regarded as a single random variable that takes values $x_{i(J_i)}$ in the possibility space $\mathcal{X}_{i(J_i)} := \bigtimes_{j \in J_i} \mathcal{X}_{ij}$. If J_i is a singleton, consisting of a single element $j \in \mathbb{N}$, then we obtain $X_{i(\{j\})} = X_{ij}$ as a special case. As another example, consider the set $N_i := \{1, \dots, n_i\}$, with $n_i \in \mathbb{N}$, for which we obtain a random variable $X_{i(N_i)} = (X_{i1}, \dots, X_{in_i})$ that takes values in $\mathcal{X}_{i(N_i)} = \bigtimes_{j=1}^{n_i} \mathcal{X}_{ij} = \mathcal{X}_i^{n_i}$. Here, generic values are denoted as $x_{i(N_i)} = (x_{i1}, \dots, x_{in_i})$. Finally, the complete sequence corresponds to $J_i = \mathbb{N}$, resulting in a random variable $X_{i(\mathbb{N})} = (X_{i1}, \dots, X_{ij}, \dots)$ that assumes values in $\mathcal{X}_{i(\mathbb{N})} = \bigtimes_{j \in \mathbb{N}} \mathcal{X}_{ij} = \mathcal{X}_i^{\mathbb{N}}$. We will denote generic elements of this set as $x_{i(\mathbb{N})} = (x_{i1}, \dots, x_{ij}, \dots)$.

Running example. Consider again our running example and suppose that four men have been tested so far, yielding the outcomes $X_{11} = 0$, $X_{12} = 1$, $X_{13} = 1$ and $X_{14} = 0$. If we choose $n_1 = 4$ and use our convention that $N_1 := \{1, \dots, n_1\}$, this can be denoted compactly as $X_{1(N_1)} = (0, 1, 1, 0)$. For this particular outcome, and with $J_1 = \{2, 4\}$, we obtain $X_{1(J_1)} = (X_{12}, X_{14}) = (1, 0)$. In much the same way, the (potentially) infinite sequence of experiments could result in $X_{1(\mathbb{N})} = (0, 1, 1, 0, 0, 1, 0, 1, \dots)$. □

We extend this notation towards multiple sequences as follows. We denote by $X_{G(J)}$ the tuple of variables (with one component $X_{i(J_i)}$ for every $i \in G$) that takes values in $\mathcal{X}_{G(J)} = \bigtimes_{i \in G} \mathcal{X}_{i(J_i)}$. With a slight abuse of notation, this also allows us to use $X_{G(N)}$ to refer to the tuple consisting of the variables $X_{i(N_i)}$, $i \in G$. Similarly, we use $X_{G(\mathbb{N})}$ to refer to the tuple consisting of the variables $X_{i(\mathbb{N})}$, $i \in G$, which is perhaps the most important special case. Indeed, this results in a single random variable $X_{G(\mathbb{N})}$, representing all the variables X_{ij}, $i \in G$ and $j \in \mathbb{N}$, in a very compact manner.

Running example. Suppose that apart from the four men, so far, three women have been tested as well. The outcomes of these additional experiments were $X_{21} = 1$, $X_{22} = 1$ and $X_{23} = 0$. With $n_2 = 3$, and using the convention that $N_2 = \{1, \dots, n_2\}$, this can be denoted compactly as $X_{2(N_2)} = (1, 1, 0)$. We refer to all seven experiments at once by writing $X_{G(N)}$, which, using our notational conventions, is equal to the tuple $(X_{1(N_1)}, X_{2(N_2)})$. As a final example, consider $J_1 = \{2, 4\}$ and $J_2 = \{2, 3\}$. Then $X_{G(J)} = (X_{1(J_1)}, X_{2(J_2)})$, with $X_{1(J_1)}$ as given earlier on and $X_{2(J_2)} = (X_{22}, X_{23}) = (1, 0)$. □

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

3.2. Modelling beliefs about countable sequences of variables

As explained in the previous section, sequences of variables can be regarded as a single (joint) variable as well. Hence, we can easily use the tools from Section 2 to model a subject's beliefs about such sequences. For the joint variable $X_{G(\mathbb{N})}$, the most straightforward approach would be to use a coherent set $\mathcal{D}_{G(\mathbb{N})}$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ or a coherent lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\mathcal{G}(\mathcal{X}_{G(\mathbb{N})})$.

However, we believe that this would make little sense from a behavioural point of view. Let us consider a single countable sequence of variables, each of which corresponds to a coin toss. Does it make sense to bet on the event that the outcome of *every* toss will be heads? We follow de Finetti in thinking that it does not: it is impossible to observe the outcome of *every* coin toss, as there are infinitely many. It does however make sense to bet on the event that the outcome of the first n coin tosses is heads, and one can do so for any $n \in \mathbb{N}$. In much the same way, it also makes sense to consider bets on any finite subset of a countable sequence of coin tosses. For this reason, we do not consider all the gambles in $\mathcal{G}(\mathcal{X}_{G(\mathbb{N})})$ as relevant. We will restrict ourselves to those gambles that are of *finite structure*,[12] meaning that they depend upon the value of a *finite subset* of the variables X_{ij} only, with $i \in G$ and $j \in \mathbb{N}$, rather than the value of $X_{G(\mathbb{N})}$, which depends upon all of them. We denote the set of all gambles of finite structure by $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$.

In order to be able to formally define this set, we introduce an important simplifying device called *cylindrical extension*. For all $i \in G$, let J_i and J_i' be disjoint subsets of $\mathbb{N}$ and denote their union as J_i^*. Then for any gamble f on $\mathcal{X}_{G(J)}$, its so-called cylindrical extension $\tilde{f}$ to $\mathcal{X}_{G(J^*)}$ is defined as follows:

$$\tilde{f}(x_{G(J^*)}) = \tilde{f}(x_{G(J)}, x_{G(J')}) := f(x_{G(J)}) \quad \text{for all } x_{G(J^*)} \in \mathcal{X}_{G(J^*)}. \tag{4}$$

Formally, f belongs to $\mathcal{G}(\mathcal{X}_{G(J)})$, whereas $\tilde{f}$ belongs to $\mathcal{G}(\mathcal{X}_{G(J^*)})$. In practice however, both gambles clearly coincide: they both depend upon the value of $X_{G(J)}$ only and, as such, correspond to the same bet. *We will therefore, in the sequel, repeatedly identify f with its cylindrical extension $\tilde{f}$, in which case we denote this extension by f as well.* Using this convention, we can for example identify $\mathcal{G}(\mathcal{X}_{G(J)})$ with a subset of $\mathcal{G}(\mathcal{X}_{G(\mathbb{N})})$. Similarly, for any $\mathcal{K} \subseteq \mathcal{G}(\mathcal{X}_{G(\mathbb{N})})$, we can write $\mathcal{K} \cap \mathcal{G}(\mathcal{X}_{G(J)})$ to denote the set of those gambles in $\mathcal{K}$ that depend upon the value of $X_{G(J)}$ only.

The set of all gambles of finite structure can now be defined as follows:

$$\begin{aligned} \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})}) &:= \left\{ f \in \mathcal{G}(\mathcal{X}_{G(\mathbb{N})}) \colon f \in \mathcal{G}(\mathcal{X}_{G(J)}) \text{ for some } J_i \Subset \mathbb{N}, i \in G \right\} \\ &= \bigcup_{J_i \Subset \mathbb{N}, i \in G} \mathcal{G}(\mathcal{X}_{G(J)}), \end{aligned} \tag{5}$$

where $J_i \Subset \mathbb{N}$ is taken to mean that J_i is a finite subset of $\mathbb{N}$. In other words, every gamble in $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is the cylindrical extension to $\mathcal{X}_{G(\mathbb{N})}$ of some gamble on $\mathcal{X}_{G(J)}$, where for all $i \in G$, J_i is an arbitrary but finite subset of $\mathbb{N}$. Clearly, $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is a linear subspace of $\mathcal{G}(\mathcal{X}_{G(\mathbb{N})})$ that includes all constant gambles. Therefore, all the results of Section 2 apply. In particular, we can model a subject's beliefs about $X_{G(\mathbb{N})}$ by means of a set of desirable gambles $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ that is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ or, alternatively, by means of a coherent lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$.

Now that we know how to model all variables at once, let us see what happens if we focus on a subset. Consider, for all $i \in G$, some $J_i \Subset \mathbb{N}$ or, in other words, choose a finite number of variables from every sequence. Then we can model a subject's beliefs about these variables by means of a coherent set $\mathcal{D}_{G(J)}$ of desirable gambles on $\mathcal{X}_{G(J)}$ or a coherent lower prevision on $\underline{P}_{G(J)}$ on $\mathcal{G}(\mathcal{X}_{G(J)})$. If besides a model for $X_{G(J)}$, we also have a joint model for $X_{G(\mathbb{N})}$, then both of these should be related. For example, if a subject considers a gamble f on $\mathcal{X}_{G(J)}$ as desirable, he should consider its cylindrical extension to $\mathcal{X}_{G(\mathbb{N})}$ as desirable as well, since, in practice, these gambles are indistinguishable. More generally, $\mathcal{D}_{G(J)}$ should be the marginalisation of $\mathcal{D}_{G(\mathbb{N})}$ to $\mathcal{G}(\mathcal{X}_{G(J)})$, as defined by[13]

$$\operatorname{marg}_{G(J)} \mathcal{D}_{G(\mathbb{N})} := \left\{ f \in \mathcal{G}(\mathcal{X}_{G(J)}) \colon f \in \mathcal{D}_{G(\mathbb{N})} \right\} = \mathcal{D}_{G(\mathbb{N})} \cap \mathcal{G}(\mathcal{X}_{G(J)}). \tag{6}$$

[12] In choosing this terminology, we were inspired by Dubins and Savage [24, Chapter 2.7], who introduce functions of finite structure as functions that depend upon only a finite subset of a countable set of coordinates.

[13] Ref. [17] includes a general introduction to marginalisation for sets of desirable gambles.

More or less the same can be said for lower previsions. Let $\underline{P}_{G(J)}$ be a lower prevision on $\mathcal{G}(\mathcal{X}_{G(J)})$ and $\underline{P}_{G(\mathbb{N})}$ be a lower prevision on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$. Then these are said to be related trough marginalisation if

$$\underline{P}_{G(J)}(f) = \underline{P}_{G(\mathbb{N})}(f) \quad \text{for all } f \text{ in } \mathcal{G}(\mathcal{X}_{G(J)}). \tag{7}$$

As is to be expected from a rationality requirement, coherence is preserved under marginalisation. Within our finitary context, where we only consider gambles of finite structure, we can even establish a converse result.

Proposition 4. *A set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ if and only if for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, the marginal set of desirable gambles $\mathcal{D}_{G(J)}$, as given by Eq. (6), is coherent.*

A similar result holds for coherent lower previsions as well.

Proposition 5. *A lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is coherent if and only if for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, the marginal lower prevision $\underline{P}_{G(J)}$, as given by Eq. (7), is coherent.*

4. Finite partially exchangeable sequences

Armed with the tools introduced in the two previous sections, we can now start with the main topic of this paper: investigating the consequences of a judgement of partial exchangeability. We start with *finite* partially exchangeable sequences. In order to treat this problem in its full generality, we will consider in all instances, for all $i \in G$, some $J_i \Subset \mathbb{N}$ or, in other words, choose a finite number of variables from every group. We use the shorthand notation $n_i = |J_i|$ to refer to the number of variables chosen from group i and collect these numbers in the vector $n_G = (n_1, \dots, n_g)$.

4.1. Defining finite partial exchangeability

As mentioned in the introduction, we model variables that are considered to be (g-fold) partially exchangeable. In terms of gambles, if a subject regards g sequences as partially exchangeable, we take this to mean that he considers the exchange of any gamble f on those sequences for its permuted form—the same gamble, after permuting the variables *within* their groups—to be equivalent to no transaction at all. In order to turn this into a formal definition, we introduce some extra notation.

For any $i \in G$, let $\mathcal{P}_{i(J_i)}$ be the set of all permutations π_i of the index set J_i. Then, for every $x \in \mathcal{X}_{i(J_i)}$ and every $\pi_i \in \mathcal{P}_{i(J_i)}$, we can consider the permuted sequence $\pi_i x$, as given by $(\pi_i x)_{ij} = x_{i\pi_i(j)}$ for every $j \in J_i$. As such, $\mathcal{P}_{i(J_i)}$ can be identified with a group of permutations of $\mathcal{X}_{i(J_i)}$. If for all $i \in G$, we have a permutation $\pi_i \in \mathcal{P}_{i(J_i)}$, we use the shorthand notation π to refer to the tuple $(\pi_1, \dots, \pi_g)$. Clearly, π is an element of $\mathcal{P}_{G(J)} := \bigtimes_{i \in G} \mathcal{P}_{i(J_i)}$. If we let $x = (x_{1(J_1)}, \dots, x_{g(J_g)})$ be a generic element of $\mathcal{X}_{G(J)}$, we define, by considering permutations within groups, $\pi x := (\pi_1 x_{1(J_1)}, \dots, \pi_g x_{g(J_g)})$.

Running example. Consider the following permutation: $\pi = (\pi_1, \pi_2)$, where $\pi_1 = (2, 1, 4, 3)$ and $\pi_2 = (3, 2, 1)$. Then for $x = (x_{1(N_1)}, x_{2(N_2)}) = ((0, 1, 1, 0), (1, 1, 0))$ we have $\pi x = ((1, 0, 0, 1), (0, 1, 1))$. We can also permute subsequences. For instance, for $J_1 = \{2, 4\}$ and $J_2 = \{2, 3\}$, we now write $x' = (x_{1(J_1)}, x_{2(J_2)}) = ((1, 0), (1, 0))$. With $\pi_1' = (4, 2)$ a permutation of J_1 and $\pi_2' = (3, 2)$ a permutation of J_2, we find that $\pi_1'(2) = 4$, $\pi_1'(4) = 2$, $\pi_2'(2) = 3$, $\pi_2'(3) = 2$ and therefore $\pi' x' = (\pi_1'(1, 0), \pi_2'(1, 0)) = ((0, 1), (0, 1))$. □

Next, we lift any permutation $\pi \in \mathcal{P}_{G(J)}$ to a linear transformation π^t of the set $\mathcal{G}(\mathcal{X}_{G(J)})$ of all gambles on $\mathcal{X}_{G(J)}$. For any gamble $f \in \mathcal{G}(\mathcal{X}_{G(J)})$, $\pi^t f := f \circ \pi$ is given, for all $x \in \mathcal{X}_{G(J)}$, by $\pi^t f(x) = f(\pi x)$, and this is the permuted gamble to which we referred above in our informal description of partial exchangeability.

Using our newly acquired notation, this description can now be formalised: if a subject judges the variables $X_{G(J)}$ to be partially exchangeable, this means that for any gamble $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ and any permutation $\pi \in \mathcal{P}_{G(J)}$, he considers the gamble $f - \pi^t f$ to be equivalent to the zero gamble. Using the terminology in Section 2.3: the gambles in

$$\mathcal{A}^{par}_{G(J)} := \left\{ f - \pi^t f : f \in \mathcal{G}(\mathcal{X}_{G(J)}), \pi \in \mathcal{P}_{G(J)} \right\}$$

belong to the subject's set of indifferent gambles. We denote by $\mathcal{I}^{par}_{G(J)}$ the smallest, most conservative, set of indifferent gambles that reflects such a judgement of partial exchangeability—includes $\mathcal{A}^{par}_{G(J)}$—and furthermore satisfies the rationality criteria of indifference. It is not hard to see that [use I2 and I3]

$$\mathcal{I}^{par}_{G(J)} = \operatorname{span}\big(\mathcal{A}^{par}_{G(J)}\big). \tag{8}$$

The only non-trivial problem is whether $\mathcal{I}^{par}_{G(J)}$ satisfies I4. The following result establishes that this is indeed the case.

Proposition 6. *$\mathcal{I}^{par}_{G(J)}$ is a linear subspace of $\mathcal{G}(\mathcal{X}_{G(J)})$ that satisfies I1–I4.*

Given the discussion in Section 2, it is now straightforward to define finite partial exchangeability, both in terms of sets of desirable gambles and in terms of lower previsions. We simply require compatibility with $\mathcal{I}^{par}_{G(J)}$.

Definition 5 *(Finite partial exchangeability).* A coherent set $\mathcal{D}_{G(J)}$ of desirable gambles on $\mathcal{X}_{G(J)}$ is called *partially exchangeable* if it is compatible with $\mathcal{I}^{par}_{G(J)}$. In that case, the variables $X_{G(J)}$ are said to be partially exchangeable with respect to $\mathcal{D}_{G(J)}$.

Definition 6 *(Finite partial exchangeability).* A coherent lower prevision $\underline{P}_{G(J)}$ on $\mathcal{G}(\mathcal{X}_{G(J)})$ is called *partially exchangeable* if it is compatible with $\mathcal{I}^{par}_{G(J)}$. In that case, the variables $X_{G(J)}$ are said to be partially exchangeable with respect to $\underline{P}_{G(J)}$.

Due to Proposition 2, both definitions are closely related to each other. Each of them is also related to a number of alternative definitions for (partial) exchangeability, some of which are perhaps better known to the reader. Indeed, Definition 5 is a generalisation of Definition 3(iii) in Ref. [15], which considers the particular case of exchangeability.[14] Similarly, the following result both generalises and strengthens Definition 3(i) in Ref. [15] by providing conditions that appear to be increasingly weaker, but are actually equivalent to the condition in Definition 5. It uses the following subset of $\mathcal{A}^{par}_{G(J)}$:

$$\mathcal{A}_{G(J)} := \big\{\mathbb{I}_{\{x\}} - \pi^t\mathbb{I}_{\{x\}} : x \in \mathcal{X}_{G(J)}, \pi \in \mathcal{P}_{G(J)}\big\}.$$

Proposition 7. *A coherent set $\mathcal{D}_{G(J)}$ of desirable gambles on $\mathcal{X}_{G(J)}$ is partially exchangeable if and only if any of the following two equivalent conditions holds:*

(i) *$f_1 + f_2 \in \mathcal{D}_{G(J)}$ for all $f_1 \in \mathcal{D}_{G(J)}$ and all $f_2 \in \mathcal{A}^{par}_{G(J)}$;*
(ii) *$f_1 + f_2 \in \mathcal{D}_{G(J)}$ for all $f_1 \in \mathcal{D}_{G(J)}$ and all $f_2 \in \mathcal{A}_{G(J)}$.*

For lower previsions, Definition 6 is a generalisation of the definition that was used in Ref. [15],[15] which—as already mentioned before—was concerned with the particular case of exchangeability. In order to relate the condition in Definition 6 to other definitions in the literature, we consider the following three equivalent conditions.

Proposition 8. *A coherent lower prevision $\underline{P}_{G(J)}$ on $\mathcal{G}(\mathcal{X}_{G(J)})$ is partially exchangeable if and only if any of the following equivalent conditions holds:*

(i) *$\underline{P}_{G(J)}(f) \geq 0$ for all $f \in \mathcal{A}^{par}_{G(J)}$;*
(ii) *$\underline{P}_{G(J)}(f) \geq 0$ for all $f \in \mathcal{A}_{G(J)}$;*
(iii) *All the linear previsions $P_{G(J)}$ in $\mathcal{M}(\underline{P}_{G(J)})$ are partially exchangeable.*

[14] Exchangeability corresponds to the special case where only one sequence of variables is considered: $g = 1$ and $G = \{1\}$.

[15] This follows by combining the condition in Definition 5 with Theorem 11(iii) in Ref. [15] or, alternatively, with Proposition 2 and Definition 3 in Ref. [15].

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

For the case of exchangeability, Ref. [21] used condition (i) as a definition. In that same setting, Walley imposes an apparently stronger version of condition (i), which replaces the inequality by an equality [11, Section 9.5.1]; it should be clear that this definition is equivalent as well. For two sequences that are considered to be partially exchangeable—$g = 2$ and $G = \{1, 2\}$—Cozman proposed condition (ii) [25, Section 3.5.3].

Conditions (i) and (ii) both imply that partial exchangeability of a lower prevision, similarly to coherence, is preserved under taking lower envelopes. Hence, by combining this with condition (iii), we find that a coherent lower prevision $\underline{P}_{G(J)}$ is partially exchangeable if and only if it is the lower envelope of some set of partially exchangeable linear previsions and, in that case, every other linear prevision that dominates $\underline{P}_{G(J)}$ will be partially exchangeable as well. This connection is important because partial exchangeability of a linear prevision is closely related to the usual, precise-probabilistic definition of partial exchangeability for probability measures.

As explained in Section 2.5, every linear prevision $P_{G(J)}$ on $\mathcal{G}(\mathcal{X}_{G(J)})$ is the expectation operator of a unique probability measure, which we denote by $P_{G(J)}$ as well. In this case, since $\mathcal{X}_{G(J)}$ is finite, the distinction between finite and countable additivity disappears and, furthermore, $P_{G(J)}$ has a unique corresponding probability mass function $p_{G(J)}$ on $\mathcal{X}_{G(J)}$ from which it can be derived in the usual way. The following result establishes that partial exchangeability of the linear prevision—in the sense of Definition 6—is equivalent to partial exchangeability of the corresponding probability measure—in the usual sense [1].

Proposition 9. *A linear prevision $P_{G(J)}$ on $\mathcal{G}(\mathcal{X}_{G(J)})$ is partially exchangeable if and only if the corresponding probability mass function $p_{G(J)}$ satisfies:*

$$p_{G(J)}(\pi x) = p_{G(J)}(x) \quad \text{for all } x \in \mathcal{X}_{G(J)} \text{ and all } \pi \in \mathcal{P}_{G(J)}. \tag{9}$$

4.2. From sequences to count vectors

In a precise-probabilistic setting, a judgement of partial exchangeability turns frequency counts into sufficient statistics. In order to show that such a result extends to our setting as well—as we will do in Section 4.3, we need tools that enable us to count, for each element of $\mathcal{X}_i$, $i \in G$, how many times it occurs within a given sequence of outcomes. We start by defining the following sets of *count vectors*:

$$\mathcal{N}^{n_i} := \left\{ m_i \in \mathbb{N}^{\mathcal{X}_i} : \sum_{z \in \mathcal{X}_i} (m_i)_z = n_i \right\} \quad \text{for all } i \in G,$$

where we denoted the z-component of m_i as $(m_i)_z$. Recalling the notational convention for n_G in the very beginning of this section, we extend this notation towards multiple sequences by defining $\mathcal{N}^{n_G} := \times_{i \in G} \mathcal{N}^{n_i}$. A generic element m of $\mathcal{N}^{n_G}$ is a G-tuple, consisting of count vectors $m_i \in \mathcal{N}^{n_i}$, one for every $i \in G$.

The reason why we call them count vectors is because they are the result of so-called *counting operators*. The counting operator $T_{i(J_i)}$ is a map from $\mathcal{X}_{i(J_i)}$ to $\mathcal{N}^{n_i}$. It maps any $x \in \mathcal{X}_{i(J_i)}$ to a count vector $T_{i(J_i)}(x) \in \mathcal{N}^{n_i}$ of which the z-component is given by

$$T_{i(J_i)}(x)_z := \left| \{ j \in J_i : x_{ij} = z \} \right| \quad \text{for all } z \in \mathcal{X}_i.$$

It should be clear that $T_{i(J_i)}(x)_z$ is the number of times the element z occurs in the sequence x, hence the term count vector. It is also useful to extend the domain of $T_{i(J_i)}$ to $\mathcal{X}_{G(J)}$. For all $x \in \mathcal{X}_{G(J)}$, this extended version is given by $T_{i(J_i)}(x) := T_{i(J_i)}(x_{i(J_i)})$. This makes it particularly easy to introduce the operator $T_{G(J)} : \mathcal{X}_{G(J)} \to \mathcal{N}^{n_G}$. For all $x \in \mathcal{X}_{G(J)}$, $T_{G(J)}(x)$ is a G-tuple that has the count vectors $T_{i(J_i)}(x)$ as its components, one for every $i \in G$.

Running example. Consider again our running example with four men and three women. For the men, we have that $n_1 = 4$, $J_1 = N_1 = \{1, 2, 3, 4\}$ and $\mathcal{X}_1 = \{0, 1\}$. Hence, the set of possible count vectors is

$$\mathcal{N}^{n_1} = \mathcal{N}^4 = \{ m_1 \in \mathbb{N}^{\{0,1\}} : (m_1)_0 + (m_1)_1 = 4 \}.$$

For the women, we have $n_2 = 3$, $J_2 = N_2 = \{1, 2, 3\}$ and $\mathcal{X}_2 = \{0, 1\}$ and therefore, the set of possible count vectors is

$$\mathcal{N}^{n_2} = \mathcal{N}^3 = \{ m_2 \in \mathbb{N}^{\{0,1\}} : (m_2)_0 + (m_2)_1 = 3 \}.$$

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

Let us denote the outcome of the experiments on both sexes as $x = (x_{1(N_1)}, x_{2(N_2)})$, and recall that in this running example, the outcome is $x_{1(N_1)} = (0, 1, 1, 0)$ for the men and $x_{2(N_2)} = (1, 1, 0)$ for the women. We then obtain a joint count vector $m := T_{G(N)}(x) = (T_{1(N_1)}(x), T_{2(N_2)}(x)) =: (m_1, m_2)$, where

$$m_1 = T_{1(N_1)}(x) = \big(T_{1(N_1)}(x)_0, T_{1(N_1)}(x)_1\big) = (2, 2)$$

and, similarly, $m_2 = T_{2(N_2)}(x) = (1, 2)$. Despite the admittedly rather cumbersome notation, the meaning of these formulas should be clear. To give an example: $T_{1(N_1)}(x)_0 = 2$ simply means that amongst the four tested men, two were "not cured". □

Count vectors are useful to identify what we call permutation invariant atoms: for all $x \in \mathcal{X}_{G(J)}$, the atom $[x] := \{\pi x : \pi \in \mathcal{P}_{G(J)}\}$ consists of all the permutations of x, obtained by permuting its elements x_{ij}—within their respective groups $i \in G$—in all possible ways. These atoms are the smallest permutation invariant subsets of $\mathcal{X}_{G(J)}$. Since y belongs to $[x]$ if and only $T_{G(J)}(y) = T_{G(J)}(x)$, the atom $[x]$ is completely determined by the count vector $m = T_{G(J)}(x)$, allowing us to identify $[x]$ with $[m] := T^{-1}_{G(J)}(m) = \{y \in \mathcal{X}_{G(J)} : T_{G(J)}(y) = m\}$. The number of elements in $[m]$ is given by

$$\big|[m]\big| = \prod_{i \in G} \big|[m_i]\big|, \quad \text{where } \big|[m_i]\big| := \binom{n_i}{m_i} := \frac{n_i!}{\prod_{z \in \mathcal{X}_i} (m_i)_z!} \quad \text{for all } i \in G. \tag{10}$$

Each $[m_i] := \{x_i \in \mathcal{X}_{i(J_i)} : T_{i(J_i)}(x_i) = m_i\}$ is the permutation invariant atom of $\mathcal{X}_{i(J_i)}$ with count vector m_i, associated with the permutation group $\mathcal{P}_{i(J_i)}$. Observe that $[m] = \times_{i \in G} [m_i]$, which justifies Eq. (10).

Running example. Consider again the particular case of our running example where $J_1 = \{2, 4\}$, $J_2 = \{2, 3\}$ and $x = (x_{1(J_1)}, x_{2(J_2)}) = ((1, 0), (1, 0)) \in \mathcal{X}_{G(J)}$. Besides x, the invariant atom $[x]$ contains three additional elements of $\mathcal{X}_{G(J)}$, obtained by permuting the individual values of x within their respective groups: $((1, 0), (0, 1))$, $((0, 1), (1, 0))$ and $((0, 1), (0, 1))$. □

4.3. *Finite representation theorems for partial exchangeability*

In the previous section, we established that there is a strong connection between permutations and count vectors. In the present section, we exploit this connection to derive finite representation theorems for partial exchangeability. We start by introducing a linear transformation $\mathrm{ex}_{G(J)}$ of the linear space $\mathcal{G}(\mathcal{X}_{G(J)})$:

$$\mathrm{ex}_{G(J)}(f) := \frac{1}{|\mathcal{P}_{G(J)}|} \sum_{\pi \in \mathcal{P}_{G(J)}} \pi^t f \quad \text{for all } f \in \mathcal{G}(\mathcal{X}_{G(J)}),$$

where $|\mathcal{P}_{G(J)}| = \prod_{i \in G} |\mathcal{P}_{i(J_i)}|$ and $|\mathcal{P}_{i(J_i)}| = n_i!$. This operator is the uniform average of the gambles $\pi^t f$, taken over all the permutations $\pi \in \mathcal{P}_{G(J)}$ and, in this way, it makes any gamble f insensitive to permutations: we invite the reader to check that for all $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ and all $\pi \in \mathcal{P}_{G(J)}$:

$$\mathrm{ex}_{G(J)}\big(\pi^t f\big) = \pi^t\big(\mathrm{ex}_{G(J)}(f)\big) = \mathrm{ex}_{G(J)}(f) = \mathrm{ex}_{G(J)}\big(\mathrm{ex}_{G(J)}(f)\big). \tag{11}$$

The third equality tells us that $\mathrm{ex}_{G(J)}$ is a projection operator, and the second equality guarantees that it maps gambles to gambles that are permutation invariant. Hence, the value of $\mathrm{ex}_{G(J)}(f)$ is constant on the permutation invariant atoms, that is, it assumes the same value in any element of such an atom. For any $x \in \mathcal{X}_{G(J)}$, the constant value that is assumed by $\mathrm{ex}_{G(J)}(f)$ on the elements of $[x]$ is

$$\begin{aligned}\frac{1}{|\mathcal{P}_{G(J)}|} \sum_{\pi \in \mathcal{P}_{G(J)}} \pi^t f(x) &= \frac{1}{|\mathcal{P}_{G(J)}|} \sum_{\pi \in \mathcal{P}_{G(J)}} f(\pi x) = \frac{1}{|\mathcal{P}_{G(J)}|} \sum_{y \in [x]} \sum_{\substack{\pi \in \mathcal{P}_{G(J)}:\\ \pi x = y}} f(y) \\ &= \frac{1}{|\mathcal{P}_{G(J)}|} \sum_{y \in [x]} \frac{|\mathcal{P}_{G(J)}|}{|[x]|} f(y) = \frac{1}{|[x]|} \sum_{y \in [x]} f(y).\end{aligned}$$

If we recall that $[x] = [m]$ for $m = T_{G(J)}(x)$, this means that we can identify $\mathrm{ex}_{G(J)}(f)$ with a gamble on the count vectors in $\mathcal{N}^{n_G}$. Inspired by this observation, we introduce an operator $\mathrm{Hy}_{G(J)}$ that maps $\mathcal{G}(\mathcal{X}_{G(J)})$ to $\mathcal{G}(\mathcal{N}^{n_G})$. For every gamble f on $\mathcal{X}_{G(J)}$, we define $\mathrm{Hy}_{G(J)}(f)$ by

$$\mathrm{Hy}_{G(J)}(f)(m) := \mathrm{Hy}_{G(J)}(f|m) := \frac{1}{|[m]|} \sum_{x \in [m]} f(x) \quad \text{for all } m \in \mathcal{N}^{n_G}$$

Consequently, we have that

$$\mathrm{ex}_{G(J)}(f) = \mathrm{Hy}_{G(J)}(f) \circ T_{G(J)}. \tag{12}$$

For each given $m \in \mathcal{N}^{n_G}$, $\mathrm{Hy}_{G(J)}(\cdot|m)$ is the expectation operator associated with the uniform distribution on $[m]$. This uniform distribution is the independent joint of the uniform distributions on the invariant atoms $[m_i]$, $i \in G$, and each of these independent uniform distributions is essentially a multivariate hypergeometric distribution, associated with random sampling, without replacement, from an urn that contains n_i balls, $(m_i)_x$ of which are of type x, with $x \in \mathcal{X}_i$. In summary, $\mathrm{Hy}_{G(J)}(\cdot|m)$ is the expectation operator for an independent joint of g multivariate hypergeometric distributions, and is associated with random sampling, without replacement, independently from g urns.

Running example. Using once more the data from our example, we illustrate the operator $\mathrm{Hy}_{G(J)}$. If we let $J_1 = \{1, 2, 3\}$ and $J_2 = \{1, 2\}$, then $x = ((0, 1, 1), (1, 1))$ and $m = T_{G(J)}(x) = ((1, 2), (0, 2))$. The atom $[m]$ contains $|[m]| = 3$ elements, one of which is x. The other two are $x' = ((1, 0, 1), (1, 1))$ and $x'' = ((1, 1, 0), (1, 1))$. If we let E be the event that, amongst the first two men and first woman tested, two people are cured and one is not, then the intersection of E and $[m]$ consists of two elements: x and x'. Hence, with $f = \mathbb{I}_E \in \mathcal{G}(\mathcal{X}_{G(J)})$, we find that

$$\mathrm{Hy}_{G(J)}(\mathbb{I}_E|m) = \frac{1}{3} \sum_{y \in [m]} \mathbb{I}_E(y) = \frac{1}{3}\big(\mathbb{I}_E(x) + \mathbb{I}_E(x') + \mathbb{I}_E(x'')\big) = \frac{1}{3}(1 + 1 + 0) = \frac{2}{3}.$$

This is the probability of extracting one white ball when drawing two balls, without replacement, from an urn with two white balls and one black ball multiplied by the probability of extracting one white ball from an urn with two white balls, if we consider the extraction of every ball equally probable. □

The linear operators $\mathrm{ex}_{G(J)}$ and $\mathrm{Hy}_{G(J)}$, which essentially transform gambles to permutation invariant ones, owe their importance to the following results, which clearly illustrate their connection to partial exchangeability.

Proposition 10. *$\mathcal{I}^{par}_{G(J)}$ is the kernel of both $\mathrm{ex}_{G(J)}$ and $\mathrm{Hy}_{G(J)}$:*

$$f \in \mathcal{I}^{par}_{G(J)} \Leftrightarrow \mathrm{ex}_{G(J)}(f) = 0 \Leftrightarrow \mathrm{Hy}_{G(J)}(f) = 0 \quad \textit{for all } f \in \mathcal{G}(\mathcal{X}_{G(J)}).$$

Proposition 11. *A coherent set $\mathcal{D}_{G(J)}$ of desirable gambles on $\mathcal{X}_{G(J)}$ is partially exchangeable if and only if*

$$f \in \mathcal{D}_{G(J)} \Leftrightarrow \mathrm{ex}_{G(J)}(f) \in \mathcal{D}_{G(J)} \quad \textit{for all } f \in \mathcal{G}(\mathcal{X}_{G(J)}). \tag{13}$$

This last result is particularly important. It means that, under an assessment of exchangeability, the desirability of a gamble f is fully determined by the desirability of its permutation invariant counterpart $\mathrm{ex}_{G(J)}(f)$. Given that, by Eq. (12), $\mathrm{ex}_{G(J)}(f)$ is completely characterised by the gamble $\mathrm{Hy}_{G(J)}(f)$ on $\mathcal{N}^{n_G}$, Proposition 11 suggests that an exchangeable set of desirable gambles can be fully represented by means of gambles on count vectors only, which significantly reduces the dimension of the model, and suggests the role of count vectors as sufficient statistics. The following representation theorem shows that this is indeed the case.

Theorem 12 *(Finite representation). A set $\mathcal{D}_{G(J)}$ of desirable gambles on $\mathcal{X}_{G(J)}$ is coherent and partially exchangeable if and only if there is some coherent set $\mathcal{R}_{n_G}$ of desirable gambles on $\mathcal{N}^{n_G}$ such that*

$$\mathcal{D}_{G(J)} = \mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$$

and, in that case, this $\mathcal{R}_{n_G}$ is uniquely determined by $\mathcal{R}_{n_G} = \mathrm{Hy}_{G(J)}(\mathcal{D}_{G(J)})$, and referred to as the count representation *of $\mathcal{D}_{G(J)}$.*

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

Using Proposition 2, it is relatively easy to derive a similar result in terms of coherent lower previsions.

Theorem 13 *(Finite representation). A lower prevision $\underline{P}_{G(J)}$ on $\mathcal{G}(\mathcal{X}_{G(J)})$ is coherent and partially exchangeable if and only if there is some coherent lower prevision $\underline{Q}_{n_G}$ on $\mathcal{G}(\mathcal{N}^{n_G})$ such that $\underline{P}_{G(J)} = \underline{Q}_{n_G} \circ \mathrm{Hy}_{G(J)}$ and, in that case, this $\underline{Q}_{n_G}$ is uniquely determined by*

$$\underline{Q}_{n_G}(r) = \underline{P}_{G(J)}(r \circ T_{G(J)}) \quad \text{for all } r \in \mathcal{G}\big(\mathcal{N}^{n_G}\big)$$

and referred to as the count representation *of $\underline{P}_{G(J)}$.*

These theorems tell us that for imprecise probability models also, making observations under partial exchangeability is essentially equivalent to independent random sampling without replacement from a number of urns with uncertain compositions.

If we apply Theorem 13 to a partially exchangeable linear prevision $P_{G(J)}$, then, clearly, the resulting count representation will be a linear prevision as well. Given the discussion in Section 2.5, this representation Q_{n_G} is the expectation operator with respect to a probability mass function on the count vectors in $\mathcal{N}^{n_G}$, so we obtain in this way the usual finite representation theorem for partial exchangeability as a special case.[16] By combining this special case with Proposition 8(iii), we see that, in general, $\underline{Q}_{n_G}$ is the lower envelope of the count representations Q_{n_G} of the linear previsions $P_{G(J)}$ that dominate $\underline{P}_{G(J)}$. Hence, we find that within our imprecise-probabilistic context, we no longer have a single representing probability mass function on $\mathcal{N}^{n_G}$, but rather a (convex) set of them.

From a practical point of view, the main advantage of Theorems 12 and 13 is that they result in representations that can be expressed in terms of a lower dimensional space. The dimension of $\mathcal{G}(\mathcal{N}^{n_G})$ is typically much smaller than that of $\mathcal{G}(\mathcal{X}_{G(J)})$. For example, with $G = \{1, 2\}$, $\mathcal{X}_1$ and $\mathcal{X}_2$ binary, and $n_1 = n_2 = n$, the dimension of $\mathcal{G}(\mathcal{X}_{G(J)})$ is 2^{2n}, whereas $\mathcal{G}(\mathcal{N}^{n_G})$ has dimension $(n + 1)^2$.

4.4. Polynomial gambles

We have just seen that judging a finite number of random variables to be partially exchangeable allows us to express beliefs about these variables by means of gambles on count vectors rather than gambles on finite sequences. In the sequel, we will show that similar results can be obtained for *countable* sequences as well. However, in order to do so, we can no longer use count vectors: it does not make sense to 'count' in infinite sequences. Instead, our infinite representation theorems will be expressed in terms of a more abstract framework, which we introduce in the present section.

For all $i \in G$, we consider the set Σ_i consisting of all probability mass functions on $\mathcal{X}_i$. We call it the $\mathcal{X}_i$-simplex, and define it as

$$\Sigma_i := \Big\{\theta_i \in \mathbb{R}^{\mathcal{X}_i} : (\forall x \in \mathcal{X}_i)\big((\theta_i)_x \geq 0\big) \text{ and } \sum_{x \in \mathcal{X}_i} (\theta_i)_x = 1\Big\}.$$

The cross product of these simplices is denoted by $\Sigma_G := \bigtimes_{i \in G} \Sigma_i$. Clearly, choosing $\theta \in \Sigma_G$ is equivalent to specifying, for every $i \in G$, a probability mass function θ_i on $\mathcal{X}_i$. The set of all gambles on Σ_G is denoted by $\mathcal{G}(\Sigma_G)$.

As a special subset of $\mathcal{G}(\Sigma_G)$, we consider the set $\mathcal{V}(\Sigma_G)$ consisting of all *polynomial gambles* h on Σ_G, which are those gambles that are the restriction to Σ_G of a (multivariate) polynomial p on $\bigtimes_{i \in G} \mathbb{R}^{\mathcal{X}_i}$, in the sense that $h(\theta) = p(\theta)$ for all $\theta \in \Sigma_G$. We call p a representation of h. Clearly, for a given polynomial gamble h, there can be many such representations.

For every polynomial p on $\bigtimes_{i \in G} \mathbb{R}^{\mathcal{X}_i}$, we denote by $\deg_i(p)$ the total degree in the variables $(\theta_i)_x$, with $x \in \mathcal{X}_i$. We use $\deg_G(p)$ to refer to the tuple that has $\deg_i(p)$ as its components. Using this notation, we define $\mathcal{V}^{n_G}(\Sigma_G)$ as the set consisting of those polynomial gambles h on Σ_G that have at least one representation p for which $\deg_G(p) \leq n_G$, by which we mean that, for all $i \in G$, $\deg_i(p) \leq n_i$. We say that $\mathcal{V}^{n_G}(\Sigma_G)$ is the set of polynomial gambles of degree up to n_G.

[16] For partially exchangeable binary random variables, a finite version of de Finetti's representation theorem can be found in Refs. [26, Section 3] and [27, p.213].

$\mathcal{V}^{n_G}(\Sigma_G)$ and $\mathcal{V}(\Sigma_G)$ are both linear subspaces of $\mathcal{G}(\Sigma_G)$ that include the constant gambles. This is important, because it allows us to apply the results discussed in Section 2.

An important subclass of polynomial gambles are the Bernstein gambles. Consider, for any $m_i \in \mathcal{N}^{n_i}$, the associated Bernstein basis polynomial B_{m_i} on $\mathbb{R}^{\mathcal{X}_i}$, given by

$$B_{m_i}(\theta_i) := \binom{n_i}{m_i} \prod_{x \in \mathcal{X}_i} (\theta_i)_x^{(m_i)_x} \quad \text{for all } \theta_i \in \mathbb{R}^{\mathcal{X}_i},$$

where $\binom{n_i}{m_i}$, as in Eq. (10), is the multinomial coefficient. We generalise this by defining, for all $m \in \mathcal{N}^{n_G}$, $B_m = \prod_{i \in G} B_{m_i}$, which we call a Bernstein basis polynomial as well. The restriction of B_m to Σ_G is called a Bernstein gamble and will also be denoted as B_m. The distinction between the polynomial and the polynomial gamble should be clear from the context. The importance of Bernstein gambles is due to the following result.

Proposition 14. *The set of Bernstein gambles $\{B_m : m \in \mathcal{N}^{n_G}\}$ constitutes a basis for the linear space $\mathcal{V}^{n_G}(\Sigma_G)$ of all polynomial gambles of degree up to n_G.*

4.5. Finite representation in terms of polynomial gambles

As a first example of the use of polynomial gambles, we translate one of the finite representation theorems of Section 4.3 into this framework. An important tool that will help us to achieve this goal is the linear map CoMn^{n_G} from $\mathcal{G}(\mathcal{N}^{n_G})$ to $\mathcal{V}^{n_G}(\Sigma_G)$ that is defined for all $r \in \mathcal{G}(\mathcal{N}^{n_G})$ by

$$\mathrm{CoMn}^{n_G}(r) := \sum_{m \in \mathcal{N}^{n_G}} r(m) B_m,$$

where, for all $\theta \in \Sigma_G$, $\mathrm{CoMn}^{n_G}(r|\theta) := \mathrm{CoMn}^{n_G}(r)(\theta)$ is the expectation of r with respect to an independent joint of multinomial distributions whose parameters are n_i and θ_i, for $i \in G$. Since by Proposition 14, every $h \in \mathcal{V}^{n_G}(\Sigma_G)$ has a unique corresponding count gamble $b_h^{n_G} \in \mathcal{G}(\mathcal{N}^{n_G})$ for which $\mathrm{CoMn}^{n_G}(b_h^{n_G}) = h$, we find that CoMn^{n_G} is a linear isomorphism between the linear spaces $\mathcal{G}(\mathcal{N}^{n_G})$ and $\mathcal{V}^{n_G}(\Sigma_G)$. Hence, $b_h^{n_G}$ is the unique gamble given by $b_h^{n_G} := (\mathrm{CoMn}^{n_G})^{-1}(h)$.

By applying the operator CoMn^{n_G} to every element of some set $\mathcal{R}_{n_G}$ of desirable gambles on $\mathcal{N}^{n_G}$, we obtain a set $\mathcal{H}_{n_G} = \mathrm{CoMn}^{n_G}(\mathcal{R}_{n_G})$ of polynomial gambles on Σ_G, which clearly is a subset of $\mathcal{V}^{n_G}(\Sigma_G)$. We show in Proposition 15 that $\mathcal{H}_{n_G}$ is *Bernstein coherent at degree n_G* if and only if $\mathcal{R}_{n_G}$ is coherent.

Definition 7 *(Bernstein coherence at degree n_G).* A set $\mathcal{H}_{n_G}$ of polynomial gambles in $\mathcal{V}^{n_G}(\Sigma_G)$ is called *Bernstein coherent at degree n_G* if

$\mathrm{B}_{n_G}1.$ $0 \notin \mathcal{H}_{n_G}$,
$\mathrm{B}_{n_G}2.$ $B_m \in \mathcal{H}_{n_G}$ for all $m \in \mathcal{N}^{n_G}$,
$\mathrm{B}_{n_G}3.$ $\lambda h \in \mathcal{H}_{n_G}$ for all $h \in \mathcal{H}_{n_G}$ and all $\lambda \in \mathbb{R}_{>0}$,
$\mathrm{B}_{n_G}4.$ $h_1 + h_2 \in \mathcal{H}_{n_G}$ for all $h_1, h_2 \in \mathcal{H}_{n_G}$.

Alternatively, as in Ref. [15],[17] $\mathrm{B}_{n_G}2$ can be replaced by

$\mathrm{B}_{n_G}2'.$ $b_h^{n_G} > 0 \Rightarrow h \in \mathcal{H}_{n_G}$,

which contains $\mathrm{B}_{n_G}2$ as a special case. Since $\mathrm{B}_{n_G}2'$ is implied by $\mathrm{B}_{n_G}2$, $\mathrm{B}_{n_G}3$ and $\mathrm{B}_{n_G}4$, both definitions of Bernstein coherence at degree n_G are equivalent. We prefer our present version because it is stated entirely in terms of polynomial gambles, without using count gambles.

[17] They consider the particular case of exchangeability rather than partial exchangeability.

Proposition 15. *Consider a set $\mathcal{R}_{n_G}$ of desirable gambles on $\mathcal{N}^{n_G}$ and a set $\mathcal{H}_{n_G}$ of polynomial gambles on Σ_G up to degree n_G such that $\mathcal{H}_{n_G} = \mathrm{CoMn}^{n_G}(\mathcal{R}_{n_G})$ or, equivalently, such that $\mathcal{R}_{n_G} = (\mathrm{CoMn}^{n_G})^{-1}(\mathcal{H}_{n_G})$. Then $\mathcal{R}_{n_G}$ is coherent if and only if $\mathcal{H}_{n_G}$ is Bernstein coherent at degree n_G.*

We also introduce the map $\mathrm{Mn}_{G(J)} := \mathrm{CoMn}^{n_G} \circ \mathrm{Hy}_{G(J)}$ from $\mathcal{G}(\mathcal{X}_{G(J)})$ to $\mathcal{V}^{n_G}(\Sigma_G)$. For all $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ and $\theta \in \Sigma_G$,

$$\mathrm{Mn}_{G(J)}(f|\theta) := \mathrm{Mn}_{G(J)}(f)(\theta) = \mathrm{CoMn}^{n_G}\big(\mathrm{Hy}_{G(J)}(f)\big)(\theta) = \sum_{x \in \mathcal{X}_{G(J)}} f(x) \prod_{i \in G} \prod_{j \in J_i} (\theta_i)_{x_{ij}}$$

is the expectation of f with respect to an independent joint of categorical distributions, $n_i = |J_i|$ of which are defined on $\mathcal{X}_i$, with θ_i as their vector of probabilities, for $i \in G$.

With these tools in hand, we can now present our finite representation theorem in terms of polynomial gambles. Due to Proposition 15, it is an almost immediate consequence of Theorem 12.

Theorem 16. *A set of desirable gambles $\mathcal{D}_{G(J)}$ on $\mathcal{X}_{G(J)}$ is coherent and partially exchangeable if and only if there is some Bernstein coherent set $\mathcal{H}_{n_G}$ of polynomial gambles on Σ_G up to degree n_G such that*

$$\mathcal{D}_{G(J)} = \mathrm{Mn}_{G(J)}^{-1}(\mathcal{H}_{n_G})$$

and, in that case, this $\mathcal{H}_{n_G}$ is uniquely determined by $\mathcal{H}_{n_G} = \mathrm{Mn}_{G(J)}(\mathcal{D}_{G(J)})$ and referred to as the polynomial representation *of $\mathcal{D}_{G(J)}$.*[18] *Furthermore, the count representation of $\mathcal{D}_{G(J)}$ is then given by $\mathcal{R}_{n_G} = (\mathrm{CoMn}^{n_G})^{-1}(\mathcal{H}_{n_G})$.*

5. Countable partially exchangeable sequences

From now on, we no longer restrict ourselves to finite numbers of variables within each group. Instead, we will consider the complete set of all variables X_{ij}, with $i \in G$ and $j \in \mathbb{N}$, compactly represented by the single variable $X_{G(\mathbb{N})}$. We will explain what it means to consider these variables as partially exchangeable and derive representation theorems that include the usual precise-probabilistic versions as special cases. As a model for $X_{G(\mathbb{N})}$, we consider either a set of desirable gambles $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ that is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ or, alternatively, a coherent lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$. For every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, the corresponding marginal models will be denoted by $\mathcal{D}_{G(J)}$ and $\underline{P}_{G(J)}$ and are taken to be derived from $\mathcal{D}_{G(\mathbb{N})}$ and $\underline{P}_{G(\mathbb{N})}$ by means of marginalisation, as defined by Eqs. (6) and (7) respectively. Also, whenever n_i is not explicitly defined, it is taken to be equal to $|J_i|$ and, as such, N_i and n_G are silently instantiated as $\{1, \ldots, |J_i|\}$ and $(|J_1|, \ldots, |J_g|)$ respectively.[19]

5.1. Defining countable partial exchangeability

Our definition for countable partial exchangeability is completely analogous to the finite version that was presented in Section 4.1. The only difference is that here, instead of using the set of indifferent gambles $\mathcal{I}^{par}_{G(J)}$, we use the following superset:

$$\mathcal{I}^{par}_{G(\mathbb{N})} := \big\{ f \in \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})}) \colon f \in \mathcal{I}^{par}_{G(J)} \text{ for some } J_i \Subset \mathbb{N},\ i \in G \big\}. \tag{14}$$

Recalling the discussion on cylindrical extension in Section 3.2, $\mathcal{I}^{par}_{G(\mathbb{N})}$ is simply the union of the sets $\mathcal{I}^{par}_{G(J)}$, taken over all possible selections of a finite number of variables X_{ij}. The next result establishes that $\mathcal{I}^{par}_{G(\mathbb{N})}$—similarly to $\mathcal{I}^{par}_{G(J)}$—is a valid set of indifferent gambles.

Proposition 17. *$\mathcal{I}^{par}_{G(\mathbb{N})}$ is a linear subspace of $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ that satisfies I1–I4.*

[18] For the particular case of regular exchangeability, Ref. [15] calls this the *frequency representation*. We prefer not to use this terminology because, in statistics, frequencies are counts. Our representation here is instead related to limiting relative frequencies. By calling it the *polynomial representation*, we avoid any confusion that might arise.

[19] Theorems 22 and 23 serve as nice examples of how we use these conventions.

Partial exchangeability is now easily defined. Similarly to what we did in the finite case, we require compatibility with $\mathcal{I}^{par}_{G(\mathbb{N})}$.

Definition 8 *(Countable partial exchangeability).* A set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ that is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is called *partially exchangeable* if it is compatible with $\mathcal{I}^{par}_{G(\mathbb{N})}$. In that case, the variables $X_{G(\mathbb{N})}$ are said to be partially exchangeable with respect to $\mathcal{D}_{G(J)}$.

Definition 9 *(Countable partial exchangeability).* A coherent lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is called *partially exchangeable* if it is compatible with $\mathcal{I}^{par}_{G(\mathbb{N})}$. In that case, the variables $X_{G(\mathbb{N})}$ are said to be partially exchangeable with respect to $\underline{P}_{G(\mathbb{N})}$.

By Proposition 2, these two definitions are tightly connected with each other. Furthermore, as is to be expected from the definition of $\mathcal{I}^{par}_{G(\mathbb{N})}$, both of these definitions for countable partial exchangeability are closely related to their finite counterparts.

Proposition 18. *A set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ that is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is partially exchangeable if and only if for every choice of $J_i \Subset \mathbb{N}$, $i \in G$, the marginal model $\mathcal{D}_{G(J)}$ is partially exchangeable.*

Proposition 19. *A coherent lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is partially exchangeable if and only if for every choice of $J_i \Subset \mathbb{N}$, $i \in G$, the marginal model $\underline{P}_{G(J)}$ is partially exchangeable.*

This connection with the finite versions of the definitions should not come as a surprise. In fact, partial exchangeability of countable sequences is often defined exactly in this way, by imposing finite partial exchangeability on every finite subset of variables.

Due to this connection, and given the discussion in Section 4.1, it is now easy to derive alternative characterisations for our notion of countable partial exchangeability. It suffices to combine Propositions 18 and 19 with Propositions 7 and 8 to obtain appropriate countable versions of the characterisations that are provided by Propositions 7 and 8. In particular, one finds that a coherent lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is partially exchangeable if and only if every dominating linear prevision $P_{G(\mathbb{N})}$ is.

For every such linear prevision $P_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$, we can furthermore establish an equivalence with the usual precise-probabilistic notion of partial exchangeability [1,2], similarly to what we did in Proposition 9. However, care should be taken in defining the probability measure that is used to state such a result. As explained in Section 2.5, $P_{G(\mathbb{N})}$ is an expectation operator with respect to a finitely additive probability measure, defined on all events $E \subseteq \mathcal{X}_{G(\mathbb{N})}$. However, since $P_{G(\mathbb{N})}$ is defined on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ rather than $\mathcal{G}(\mathcal{X}_{G(\mathbb{N})})$, many such measures may (and do) exist. Uniqueness can be obtained by considering the restriction of these measures to what could be called *events of finite structure*: those events that depend on the outcome a finite number of variables only. These are the only events whose indicators are of finite structure and, therefore, elements of $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$, thereby allowing us to define the probability of E as $P_{G(\mathbb{N})}(E) := P_{G(\mathbb{N})}(\mathbb{I}_E)$. As such, $P_{G(\mathbb{N})}$ has a unique corresponding probability measure of which it is the expectation functional. It is denoted by $P_{G(\mathbb{N})}$ as well, and is defined on the set of all events $E \subseteq \mathcal{X}_{G(\mathbb{N})}$ that are of finite structure.[20] For every choice of $J_i \Subset \mathbb{N}$, $i \in G$, the corresponding marginal measure is denoted by $P_{G(J)}$—just like its corresponding marginal linear prevision—and defined on all events $E \subseteq \mathcal{X}_{G(J)}$. Note that $P_{G(\mathbb{N})}$ is completely characterised by these marginals, and vice versa. With these correspondence in mind, we can now state the following countable version of Proposition 9.

Proposition 20. *A linear prevision $P_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is partially exchangeable if and only if for every choice of $J_i \Subset \mathbb{N}$, $i \in G$, the corresponding marginal probability measure $P_{G(J)}$ is partially exchangeable in the usual sense, meaning that its probability mass function $p_{G(J)}$ satisfies Eq. (9).*

[20] See Ref. [11, Section 3.2] for a general discussion of the connection between linear previsions, expectations and probability measures, and the conditions under which the correspondence is unique.

5.2. Representation theorems for countable partial exchangeability

We now turn to the final and perhaps most important contribution of this paper: representation theorems for countable sequences of random variables that are considered to be partially exchangeable. These variables will be modelled by means of sets of desirable gambles or lower previsions. For the representations themselves, we use the framework of polynomial gambles, as introduced in Section 4.4. In order to move from the finite case—as treated in Section 4.5—to the countable one, we start by extending the notion of Bernstein coherence (at degree n_G) to sets of polynomials that may be of arbitrary degree.

Definition 10 *(Bernstein coherence).* A set $\mathcal{H}$ of polynomial gambles on Σ_G is called *Bernstein coherent* if

B1. $0 \notin \mathcal{H}$,
B2. $B_m \in \mathcal{H}$ for all $n_G \in \mathbb{N}^G$ and $m \in \mathcal{N}^{n_G}$,
B3. $\lambda h \in \mathcal{H}$ for all $h \in \mathcal{H}$ and all $\lambda \in \mathbb{R}_{>0}$,
B4. $h_1 + h_2 \in \mathcal{H}$ for all $h_1, h_2 \in \mathcal{H}$.

Equivalently, B2 can be replaced by a more stringent axiom, similar to what is done in Ref. [15]; see the comment after Definition 7 as well. The following proposition establishes a connection between Bernstein coherence and Bernstein coherence at degree n_G.

Proposition 21. *A set $\mathcal{H}$ of polynomial gambles on Σ_G is Bernstein coherent if and only if, for all $n_G \in \mathbb{N}^G$, $\mathcal{H}_{n_G} := \mathcal{H} \cap \mathcal{V}^{n_G}(\Sigma_G)$ is Bernstein coherent at degree n_G.*

Using this connection, and by combining it with Theorem 16, we obtain the following representation theorem for countable partial exchangeability, in terms of sets of desirable gambles.

Theorem 22 *(Countable representation). A set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and partially exchangeable if and only if there is a Bernstein coherent set $\mathcal{H}$ of polynomial gambles on Σ_G such that for every choice of $J_i \Subset \mathbb{N}$, $i \in G$,*

$$\mathcal{D}_{G(J)} = \mathrm{Mn}_{G(J)}^{-1}(\mathcal{H}_{n_G}), \quad \textit{with } \mathcal{H}_{n_G} := \mathcal{H} \cap \mathcal{V}^{n_G}(\Sigma_G)$$

and, in that case, this $\mathcal{H}$ is uniquely determined by $\mathcal{H} = \bigcup_{n_G \in \mathbb{N}^G} \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)})$ and referred to as the polynomial representation *of $\mathcal{D}_{G(\mathbb{N})}$.*

Although this result may come across as less intuitive than its finite version, the main idea is identical to that of Theorem 12. Due to the assumption of partial exchangeability, every gamble $f \in \mathcal{D}_{G(J)}$ has a corresponding polynomial gamble $\mathrm{Mn}_{G(J)}(f) \in \mathcal{V}(\Sigma_G)$ that acts as a 'condensed representation'[21] for f. The connection between f and $\mathrm{Mn}_{G(J)}(f)$ is established by means of the multinomial distribution. Furthermore, coherence of $\mathcal{D}_{G(\mathbb{N})}$ relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ corresponds to Bernstein coherence of the polynomial representation $\mathcal{H}$.

By combining Theorem 22 with Proposition 2, we obtain a similar result in terms of lower previsions. Interestingly, in this case, Bernstein coherence of the representation is replaced by 'normal' coherence.

Theorem 23 *(Countable representation). A lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ is coherent and partially exchangeable if and only if there is a coherent lower prevision $\underline{R}$ on $\mathcal{V}(\Sigma_G)$ such that for every choice of $J_i \Subset \mathbb{N}$, $i \in G$,*

$$\underline{P}_{G(J)}(f) = \underline{R}\big(\mathrm{Mn}_{G(J)}(f)\big) \quad \textit{for all } f \in \mathcal{G}(\mathcal{X}_{G(J)})$$

and, in that case, this $\underline{R}$ is uniquely determined by

$$\underline{R}(h) = \underline{P}_{G(N)}\big(b_h^{n_G} \circ T_{G(N)}\big) \quad \textit{for all } n_G \in \mathbb{N}^G \textit{ and } h \in \mathcal{V}^{n_G}(\Sigma_G)$$

and referred to as the polynomial representation *of $\underline{P}_{G(\mathbb{N})}$.*

[21] We could even call $\mathrm{Mn}_{G(J)}(f)$ a 'sufficient statistic' for f as it allows us to perform all relevant inferences about the function f of the observations or data.

If we apply Theorem 23 to a partially exchangeable *linear* prevision $P_{G(\mathbb{N})}$, then the resulting polynomial representation will be a linear prevision R on $\mathcal{V}(\Sigma_G)$, the set of all polynomial gambles on Σ_G. Due to its coherence, R can also be extended to the set $\mathcal{C}(\Sigma_G)$ of all continuous gambles on Σ_G: by the Stone–Weierstraß theorem, every continuous gamble on Σ_G is the uniform limit of some sequence of polynomial gambles on Σ_G and, by coherence [P6], the prevision of this uniform limit is the limit of the previsions. By the Riesz–Markov–Kakutani representation theorem, this extension of R to $\mathcal{C}(\Sigma_G)$ has a unique corresponding σ-additive probability measure on the Borel sets of Σ_G such that, for all $f \in \mathcal{C}(\Sigma_G)$, $R(f)$ is the (Lebesgue) integral of f with respect to this measure. Hence, we see that the usual, precise-probabilistic, measure-theoretic representation theorems for partial exchangeability [2,3]—at least for finite possibility spaces $\mathcal{X}_i$, $i \in G$—correspond to a special case of Theorem 23.

However, this should not be taken to mean that they are equivalent. In order to obtain a unique representing probability measure, one needs to *assume* σ-additivity; we fail to see how, in the present context, such an assumption could be motivated by anything other than mathematical convenience. If this assumption is dropped, there will be many (finitely additive) representing probability measures, only one of which will be σ-additive on the Borel sets. Their expectation operators coincide on $\mathcal{C}(\Sigma_G)$, but may differ for gambles that are not continuous; see for example Ref. [28]. Our version of the representation theorem has the advantage of avoiding both (a) an assumption of σ-additivity and (b) the non-uniqueness that usually comes with dropping that assumption. As can be seen from Theorem 23, all that is really needed in order to fully represent a partially exchangeable (lower) prevision on the set $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of all gambles of finite structure is a (lower) prevision on the polynomial gambles on Σ_G or, equivalently, the continuous ones.

De Finetti avoided making an assumption of σ-additivity as well [8, Section 6.3], which is why in his original version of the representation theorem for (partially) exchangeable random variables, the representing object was a distribution function rather than a measure [1].[22] However, here as well, one needs to impose additional technical assumptions—regarding discontinuity points of the representing distribution function—to obtain uniqueness.[23] One particular option is to consider distribution functions as indeterminate at discontinuity points, as suggested by de Finetti [8, Section 6.5].[24] In any case, again, by using a linear prevision on $\mathcal{V}(\Sigma_G)$ instead of a distribution function, these issues need not be considered, and uniqueness is obtained automatically, without the need for any additional technical assumptions.

Let us now leave the issue of uniqueness, and take a closer look at the general case, where $\underline{P}_{G(\mathbb{N})}$ is coherent and partially exchangeable, but not necessarily linear. In that case, we find that the representation consists of a coherent lower prevision $\underline{R}$ on $\mathcal{V}(\Sigma_G)$. Taking into account the discussion after Proposition 19, we see that this $\underline{R}$ is the lower envelope of the polynomial representations R that correspond to the linear previsions $P_{G(\mathbb{N})}$ that dominate $\underline{P}_{G(\mathbb{N})}$. Therefore, one could also consider using this (convex) set of linear previsions R as a representation. The advantage of using their lower envelope $\underline{R}$ is that it provides a far more compact representation—a single operator instead of a possibly infinite number of them—which, nevertheless, still contains all relevant information. Similarly, given the connection described above, one could also consider using a set of distribution functions or a set of σ-additive probability measures. However, on top of the aforementioned issues related to uniqueness, these sets have the additional disadvantage that it is not always possible to recover $\underline{R}$ or, equivalently, $\underline{P}_{G(\mathbb{N})}$ from their lower (and/or upper) envelope.[25]

6. Conclusions

This paper has studied the notion of partial exchangeability within the frameworks of sets of desirable gambles and lower previsions. Its main contributions are four representation theorems: two for every framework, covering both the finite and countable case. The usual precise-probabilistic representation theorems for partially exchangeable random variables were obtained as special cases, as were the representation theorems in Refs. [15,21], which considered regular exchangeability. Apart from their generality, and the fact that they allow for imprecision, a distinctive feature of

[22] Similar representation theorems for partially exchangeable random variables can also be found in Refs. [26] and [27, p. 212].

[23] For example, for the particular case of exchangeable random variables that are binary, uniqueness of the representing distribution function can be obtained by assuming that it is right-continuous; see for example Ref. [29, p. 675].

[24] See Ref. [30] for some related results.

[25] This is because lower previsions are more general than lower and upper probabilities [19], of which lower and upper distribution functions (which, when combined, are sometimes referred to as a probability box or p-box [31]) are in turn a special case [11, Section 4.6.6].

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

our results is that they clearly indicate what the representation actually consists of. It all comes down to the simple fact that, due to the assumption of partial exchangeability, every gamble on the observable variables has a corresponding gamble on some lower-dimensional parameter space. Therefore, all relevant information can be represented in terms of (gambles on) this parameter space. For the countable case, it suffices to consider polynomial gambles on a cross-product of simplices.

Although our representation theorems in terms of lower previsions might come across as more intuitive—in part due to our effort of relating them to the usual precise-probabilistic case—we want to stress that the versions in terms of sets of desirable gambles are more fundamental. In the introduction, we already mentioned a number of general advantages of sets of desirable gambles, some of which they share with lower previsions. However, there is one feature that they do not share: a set of desirable gambles can always be conditioned in a unique way, even if the conditioning event has (lower) probability zero [16,19].[26] Consequently, our representation theorems for sets of desirable gambles lead to a representation—or 'prior'—that truly represents *all* relevant information about a sequence of partially exchangeable variables, including all conditional models.[27] For representations that are expressed in terms of (lower) previsions or probability measures, this is only true if the conditioning event has positive (lower) probability.[28]

From a more practical point of view, sets of desirable gambles and lower previsions have an important advantage in common: they both allow for imprecision.[29] Consequently, these frameworks allow for the use of imprecise-probabilistic priors or, loosely speaking, sets of precise-probabilistic priors. For the case of regular exchangeability, Walley's IDM(M)—Imprecise Dirichlet (Multinomial) Model—is a prime example [36,37][30]; see Ref. [42] for an introduction, including an overview of some of its applications. Rather than using a single Dirichlet prior, the IDM considers the set of all Dirichlet priors with some fixed strength.A key feature of such imprecise-probabilistic priors is that they are able to distinguish between ignorance and symmetry, a distinction which cannot be made using a single precise-probabilistic prior [20,36]. The corresponding inferences are, or rather can be, initially imprecise—even vacuous, in case of prior near-ignorance—but become more and more precise as the number of observations increases.

We believe that the use of imprecise-probabilistic priors—be it sets of probability measures, lower prevision or sets of desirable gambles—is particularly relevant in case of a judgement of *partial* exchangeability. The main reason is that, when modelling partially exchangeable data, the assessment and interpretation of a prior is usually more difficult than in situations where regular exchangeability is assumed. While in the latter setting one could come up with a prior density over the parameter space (a single simplex)—which is often easier to interpret—in the present context the subject has to specify a prior distribution over a cross-product of simplices. One approach is, for instance, to combine separate priors on the individual simplices by means of a copula [43]. However, the interpretation of these copula's is far from straightforward, making it difficult to choose one. Other approaches for coming up with a prior over a cross-product of simplices can be found in, amongst others, Ref. [44, chap. 9] and Refs. [45–50]. A common characteristic of these methods is that they are rather technical, and that the resulting priors are hard to interpret from a behavioural point of view. In order to reflect the arbitrariness that is implied by these difficulties in choosing a prior, one would be inclined to assess his beliefs more cautiously, leading to the use of an imprecise-probabilistic prior. Related to this, one could also consider using sets of copulas rather than a single one [51]; see Ref. [52] for some recent work on so-called imprecise copulas.

In order to truly exploit the aforementioned advantages, much work remains to be done. As far as theory is concerned, the results in the present paper should be used as a starting point to develop a general theory of conservative predictive inference under partial exchangeability. An important first step would be to extend some of the results in Ref. [15, Sections 5.2, 5.5 and 6] to the present context of partial exchangeability: in particular, the results on how to update an infinite representation, and how to extend (local) expert assessments—natural extension. In a next step,

[26] Within the context of imprecise probabilities, such a condition is rather weak: events with lower probability zero can have positive upper probability. This in contrast with precise probabilities, where the implications of assessing a (lower) probability to be zero are much stronger.

[27] For the particular case of exchangeable variables, see Refs. [15,32] for more information.

[28] Problems with probability zero can be avoided within those frameworks by considering conditional (lower) previsions and probabilities as primitive notions, rather than as derived concepts, relating them—not necessarily uniquely—to unconditional ones by means of coherence; see for instance Refs. [8,11,33–35]. However, even within these extended theories, the problem still persists that, in the presence of (lower) probability zero, an unconditional prior does not lead to unique conditional posterior models.

[29] Furthermore, they generalise almost every other framework that is capable of doing so [19].

[30] Other examples can be found in, amongst others, Refs. [32,38–41].

these results could then be applied to predictive inference. Reasonable inference principles could be imposed, and the implications of combining these (global) principles with (local) expert assessments could be studied; for the particular case of regular exchangeability, see for example Refs. [32,41]. From a more applied side, the most important question is whether, in the context of partial exchangeability, using imprecise-probabilistic priors can lead to useful inferences that are, despite the conservatism of the prior, informative enough to be used in practice. In order to answer this question, it is necessary to develop such priors, and to apply them to real-life problems. An obvious first suggestion is to generalise the IDM, extending it to the framework of partial exchangeability.

Acknowledgements

Jasper De Bock is a PhD Fellow of the Research Foundation Flanders (FWO) and wishes to acknowledge its financial support. Márcio Diniz was supported by FAPESP (São Paulo Research Foundation), under the project 2012/14764-0 and wishes to thank the SYSTeMS Research Group at Ghent University for its hospitality and support. Gert de Cooman's research was partially funded through project number 3G012512 of the Research Foundation Flanders (FWO).

Appendix A. Proofs

Proof of Proposition 1. First, consider any set $\mathcal{D} \subseteq \mathcal{K}$ that is coherent relative to $\mathcal{K}$ and let $\underline{P}_{\mathcal{D}}$ be the corresponding lower prevision on $\mathcal{K}$, as given by Eq. (1). We show that $\underline{P}_{\mathcal{D}}$ is coherent, meaning that it satisfies P1–P3. For P1, it suffices to realise that due to D2, $f - \mu \in \mathcal{D}$ for all $f \in \mathcal{K}$ and $\mu < \inf f$. P2 follows easily from D3. For P3, consider any $f_1, f_2 \in \mathcal{K}$. Then

$$\begin{aligned}\underline{P}_{\mathcal{D}}(f_1) + \underline{P}_{\mathcal{D}}(f_2) &= \sup\{\mu_1 + \mu_2 \colon \mu_1, \mu_2 \in \mathbb{R}, f_1 - \mu_1 \in \mathcal{D}, f_2 - \mu_2 \in \mathcal{D}\} \\ &\leq \sup\big\{\mu_1 + \mu_2 \colon \mu_1, \mu_2 \in \mathbb{R}, f_1 + f_2 - (\mu_1 + \mu_2) \in \mathcal{D}\big\} \\ &= \sup\{\alpha \in \mathbb{R} \colon f_1 + f_2 - \alpha \in \mathcal{D}\} = \underline{P}_{\mathcal{D}}(f_1 + f_2),\end{aligned}$$

where the inequality is due to D4.

Next, consider any coherent lower prevision $\underline{P}$ on $\mathcal{K}$ and define

$$\mathcal{D}' := \big\{f \in \mathcal{K} \colon f \in \mathcal{K}_{>0} \text{ or } \underline{P}(f) > 0\big\}. \tag{A.1}$$

We then find that, for all $f \in \mathcal{K}$,

$$\begin{aligned}\underline{P}_{\mathcal{D}'}(f) = \sup\big\{\mu \in \mathbb{R} \colon f - \mu \in \mathcal{D}'\big\} &= \sup\big\{\mu \in \mathbb{R} \colon f - \mu > 0 \text{ or } \underline{P}(f - \mu) > 0\big\} \\ &= \sup\big\{\mu \in \mathbb{R} \colon f > \mu \text{ or } \underline{P}(f) > \mu\big\}\end{aligned}$$

where the third equality is due to P5. Since $f > \mu$ implies $\inf f \geq \mu$, which, due to P1, in turn implies $\underline{P}(f) \geq \mu$, we find that

$$\underline{P}(f) = \sup\big\{\mu \in \mathbb{R} \colon \underline{P}(f) > \mu\big\} \leq \underline{P}_{\mathcal{D}'}(f) \leq \sup\big\{\mu \in \mathbb{R} \colon \underline{P}(f) \geq \mu\big\} = \underline{P}(f).$$

To conclude the proof, we are left to show that $\mathcal{D}'$ is coherent relative to $\mathcal{K}$, meaning that it satisfies D1–D4. D1 follows from the fact that, due to P5, $\underline{P}(0) = 0$. D2 is trivial and D3 follows from P2. For D4, it suffices to combine P3 and P1. □

Proof of Proposition 2. First, consider any set $\mathcal{D} \subseteq \mathcal{K}$ that is coherent relative to $\mathcal{K}$ and compatible with $\mathcal{I}$ and let $\underline{P}_{\mathcal{D}}$ be the corresponding lower prevision on $\mathcal{K}$, as given by Eq. (1). We know from Proposition 1 that $\underline{P}_{\mathcal{D}}$ is coherent. We now show that it is compatible with $\mathcal{I}$ as well. Consider therefore any $f \in \mathcal{I}$. For all $\epsilon < 0$, we know from D2 that $-\epsilon \in \mathcal{D}$ and therefore, due to ID1, that $f - \epsilon \in \mathcal{D}$. Hence, $\underline{P}_{\mathcal{D}}(f) \geq 0$. Assume *ex absurdo* that $\underline{P}_{\mathcal{D}}(f) > 0$, implying the existence of some $\mu > 0$ for which $f - \mu \in \mathcal{D}$. Since $\mu \in \mathcal{D}$ because of D2, we can use D4 to find that $f \in \mathcal{D}$, contradicting ID2.

Next, consider any coherent lower prevision $\underline{P}$ on $\mathcal{K}$ that is compatible with $\mathcal{I}$. Let $\mathcal{D}'$ be the set of desirable gambles that is given by Eq. (A.1) and define $\mathcal{D} := \mathcal{D}' + \mathcal{I}$. Recall from the proof of Proposition 1 that $\mathcal{D}'$ is coherent and that $\underline{P} = \underline{P}_{\mathcal{D}'}$. Since $\mathcal{I}$ satisfies I3, we have that $\mathcal{I} + \mathcal{I} = \mathcal{I}$ and therefore $\mathcal{D} + \mathcal{I} = \mathcal{D}$, implying that $\mathcal{D}$ and $\mathcal{I}$ satisfy ID1. Using ID1 and I1–I3, it is easy to infer from the coherence of $\mathcal{D}'$ that $\mathcal{D}$ is coherent as well. D2 and D3 are trivial [use I2 and I3]. For D1, assume *ex absurdo* that $0 \in \mathcal{D}$, implying the existence of some $f \in \mathcal{D}'$ such that $-f \in \mathcal{I}$ and therefore, due to I2, $f \in \mathcal{I}$. Since $f \in \mathcal{D}'$, we know that either $f > 0$ or $\underline{P}(f) > 0$, which, since $f \in \mathcal{I}$, contradicts either I4 or IP1, respectively. Hence, we know that $\mathcal{D}$ is both coherent relative to $\mathcal{K}$ and compatible with $\mathcal{I}$. To conclude the proof, we show that $\underline{P} = \underline{P}_{\mathcal{D}}$. So choose any $f \in \mathcal{K}$. Since, due to I1, $0 \in \mathcal{I}$ and therefore $\mathcal{D}' \subseteq \mathcal{D}$, we have that $\underline{P}_{\mathcal{D}}(f) \geq \underline{P}_{\mathcal{D}'}(f) = \underline{P}(f)$. Consider now any $\mu \in \mathbb{R}$ for which $f - \mu \in \mathcal{D}$, implying the existence of $f_1 \in \mathcal{D}'$ and $f_2 \in \mathcal{I}$ such that $f - \mu = f_1 + f_2$ and therefore $f = \mu + f_1 + f_2$. Using P3 and P5, we find that $\underline{P}(f) \geq \mu + \underline{P}(f_1) + \underline{P}(f_2)$. Since $f_1 \in \mathcal{D}'$, we have that $\underline{P}(f_1) = \underline{P}_{\mathcal{D}'}(f_1) \geq 0$ and, since $f_2 \in \mathcal{I}$, $\underline{P}(f_2) = 0$ because of IP1. Hence, $\underline{P}(f) \geq \mu$. Since this holds for all μ such that $f - \mu \in \mathcal{D}$, we find that $\underline{P}_{\mathcal{D}}(f) \leq \underline{P}(f)$. □

Proof of Proposition 4. The 'only if' part is a trivial consequence of Eq. (6). For the 'if' part, consider any set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ and assume that, for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, the marginal set of desirable gambles $\mathcal{D}_{G(J)}$, as given by Eq. (6), is coherent. We need to prove that $\mathcal{D}_{G(\mathbb{N})}$ is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ or, equivalently, that $\mathcal{D}_{G(\mathbb{N})}$ satisfies D1–D4. Properties D1–D3 follow directly from Eqs. (5) and (6) and the coherence of the marginal models $\mathcal{D}_{G(J)}$. For D4, consider any $f_1, f_2 \in \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$. Then due to Eq. (5), f_1 and f_2 each depend upon the value of a finite number of variables X_{ij}, with $i \in G$ and $j \in \mathbb{N}$. Hence, it is clearly possible to find $J_i \Subset \mathbb{N}$, with $i \in G$, such that f_1 and f_2 are both (cylindrical extensions of) elements of $\mathcal{G}(\mathcal{X}_{G(J)})$ and therefore also, due to Eq. (6), elements of $\mathcal{D}_{G(J)}$. Due to the coherence of $\mathcal{D}_{G(J)}$, $f_1 + f_2 \in \mathcal{D}_{G(J)}$ and hence, by applying Eq. (6) again, we find that $f_1 + f_2 \in \mathcal{D}_{G(\mathbb{N})}$. □

Proof of Proposition 5. The 'only if' part is a trivial consequence of Eq. (7). For the 'if' part, consider any lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and assume that, for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, the marginal lower prevision $\underline{P}_{G(J)}$, as given by Eq. (7), is coherent. We need to prove that $\underline{P}_{G(\mathbb{N})}$ is coherent, or equivalently, that it satisfies P1–P3. Properties P1 and P2 follow directly from Eqs. (5) and (7) and the coherence of the marginal models $\underline{P}_{G(J)}$. For P3, consider any $f_1, f_2 \in \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$. Then due to Eq. (5), and as explained in the proof of Proposition 4, there are $J_i \Subset \mathbb{N}$, $i \in G$, such that f_1 and f_2 are both elements of $\mathcal{G}(\mathcal{X}_{G(J)})$. Hence, we infer from the coherence of $\underline{P}_{G(J)}$ that $\underline{P}_{G(J)}(f_1 + f_2) \geq \underline{P}_{G(J)}(f_1) + \underline{P}_{G(J)}(f_2)$ and therefore also, due to Eq. (7), that $\underline{P}_{G(\mathbb{N})}(f_1 + f_2) \geq \underline{P}_{G(\mathbb{N})}(f_1) + \underline{P}_{G(\mathbb{N})}(f_2)$. □

Proof of Proposition 6. It suffices to prove I1–I4. Note that I1–I3 are direct consequences of Eq. (8). For I4, consider any gamble $f \in \mathcal{I}^{par}_{G(J)}$, and assume *ex absurdo* that $f > 0$ or $f < 0$. If $f > 0$, then $\pi^t f > 0$ for all permutations $\pi \in \mathcal{P}_{G(J)}$. This tells us that $\sum_{\pi \in \mathcal{P}_{G(J)}} \pi^t f > 0$, whence $\mathrm{ex}_{G(J)}(f) > 0$, contradicting Proposition 10. If $f < 0$, then similarly, $\mathrm{ex}_{G(J)}(f) < 0$, again contradicting Proposition 10. □

Proof of Proposition 7. Consider any coherent set $\mathcal{D}_{G(J)}$ of desirable gambles on $\mathcal{X}_{G(J)}$. Then taking into account Definitions 5 and 2, and because $\mathcal{A}_{G(J)} \subseteq \mathcal{A}^{par}_{G(J)} \subseteq \mathcal{I}^{par}_{G(J)}$, we immediately find that the partial exchangeability of $\mathcal{D}_{G(J)}$ implies condition (i), which in turn implies condition (ii). It therefore suffices to prove that condition (ii) implies the partial exchangeability of $\mathcal{D}_{G(J)}$. Assume that condition (ii) holds and consider any $f_1 \in \mathcal{D}_{G(J)}$ and any $f_2 \in \mathcal{I}^{par}_{G(J)}$. We need to prove that $f_1 + f_2 \in \mathcal{D}_{G(J)}$. Due to Lemma 24, f_2 is a finite sum of gambles λf, with $\lambda \in \mathbb{R}_{>0}$ and $f \in \mathcal{A}_{G(J)}$. Hence, the proof will follow by induction if we can show that $f_1 + \lambda f \in \mathcal{D}_{G(J)}$ for any such λ and f. By D3, $1/\lambda f_1 \in \mathcal{D}_{G(J)}$ and therefore, by condition (ii), $1/\lambda f_1 + f \in \mathcal{D}_{G(J)}$. Using D3 again, this leads to the desired result: $f_1 + \lambda f = \lambda(1/\lambda f_1 + f) \in \mathcal{D}_{G(J)}$. □

Lemma 24. *Every $f \in \mathcal{I}^{par}_{G(J)}$ is a finite positive linear combination of elements in $\mathcal{A}_{G(J)}$: $f = \sum_{k=1}^{m} \lambda_k f_k$, with $m \in \mathbb{N}$ and, for all $k \in \{1, \ldots, m\}$, $\lambda_k \in \mathbb{R}_{>0}$ and $f_k \in \mathcal{A}_{G(J)}$.*

Proof. Consider any $f \in \mathcal{I}^{par}_{G(J)}$. Then due to Eq. (8), f is a finite sum of gambles $\lambda(f' - \pi^t f') = f'' - \pi^t f''$, with $\lambda \in \mathbb{R}$, $f' \in \mathcal{G}(\mathcal{X}_{G(J)})$, $f'' := \lambda f' \in \mathcal{G}(\mathcal{X}_{G(J)})$ and $\pi \in \mathcal{P}_{G(J)}$. For any such gamble $f'' - \pi^t f''$, we have that

$$f'' - \pi^t f'' = \sum_{x \in \mathcal{X}_{G(J)}} f''(x)\mathbb{I}_{\{x\}} - \pi^t \sum_{x \in \mathcal{X}_{G(J)}} f''(x)\mathbb{I}_{\{x\}} = \sum_{x \in \mathcal{X}_{G(J)}} f''(x)\left[\mathbb{I}_{\{x\}} - \pi^t \mathbb{I}_{\{x\}}\right].$$

Furthermore, for any $x \in \mathcal{X}_{G(J)}$, we have that

$$-\left[\mathbb{I}_{\{x\}} - \pi^t \mathbb{I}_{\{x\}}\right] = \pi^t \mathbb{I}_{\{x\}} - \mathbb{I}_{\{x\}} = \mathbb{I}_{\{x'\}} - \left(\pi'\right)^t \mathbb{I}_{\{x'\}},$$

where $\pi' := \pi^{-1} \in \mathcal{P}_{G(J)}$ and $x' := \pi^{-1}x = \pi' x \in \mathcal{X}_{G(J)}$. Hence, f can be written as a non-negative linear combination of gambles in $\mathcal{A}_{G(J)}$.

This implies the desired result because we can write any term whose coefficient is zero as a new one that has a positive coefficient and $0 = \mathbb{I}_{\{x\}} - (\pi^*)^t \mathbb{I}_{\{x\}} \in \mathcal{A}_{G(J)}$ as its gamble, with π^* the identity permutation. □

Proof of Proposition 8. Partial exchangeability of $\underline{P}_{G(J)}$ means that $\underline{P}_{G(J)}$ is compatible with $\mathcal{I}^{par}_{G(J)}$. As mentioned in Definition 4, this is equivalent to IP2: $\underline{P}_{G(J)}(f) \geq 0$ for all $f \in \mathcal{I}^{par}_{G(J)}$. We prove that IP2 $\Leftrightarrow$ (iii) and IP2 $\Rightarrow$ (i) $\Rightarrow$ (ii) $\Rightarrow$ IP2.

We start with IP2 $\Leftrightarrow$ (iii). First, assume that $\underline{P}_{G(J)}$ satisfies IP2. Then for all $P_{G(J)} \in \mathcal{M}(\underline{P}_{G(J)})$, we can use Eq. (3) to infer that $P_{G(J)}(f) \geq \underline{P}_{G(J)}(f) \geq 0$ for all $f \in \mathcal{I}^{par}_{G(J)}$. Hence, $P_{G(J)}$ satisfies IP2 and therefore, is partially exchangeable. Conversely, if every $P_{G(J)} \in \mathcal{M}(\underline{P}_{G(J)})$ is partially exchangeable and therefore, satisfies IP2, we infer from Eq. (2) that $\underline{P}_{G(J)}$ does as well.

Since clearly, IP2 $\Rightarrow$ (i) $\Rightarrow$ (ii), we are left to prove that (ii) $\Rightarrow$ IP2. So assume that (ii) holds and consider any $f \in \mathcal{I}^{par}_{G(J)}$. We need to prove that $\underline{P}_{G(J)}(f) \geq 0$. Due to Lemma 24, we know that $f = \sum_{k=1}^{m} \lambda_k f_k$, with $m \in \mathbb{N}$ and, for all $k \in \{1, \ldots, m\}$, $\lambda_k \in \mathbb{R}_{>0}$ and $f_k \in \mathcal{A}_{G(J)}$. Hence, the result follows from P3, P2 and condition (ii). □

Proof of Proposition 9. We start by reformulating Eq. (9). For all $x \in \mathcal{X}_{G(J)}$ and all $\pi \in \mathcal{P}_{G(J)}$, we know from the discussion in Section 2.5 that $p_{G(J)}(x) := P_{G(J)}(\mathbb{I}_{\{x\}})$ and $p_{G(J)}(\pi x) := P_{G(J)}(\mathbb{I}_{\{\pi x\}})$. Furthermore, with $y = \pi x$, we have that $\pi^t \mathbb{I}_{\{y\}} = \mathbb{I}_{\{\pi^{-1} y\}} = \mathbb{I}_{\{x\}}$. Hence, given the linearity of $P_{G(J)}$, we find that Eq. (9) is equivalent to

$$P_{G(J)}\left(\mathbb{I}_{\{y\}} - \pi^t \mathbb{I}_{\{y\}}\right) = 0 \quad \text{for all } y \in \mathcal{X}_{G(J)} \text{ and all } \pi \in \mathcal{P}_{G(J)}. \tag{A.2}$$

This makes the 'only if'-part of this proof a trivial consequence of Definition 4. The 'if'-part follows directly from Eq. (A.2) and Proposition 8(ii). □

Proof of Proposition 10. The second equivalence follows directly from Eq. (12). Also, for any $f \in \mathcal{I}^{par}_{G(J)}$, $\mathrm{ex}_{G(J)}(f) = 0$ because of the linearity of $\mathrm{ex}_{G(J)}$ and Eqs. (8) and (11). So consider any $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ for which $\mathrm{ex}_{G(J)}(f) = 0$, then

$$f = f - \mathrm{ex}_{G(J)}(f) = f - \frac{1}{|\mathcal{P}_{G(J)}|} \sum_{\pi \in \mathcal{P}_{G(J)}} \pi^t f = \frac{1}{|\mathcal{P}_{G(J)}|} \sum_{\pi \in \mathcal{P}_{G(J)}} \left(f - \pi^t f\right),$$

and therefore, by Eq. (8), $f \in \mathcal{I}^{par}_{G(J)}$. □

Proof of Proposition 11. We begin by assuming that $\mathcal{D}_{G(J)}$ is partially exchangeable, implying that ID1 holds. Consider any $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ and let $f' := f - \mathrm{ex}_{G(J)}(f)$. Then by Eq. (11), $\mathrm{ex}_{G(J)}(f') = \mathrm{ex}_{G(J)}(-f') = 0$ and therefore, by Proposition 10, $f' \in \mathcal{I}^{par}_{G(J)}$ and $-f' \in \mathcal{I}^{par}_{G(J)}$. If $f \in \mathcal{D}_{G(J)}$, then $\mathrm{ex}_{G(J)}(f) = f - f' \in \mathcal{D}_{G(J)}$ because of ID1. Conversely, if $\mathrm{ex}_{G(J)}(f) \in \mathcal{D}_{G(J)}$, then $f = \mathrm{ex}_{G(J)}(f) + f' \in \mathcal{D}_{G(J)}$, again because of ID1.

Conversely, suppose that Eq. (13) holds and let us prove that $\mathcal{D}_{G(J)}$ satisfies ID1. So consider any $f \in \mathcal{D}_{G(J)}$ and any $f' \in \mathcal{I}^{par}_{G(J)}$. Then $\mathrm{ex}_{G(J)}(f) \in \mathcal{D}_{G(J)}$ by Eq. (13), and $\mathrm{ex}_{G(J)}(f') = 0$ by Proposition 10. By combining the two, we find that $\mathrm{ex}_{G(J)}(f + f') = \mathrm{ex}_{G(J)}(f) + \mathrm{ex}_{G(J)}(f') = \mathrm{ex}_{G(J)}(f) \in \mathcal{D}_{G(J)}$ and therefore, by Eq. (13), that $f + f' \in \mathcal{D}_{G(J)}$. □

Proof of Theorem 12. We start with sufficiency. Let $\mathcal{R}_{n_G}$ be a coherent set of desirable gambles on $\mathcal{N}^{n_G}$ such that $\mathcal{D}_{G(J)} = \mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$, or in other words,

$$f \in \mathcal{D}_{G(J)} \Leftrightarrow \mathrm{Hy}_{G(J)}(f) \in \mathcal{R}_{n_G} \quad \text{for all } f \in \mathcal{G}(\mathcal{X}_{G(J)}). \tag{A.3}$$

We need to show that $\mathcal{D}_{G(J)}$ is both coherent and exchangeable. For D1, infer from Eq. (A.3) that $0 \in \mathcal{D}_{G(J)} \Leftrightarrow 0 = \mathrm{Hy}_{G(J)}(0) \in \mathcal{R}_{n_G}$, and use the coherence of $\mathcal{R}_{n_G}$. For D2, notice that $f > 0$ implies that $\mathrm{Hy}_{G(J)}(f) > 0$, which in turn implies, since $\mathcal{R}_{n_G}$ is coherent, that $\mathrm{Hy}_{G(J)}(f) \in \mathcal{R}_{n_G}$ and therefore by Eq. (A.3) that $f \in \mathcal{D}_{G(J)}$. For D4, consider any $f_1, f_2 \in \mathcal{D}_{G(J)}$, meaning that $\mathrm{Hy}_{G(J)}(f_1)$ and $\mathrm{Hy}_{G(J)}(f_2)$ both belong to $\mathcal{R}_{n_G}$. Using the coherence of $\mathcal{R}_{n_G}$ and the linearity of $\mathrm{Hy}_{G(J)}$, we find that $\mathrm{Hy}_{G(J)}(f_1 + f_2) = \mathrm{Hy}_{G(J)}(f_1) + \mathrm{Hy}_{G(J)}(f_2) \in \mathcal{R}_{n_G}$ and therefore, that $f_1 + f_2 \in \mathcal{D}_{G(J)}$, again using Eq. (A.3). The proof for D3 is similar. To show that $\mathcal{D}_{G(J)}$ is partially exchangeable, consider any $f \in \mathcal{D}_{G(J)}$ and $f' \in \mathcal{I}^{par}_{G(J)}$. We prove that $f + f' \in \mathcal{D}_{G(J)}$. Using Proposition 10 and the linearity of $\mathrm{Hy}_{G(J)}$, we find that $\mathrm{Hy}_{G(J)}(f + f') = \mathrm{Hy}_{G(J)}(f) \in \mathcal{R}_{n_G}$ and, therefore indeed, that $f + f' \in \mathcal{D}_{G(J)}$.

We now turn to necessity. Assume that $\mathcal{D}_{G(J)}$ is coherent and partially exchangeable. We will show that $\mathcal{R}_{n_G} := \mathrm{Hy}_{G(J)}(\mathcal{D}_{G(J)})$ is a coherent set of desirable gambles on $\mathcal{N}^{n_G}$ for which $\mathcal{D}_{G(J)} = \mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$. Firstly, we prove coherence of $\mathcal{R}_{n_G}$. For D1, assume, *ex absurdo*, that $0 \in \mathcal{R}_{n_G}$ implying that there is some $f \in \mathcal{D}_{G(J)}$ such that $\mathrm{Hy}_{G(J)}(f) = 0$. Proposition 10 implies that $f \in \mathcal{I}^{par}_{G(J)}$, which contradicts ID2. For D2, consider any $r \in \mathcal{G}(\mathcal{N}^{n_G})_{>0}$. Then one can easily construct a gamble $f > 0$ for which $\mathrm{Hy}_{G(J)}(f) = r$. Due to the coherence of $\mathcal{D}_{G(J)}$, $f \in \mathcal{D}_{G(J)}$ and therefore indeed $r \in \mathcal{R}_{n_G}$. For D4, consider $r_1, r_2 \in \mathcal{R}_{n_G}$, implying that there are $f_1, f_2 \in \mathcal{D}_{G(J)}$ such that $r_1 = \mathrm{Hy}_{G(J)}(f_1)$ and $r_2 = \mathrm{Hy}_{G(J)}(f_2)$. Since $\mathcal{D}_{G(J)}$ is coherent, $f_1 + f_2 \in \mathcal{D}_{G(J)}$ and therefore, using the linearity of $\mathrm{Hy}_{G(J)}$, it follows that, indeed, $r_1 + r_2 = \mathrm{Hy}_{G(J)}(f_1) + \mathrm{Hy}_{G(J)}(f_2) = \mathrm{Hy}_{G(J)}(f_1 + f_2) \in \mathcal{R}_{n_G}$. The proof for D3 is similar. We finish the proof by showing that $\mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G}) = \mathcal{D}_{G(J)}$. Clearly, $\mathcal{D}_{G(J)}$ is a subset of $\mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$. To show that $\mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G}) \subseteq \mathcal{D}_{G(J)}$, choose any $f \in \mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$, implying that $r := \mathrm{Hy}_{G(J)}(f) \in \mathcal{R}_{n_G}$. Then, by definition of $\mathcal{R}_{n_G}$, there is some $f' \in \mathcal{D}_{G(J)}$ such that $\mathrm{Hy}_{G(J)}(f') = r$, whence $\mathrm{Hy}_{G(J)}(f - f') = 0$. Proposition 10 now implies that $f - f' \in \mathcal{I}^{par}_{G(J)}$, which, since $\mathcal{D}_{G(J)}$ is partially exchangeable, in turn implies that $f = f - f' + f' \in \mathcal{D}_{G(J)}$. To prove that $\mathcal{R}_{n_G}$ is unique, it suffices to realise that $\mathrm{Hy}_{G(J)}$ is a surjective (onto) map and that, consequently, $\mathcal{D}_{G(J)} = \mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$ implies that $\mathcal{R}_{n_G} = \mathrm{Hy}_{G(J)}(\mathcal{D}_{G(J)})$. □

Proof of Theorem 13. We start with sufficiency. Let $\underline{Q}_{n_G}$ be a coherent lower prevision on $\mathcal{G}(\mathcal{N}^{n_G})$ such that $\underline{P}_{G(J)} = \underline{Q}_{n_G} \circ \mathrm{Hy}_{G(J)}$. We first prove that $\underline{P}_{G(J)}$ is coherent. For P1, consider any $f \in \mathcal{G}(\mathcal{X}_{G(J)})$. Then $\underline{P}_{G(J)}(f) = \underline{Q}_{n_G}(\mathrm{Hy}_{G(J)}(f)) \geq \min(\mathrm{Hy}_{G(J)}(f)) \geq \min f$, where the first inequality follows from the coherence of $\underline{Q}_{n_G}$ and the second from the definition of $\mathrm{Hy}_{G(J)}(f)$. For P3, consider any $f_1, f_2 \in \mathcal{G}(\mathcal{X}_{G(J)})$. Then

$$\begin{aligned}\underline{P}_{G(J)}(f_1 + f_2) = \underline{Q}_{n_G}\big(\mathrm{Hy}_{G(J)}(f_1 + f_2)\big) &= \underline{Q}_{n_G}\big(\mathrm{Hy}_{G(J)}(f_1) + \mathrm{Hy}_{G(J)}(f_2)\big) \\ &\geq \underline{Q}_{n_G}\big(\mathrm{Hy}_{G(J)}(f_1)\big) + \underline{Q}_{n_G}\big(\mathrm{Hy}_{G(J)}(f_2)\big),\end{aligned}$$

where the final inequality follows, again, from the coherence of $\underline{Q}_{n_G}$. The proof for P2 is similar. To prove that $\underline{P}_{G(J)}$ is partially exchangeable, it is enough to realise that, for any $f \in \mathcal{I}^{par}_{G(J)}$, $\underline{P}_{G(J)}(f) = \underline{Q}_{n_G}(\mathrm{Hy}_{G(J)}(f)) = \underline{Q}_{n_G}(0) = 0$, where the second equality follows from Proposition 10 and the last one from the coherence of $\underline{Q}_{n_G}$.

For necessity, we exploit Proposition 2 to find that, given the coherence and partial exchangeability of $\underline{P}_{G(J)}$, there is a coherent and partially exchangeable $\mathcal{D}_{G(J)}$ for which $\underline{P}_{\mathcal{D}_{G(J)}} = \underline{P}_{G(J)}$. Then Theorem 12 guarantees that there is a coherent set $\mathcal{R}_{n_G} := \mathrm{Hy}_{G(J)}(\mathcal{D}_{G(J)})$ of desirable gambles on $\mathcal{N}^{n_G}$ such that $\mathcal{D}_{G(J)} = \mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$. Furthermore, by Proposition 1, $\underline{Q}_{n_G} := \underline{P}_{\mathcal{R}_{n_G}}$ is a coherent lower prevision on $\mathcal{G}(\mathcal{N}^{n_G})$. Given the relationship between $\mathcal{R}_{n_G}$ and $\mathcal{D}_{G(J)}$, it holds for all $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ that

$$\underline{P}_{G(J)}(f) = \underline{P}_{\mathcal{D}_{G(J)}}(f) = \sup\{\alpha\colon f - \alpha \in \mathcal{D}_{G(J)}\} = \sup\big\{\alpha\colon \mathrm{Hy}_{G(J)}(f) - \alpha \in \mathcal{R}_{n_G}\big\} = \underline{Q}_{n_G}\big(\mathrm{Hy}_{G(J)}(f)\big).$$

Hence, $\underline{P}_{G(J)} = \underline{Q}_{n_G} \circ \mathrm{Hy}_{G(J)}$.

To prove uniqueness, assume that $\underline{P}_{G(J)} = \underline{Q}_{n_G} \circ \mathrm{Hy}_{G(J)}$ and consider any gamble $r \in \mathcal{G}(\mathcal{N}^{n_G})$. Since, for all $m \in \mathcal{N}^{n_G}$

$$\mathrm{Hy}_{G(J)}(r \circ T_{G(J)})(m) = \frac{1}{|[m]|} \sum_{x \in [m]} r\big(T_{G(J)}(x)\big) = \frac{1}{|[m]|} \sum_{x \in [m]} r(m) = r(m),$$

we find that $\mathrm{Hy}_{G(J)}(r \circ T_{G(J)}) = r$, whence indeed

$$\underline{Q}_{n_G}(r) = \underline{Q}_{n_G}\big(\mathrm{Hy}_{G(J)}(r \circ T_{G(J)})\big) = \underline{P}_{G(J)}(r \circ T_{G(J)}). \qquad \square$$

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

Proof of Proposition 14. We first prove that the set $\{B_m : m \in \mathcal{N}^{n_G}\}$ generates $\mathcal{V}^{n_G}(\Sigma_G)$, meaning that every $h \in \mathcal{V}^{n_G}(\Sigma_G)$ can be written as a linear combination of Bernstein gambles B_m, $m \in \mathcal{N}^{n_G}$. So consider any $h \in \mathcal{V}^{n_G}(\Sigma_G)$. Then, by definition of $\mathcal{V}^{n_G}(\Sigma_G)$, h is the restriction to Σ_G of a polynomial p on $\times_{i\in G}\mathbb{R}^{\mathcal{X}_i}$ for which $\deg(p) \le n_G$. Since p is a polynomial, it is a linear combination of monomials in the variables $(\theta_i)_x$, $i \in G$ and $x \in \mathcal{X}_i$. Clearly, up to some non-zero coefficient, such a monomial is simply a Bernstein polynomial B_{m^*}, with $m^* \in \mathcal{N}^{n_G^*}$ for some $n_G^* \in \mathbb{N}^G$. Since $\deg(p) \le n_G$, we also know that $n_G^* \le n_G$. Consequently, h is a linear combination of Bernstein gambles B_{m^*}, with $m^* \in \mathcal{N}^{n_G^*}$ for some $n_G^* \in \mathbb{N}^G$ such that $n_G^* \le n_G$. Hence, it suffices to show that any such B_{m^*} can be written as a linear combination of Bernstein gambles B_m, $m \in \mathcal{N}^{n_G}$. Since $B_{m^*} = \prod_{i\in G} B_{m_i^*}$ and $B_m = \prod_{i\in G} B_{m_i}$, it clearly even suffices for each $B_{m_i^*}$ to be a linear combination of B_{m_i}, $m_i \in \mathcal{N}^{n_i}$. So consider any of those $B_{m_i^*}$. Then since $n_i^* \le n_i$ and $\sum_{x\in\mathcal{X}_i}(\theta_i)_x = 1$, we have for all $\theta_i \in \Sigma_i$ that

$$B_{m_i^*}(\theta_i) = \binom{n_i^*}{m_i^*}\prod_{x\in\mathcal{X}_i}(\theta_i)_x^{(m_i^*)_x} = \binom{n_i^*}{m_i^*}\prod_{x\in\mathcal{X}_i}(\theta_i)_x^{(m_i^*)_x}\left(\sum_{x\in\mathcal{X}_i}(\theta_i)_x\right)^{n_i-n_i^*},$$

where the last expression, after some elaborate but obvious and therefore omitted steps, can be written as a linear combination of $B_{m_i}(\theta_i)$, $m_i \in \mathcal{N}^{n_i}$. Hence, $B_{m_i^*}$ is a linear combination of B_{m_i}, $m_i \in \mathcal{N}^{n_i}$.

By the definition of a basis, the only thing left to prove now is that the Bernstein gambles B_m, $m \in \mathcal{N}^{n_G}$ are linearly independent. So consider real coefficients $b(m)$, $m \in \mathcal{N}^{n_G}$ such that $\sum_{m\in\mathcal{N}^{n_G}} b(m)B_m(\theta) = 0$ for all $\theta \in \Sigma_G$ and assume *ex absurdo* that $b(m) \neq 0$ for at least one $m \in \mathcal{N}^{n_G}$. Consider the polynomial p on $\times_{i\in G}\mathbb{R}^{\mathcal{X}_i}$ given by $p = \sum_{m\in\mathcal{N}^{n_G}} b(m)B_m$. Clearly, $p(\theta) = 0$ for all $\theta \in \Sigma_G$. Furthermore, if we denote the non-negative reals by $\mathbb{R}_{\ge 0}$, choose any $\lambda_G = (\lambda_1, \dots, \lambda_g) \in \mathbb{R}^G_{\ge 0}$ and define $\lambda_G\theta := (\lambda_1\theta_1, \dots, \lambda_g\theta_g)$, then for every $\theta \in \Sigma_G$, we have that $B_m(\lambda_G\theta) = B_m(\theta)\prod_{i\in G}\lambda_i^{n_i}$ for all $m \in \mathcal{N}^{n_G}$ and hence also $p(\lambda_G\theta) = p(\theta)\prod_{i\in G}\lambda_i^{n_i} = 0$. Since every $\theta \in \times_{i\in G}\mathbb{R}^{\mathcal{X}_i}_{\ge 0}$ can be written as $\lambda_G\theta$ for some $\theta \in \Sigma_G$ and $\lambda_G \in \mathbb{R}^G_{\ge 0}$, we find that p is zero on $\times_{i\in G}\mathbb{R}^{\mathcal{X}_i}_{\ge 0}$ and therefore, since p is a polynomial, that all of its coefficients are zero[31]: $b(m) = 0$ for all $m \in \mathcal{N}^{n_G}$. This is a contradiction. □

Proof of Proposition 15. The equivalence of the equalities $\mathcal{H}_{n_G} = \mathrm{CoMn}^{n_G}(\mathcal{R}_{n_G})$ and $\mathcal{R}_{n_G} = (\mathrm{CoMn}^{n_G})^{-1}(\mathcal{H}_{n_G})$ follows from CoMn^{n_G} being a linear isomorphism between the linear spaces $\mathcal{G}(\mathcal{N}^{n_G})$ and $\mathcal{V}^{n_G}(\Sigma_G)$. For that same reason, $\mathcal{H}_{n_G}$ satisfies $\mathrm{B}_{n_G}1$, $\mathrm{B}_{n_G}3$ and $\mathrm{B}_{n_G}4$ if and only if $\mathcal{R}_{n_G}$ satisfies D1, D3 and D4. Hence, the proof is concluded if we can show that $\mathcal{H}_{n_G}$ satisfies $\mathrm{B}_{n_G}2'$ if and only if $\mathcal{R}_{n_G}$ satisfies D2. Suppose that $\mathcal{R}_{n_G}$ satisfies D2 and consider any $h \in \mathcal{V}^{n_G}(\Sigma_G)$ such that $b_h^{n_G} > 0$. Then due to D2, $b_h^{n_G} \in \mathcal{R}_{n_G}$ and therefore $h = \mathrm{CoMn}^{n_G}(b_h^{n_G}) \in \mathrm{CoMn}^{n_G}(\mathcal{R}_{n_G}) = \mathcal{H}_{n_G}$. Conversely, suppose that $\mathcal{H}_{n_G}$ satisfies $\mathrm{B}_{n_G}2'$ and consider any $r \in \mathcal{G}(\mathcal{N}^{n_G})$ such that $r > 0$. Then since $h = \mathrm{CoMn}^{n_G}(r)$ is such that $b_h^{n_G} = r > 0$, we can use $\mathrm{B}_{n_G}2'$ to infer that $h \in \mathcal{H}_{n_G}$, which in turn implies that $r = b_h^{n_G} = (\mathrm{CoMn}^{n_G})^{-1}(h) \in (\mathrm{CoMn}^{n_G})^{-1}(\mathcal{H}_{n_G}) = \mathcal{R}_{n_G}$. □

Proof of Theorem 16. Due to Theorem 12, $\mathcal{D}_{G(J)}$ is coherent and partially exchangeable if and only if there is some coherent set $\mathcal{R}_{n_G}$ of desirable gambles on $\mathcal{N}^{n_G}$ such that $\mathcal{D}_{G(J)} = \mathrm{Hy}^{-1}_{G(J)}(\mathcal{R}_{n_G})$. Due to Proposition 15, and since $\mathrm{Mn}^{-1}_{G(J)} = \mathrm{Hy}^{-1}_{G(J)} \circ (\mathrm{CoMn}^{n_G})^{-1}$, this last requirement is in turn equivalent to the existence of some Bernstein coherent set $\mathcal{H}_{n_G}$ of polynomial gambles on Σ_G up to degree n_G such that $\mathcal{D}_{G(J)} = \mathrm{Mn}^{-1}_{G(J)}(\mathcal{H}_{n_G})$. Since $\mathrm{Hy}_{G(J)}$ and CoMn^{n_G} and therefore also $\mathrm{Mn}_{G(J)}$ are surjective maps, $\mathrm{Hy}_{G(J)} \circ \mathrm{Hy}^{-1}_{G(J)}$ and $\mathrm{Mn}_{G(J)} \circ \mathrm{Mn}^{-1}_{G(J)}$ are identity maps. Therefore, $\mathcal{D}_{G(J)} = \mathrm{Mn}^{-1}_{G(J)}(\mathcal{H}_{n_G})$ implies both $\mathcal{H}_{n_G} = \mathrm{Mn}_{G(J)}(\mathcal{D}_{G(J)})$ and $\mathcal{R}_{n_G} := \mathrm{Hy}_{G(J)}(\mathcal{D}_{G(J)}) = (\mathrm{CoMn}^{n_G})^{-1}(\mathcal{H}_{n_G})$. □

Proof of Proposition 17. It suffices to prove I1–I4. I1 is trivial. For I2 and I3, consider any $f, f_1, f_2 \in \mathcal{I}^{par}_{G(\mathbb{N})}$ and $\lambda \in \mathbb{R}$. Then, by Lemma 25, for large enough $J_i^* \in \mathbb{N}$, $i \in G$, we have that $f, f_1, f_2 \in \mathcal{I}^{par}_{G(J^*)}$. Applying Proposition 6,

[31] This intuitive result can be proven in multiple ways. One way is to consider a point $\theta^* \in \times_{i\in G}\mathbb{R}^{\mathcal{X}_i}_{>0}$ and to realise that, since p is zero on $\times_{i\in G}\mathbb{R}^{\mathcal{X}_i}_{\ge 0}$, all of the (mixed) partial derivatives of p in θ^* are zero. Evaluating these partial derivatives in θ^* leads to a homogeneous linear system of equalities in the coefficients of p which, when solved, leads to the desired result. Alternatively, one can also apply the Combinatorial Nullstellensatz [53, Theorem 1.2].

we find that $\lambda f \in \mathcal{I}^{par}_{G(J^*)}$ and $f_1 + f_2 \in \mathcal{I}^{par}_{G(J^*)}$. Hence, by Eq. (14), $\lambda f \in \mathcal{I}^{par}_{G(\mathbb{N})}$ and $f_1 + f_2 \in \mathcal{I}^{par}_{G(\mathbb{N})}$. For I4, consider any $f \in \mathcal{I}^{par}_{G(\mathbb{N})}$. Then, there are $J_i \Subset \mathbb{N}$, $i \in G$, for which $f \in \mathcal{I}^{par}_{G(J)}$. By Proposition 6, we know that $f \not> 0$ and $f \not< 0$. □

Lemma 25. *For all $i \in G$, consider $J_i, J_i^* \Subset \mathbb{N}$ such that $J_i \subseteq J_i^*$. Then $\mathcal{I}^{par}_{G(J)} \subseteq \mathcal{I}^{par}_{G(J^*)}$.*

Proof. Consider any gamble $f' \in \mathcal{A}^{par}_{G(J)}$, so $f' = f - \pi^t f$ for some $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ and some $\pi \in \mathcal{P}_{G(J)}$. Then $f' \in \mathcal{I}^{par}_{G(J^*)}$ because its cylindrical extension to $\mathcal{X}_{G(J^*)}$ is equal to $\tilde{f}' = \tilde{f} - \tilde{\pi}^t \tilde{f}$, where $\tilde{f}$ is the cylindrical extension of f to $\mathcal{X}_{G(J^*)}$ and $\tilde{\pi} \in \mathcal{P}_{G(J^*)}$ is chosen such that it coincides with π on the indices in J_i, $i \in G$, and, on all other indices, coincides with the identity permutation. The result now follows from Eq. (8) and the fact that $\mathcal{I}^{par}_{G(J^*)}$ is a linear space [Proposition 6]. □

Proof of Proposition 18. For necessity, assume that $\mathcal{D}_{G(\mathbb{N})}$ is partially exchangeable, and consider any $J_i \Subset \mathbb{N}$, with $i \in G$. It suffices to prove that $\mathcal{D}_{G(J)} + \mathcal{I}^{par}_{G(J)} \subseteq \mathcal{D}_{G(J)}$. Take any $f \in \mathcal{D}_{G(J)}$ and any $f' \in \mathcal{I}^{par}_{G(J)}$. Due to marginalisation [Eq. (6)], we know that $f \in \mathcal{D}_{G(\mathbb{N})} \cap \mathcal{G}(\mathcal{X}_{G(J)})$ and, due to Eq. (14), we know that $f' \in \mathcal{I}^{par}_{G(\mathbb{N})} \cap \mathcal{G}(\mathcal{X}_{G(J)})$. Combining these observations, we infer that $f + f' \in (\mathcal{D}_{G(\mathbb{N})} + \mathcal{I}^{par}_{G(\mathbb{N})}) \cap \mathcal{G}(\mathcal{X}_{G(J)}) = \mathcal{D}_{G(\mathbb{N})} \cap \mathcal{G}(\mathcal{X}_{G(J)})$, where the last equality holds because $\mathcal{D}_{G(\mathbb{N})}$ is partially exchangeable by assumption. Hence, using marginalisation [Eq. (6)], we find that $f + f' \in \mathcal{D}_{G(J)}$.

Next, we turn to sufficiency. Consider any $f \in \mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and $f' \in \mathcal{I}^{par}_{G(\mathbb{N})}$. Clearly, both f and f' depend only upon the values of a finite number of variables. This means that there are $J_i, J_i' \Subset \mathbb{N}$, $i \in G$, such that $f \in \mathcal{D}_{G(J)}$ and $f' \in \mathcal{I}^{par}_{G(J')}$. Let $J_i^* := J_i \cup J_i'$ for all $i \in G$. Then $f \in \mathcal{D}_{G(J^*)}$ because of marginalisation [Eq. (6)], and $f' \in \mathcal{I}^{par}_{G(J^*)}$ by Lemma 25. Since $\mathcal{D}_{G(J^*)}$ is partially exchangeable by assumption, we have that $f + f' \in \mathcal{D}_{G(J^*)}$. Using marginalisation [Eq. (6)], we find that $f + f' \in \mathcal{D}_{G(\mathbb{N})}$. Hence, $\mathcal{D}_{G(\mathbb{N})}$ is partially exchangeable. □

Proof of Proposition 19. For necessity, assume that $\underline{P}_{G(\mathbb{N})}$ is partially exchangeable, meaning that $\underline{P}_{G(\mathbb{N})}(f) = \overline{P}_{G(\mathbb{N})}(f) = 0$ for all $f \in \mathcal{I}^{par}_{G(\mathbb{N})}$. Fix any $J_i \Subset \mathbb{N}$, with $i \in G$. Then, due to Eq. (14) and marginalisation [Eq. (7)], we find that $\underline{P}_{G(J)}(f) = \overline{P}_{G(J)}(f) = 0$ for all $f \in \mathcal{I}^{par}_{G(J)}$, meaning that $\underline{P}_{G(J)}$ is partially exchangeable.

Next, we turn to sufficiency. Consider any $f \in \mathcal{I}^{par}_{G(\mathbb{N})}$. By definition [Eq. (14)], there are $J_i \Subset \mathbb{N}$, $i \in G$ such that $f \in \mathcal{I}^{par}_{G(J)}$. Because $\underline{P}_{G(J)}$ is partially exchangeable by assumption, we have that $\underline{P}_{G(J)}(f) = \overline{P}_{G(J)}(f) = 0$. By Eq. (7), and because $f \in \mathcal{G}(\mathcal{X}_{G(J)})$, this implies that $\underline{P}_{G(\mathbb{N})}(f) = \overline{P}_{G(\mathbb{N})}(f) = 0$. Hence, $\underline{P}_{G(\mathbb{N})}$ is partially exchangeable. □

Proof of Proposition 20. This is a direct consequence of Propositions 9 and 19. □

Proof of Proposition 21. We only prove the converse implication since the direct one is trivial. So consider any set $\mathcal{H}$ of polynomial gambles on Σ_G such that, for all $n_G \in \mathbb{N}^G$, $\mathcal{H}_{n_G} := \mathcal{H} \cap \mathcal{V}^{n_G}(\Sigma_G)$ is Bernstein coherent at degree n_G. Then clearly, $\mathcal{H}$ satisfies B2 and B1. For B3 and B4, consider any $h, h_1, h_2 \in \mathcal{H}$ and $\lambda \in \mathbb{R}_{>0}$. If we choose $n_G \in \mathbb{N}^G$ such that its components n_i, $i \in G$ are high enough, we have that $h, h_1, h_2 \in \mathcal{H}_{n_G}$. Since, by assumption, $\mathcal{H}_{n_G}$ satisfies $\mathrm{B}_{n_G}3$ and $\mathrm{B}_{n_G}4$, we find that $\lambda h \in \mathcal{H}_{n_G} \subseteq \mathcal{H}$ and $h_1 + h_2 \in \mathcal{H}_{n_G} \subseteq \mathcal{H}$. □

Proof of Theorem 22. Consider any set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ and recall that for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, $\mathcal{D}_{G(J)}$ is the corresponding marginal model for $X_{G(J)}$, as given by Eq. (6).

We start by proving the converse implication (sufficiency). Assume that there is some Bernstein coherent set $\mathcal{H}$ of polynomial gambles on Σ_G such that for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, $\mathcal{D}_{G(J)} = \mathrm{Mn}^{-1}_{G(J)}(\mathcal{H}_{n_G})$, with $\mathcal{H}_{n_G} := \mathcal{H} \cap \mathcal{V}^{n_G}(\Sigma_G)$. Now consider any such choice of $J_i \Subset \mathbb{N}$, with $i \in G$. Then, by Proposition 21, $\mathcal{H}_{n_G}$ is Bernstein coherent at degree n_G and therefore, by Theorem 16, $\mathcal{D}_{G(J)}$ is coherent and partially exchangeable. Since this holds for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, we derive from Proposition 4 that $\mathcal{D}_{G(\mathbb{N})}$ is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and from Proposition 18 that $\mathcal{D}_{G(\mathbb{N})}$ is partially exchangeable.

Next, we show that $\mathcal{H} = \bigcup_{n_G \in \mathbb{N}^G} \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)})$, thus making it unique. Consider any $h \in \mathcal{H}$. Then, by the definition of a polynomial gamble, there is some $n_G^* \in \mathbb{N}^G$ such that $h \in \mathcal{H} \cap \mathcal{V}^{n_G^*}(\Sigma_G) = \mathcal{H}_{n_G^*}$. Since, by as-

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

sumption, $\mathcal{D}_{G(N^*)} = \mathrm{Mn}^{-1}_{G(N^*)}(\mathcal{H}_{n^*_G})$, there is some $f \in \mathcal{D}_{G(N^*)}$ for which $\mathrm{Mn}_{G(N^*)}(f) = h$. Hence, we find that $h \in \mathrm{Mn}_{G(N^*)}(\mathcal{D}_{G(N^*)}) \subseteq \bigcup_{n_G\in\mathbb{N}^G} \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)})$. Conversely, consider any $n_G \in \mathbb{N}^G$. Then, since by assumption, $\mathcal{D}_{G(N)} = \mathrm{Mn}^{-1}_{G(N)}(\mathcal{H}_{n_G})$, we find that

$$\mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)}) = \mathrm{Mn}_{G(N)}\big(\mathrm{Mn}^{-1}_{G(N)}(\mathcal{H}_{n_G})\big) \subseteq \mathcal{H}_{n_G} \subseteq \mathcal{H}.$$

We complete the proof by proving the direct implication (necessity). Assume that $\mathcal{D}_{G(\mathbb{N})}$ is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and partially exchangeable. We then derive from Propositions 4 and 18 that, for every choice of $J_i \Subset \mathbb{N}$ with $i \in G$, $\mathcal{D}_{G(J)}$ is coherent and partially exchangeable. Therefore, by Theorem 16, $\mathcal{D}_{G(J)}$ has a polynomial representation $\mathcal{H}_{n_G} = \mathrm{Mn}_{G(J)}(\mathcal{D}_{G(J)})$ that is Bernstein coherent at degree n_G and for which $\mathcal{D}_{G(J)} = \mathrm{Mn}^{-1}_{G(J)}(\mathcal{H}_{n_G})$. Furthermore, by Lemma 27, $\mathcal{H}_{n_G}$ does not depend on the particular choice of the J_i, $i \in G$, as long as $|J_i| = n_i$. Now let $\mathcal{H} := \bigcup_{n_G\in\mathbb{N}^G} \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)}) = \bigcup_{n_G\in\mathbb{N}^G} \mathcal{H}_{n_G}$. Then the direct implication will follow from Proposition 21, provided we can show that, for all $n^*_G \in \mathbb{N}^G$, $\mathcal{H} \cap \mathcal{V}^{n^*_G}(\Sigma_G) = \mathcal{H}_{n^*_G}$. So consider any $n^*_G \in \mathbb{N}^G$. Then $\mathcal{H}_{n^*_G} = \mathcal{H}_{n^*_G} \cap \mathcal{V}^{n^*_G}(\Sigma_G) \subseteq \mathcal{H} \cap \mathcal{V}^{n^*_G}(\Sigma_G)$, so we are left to prove that, conversely, $\mathcal{H} \cap \mathcal{V}^{n^*_G}(\Sigma_G) \subseteq \mathcal{H}_{n^*_G}$ or, equivalently, that, for all $n_G \in \mathbb{N}^G$, $\mathcal{H}_{n_G} \cap \mathcal{V}^{n^*_G}(\Sigma_G) \subseteq \mathcal{H}_{n^*_G}$. So consider any $n_G \in \mathbb{N}^G$ and any $h \in \mathcal{H}_{n_G} \cap \mathcal{V}^{n^*_G}(\Sigma_G)$ and let n'_G be the pointwise maximum of n_G and n^*_G. Since $h \in \mathcal{H}_{n_G}$, there is some $f \in \mathcal{D}_{G(N)}$ such that $\mathrm{Mn}_{G(N)}(f) = h$. Now let $\tilde{f}$ be the cylindrical extension of f to $\mathcal{X}_{G(N')}$. Then, due to marginalisation, $\tilde{f} \in \mathcal{D}_{G(N')}$ and, by Lemma 26, $\mathrm{Mn}_{G(N')}(\tilde{f}) = \mathrm{Mn}_{G(N)}(f) = h$. Hence, $h \in \mathcal{H}_{n'_G}$ and $\mathrm{Mn}^{-1}_{G(N')}(\{h\}) \subseteq \mathcal{D}_{G(N')}$. Since $h \in \mathcal{V}^{n^*_G}(\Sigma_G)$, there is some $f^* \in \mathcal{G}(\mathcal{X}_{G(N^*)})$ such that $\mathrm{Mn}_{G(N^*)}(f^*) = h$. Now let $\tilde{f}^*$ be the cylindrical extension of f^* to $\mathcal{X}_{G(N')}$. Then we infer from Lemma 26 that $\mathrm{Mn}_{G(N')}(\tilde{f}^*) = \mathrm{Mn}_{G(N^*)}(f^*) = h$, implying that $\tilde{f}^* \in \mathrm{Mn}^{-1}_{G(N')}(\{h\}) \subseteq \mathcal{D}_{G(N')}$. Since $\mathcal{D}_{G(N')}$ and $\mathcal{D}_{G(N^*)}$ are related through marginalisation, this in turn implies that $f^* \in \mathcal{D}_{G(N^*)}$ and therefore $h \in \mathcal{H}_{n^*_G}$. □

Lemma 26. *For all $i \in G$, consider subsets $J_i, J^*_i \Subset \mathbb{N}$ such that $J_i \subseteq J^*_i$. It then holds for all $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ that $\mathrm{Mn}_{G(J^*)}(\tilde{f}) = \mathrm{Mn}_{G(J)}(f)$, where $\tilde{f}$ is the cylindrical extension of f to $\mathcal{X}_{G(J^*)}$.*

Proof. For all $i \in G$, let $J'_i := J^*_i \setminus J_i$. Then, for all θ in Σ_G,

$$\begin{aligned}
\mathrm{Mn}_{G(J^*)}(\tilde{f})(\theta) &= \sum_{x_{G(J^*)}\in\mathcal{X}_{G(J^*)}} \tilde{f}(x_{G(J^*)}) \prod_{i\in G}\prod_{j\in J^*_i} (\theta_i)_{x_{ij}} \\
&= \sum_{x_{G(J)}\in\mathcal{X}_{G(J)}} \sum_{x_{G(J')}\in\mathcal{X}_{G(J')}} \tilde{f}(x_{G(J)}, x_{G(J')}) \prod_{i\in G}\Big(\prod_{j\in J_i} (\theta_i)_{x_{ij}}\Big)\Big(\prod_{j\in J'_i} (\theta_i)_{x_{ij}}\Big) \\
&= \sum_{x_{G(J)}\in\mathcal{X}_{G(J)}} f(x_{G(J)}) \prod_{i\in G}\prod_{j\in J_i} (\theta_i)_{x_{ij}} \sum_{x_{G(J')}\in\mathcal{X}_{G(J')}} \prod_{i\in G}\prod_{j\in J'_i} (\theta_i)_{x_{ij}} \\
&= \mathrm{Mn}_{G(J)}(f)(\theta) \prod_{i\in G}\prod_{j\in J'_i} \sum_{x_{ij}\in\mathcal{X}_{ij}} (\theta_i)_{x_{ij}} = \mathrm{Mn}_{G(J)}(f)(\theta),
\end{aligned}$$

where the third equality follows from Eq. (4) and the final equality from the fact that, for all $i \in G$ and $j \in J'_i$, $\sum_{x_{ij}\in\mathcal{X}_{ij}} (\theta_i)_{x_{ij}} = \sum_{x\in\mathcal{X}_i} (\theta_i)_x = 1$. □

Lemma 27. *Consider a set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ that is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and partially exchangeable. Then, for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, $\mathrm{Mn}_{G(J)}(\mathcal{D}_{G(J)}) = \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)})$.*

Proof. We start by proving that $\mathrm{Mn}_{G(J)}(\mathcal{D}_{G(J)}) \subseteq \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)})$ or, equivalently, that $h \in \mathrm{Mn}_{G(J)}(\mathcal{D}_{G(J)})$ implies that $h \in \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)})$. So consider any $h \in \mathrm{Mn}_{G(J)}(\mathcal{D}_{G(J)})$. Then there is some $f \in \mathcal{D}_{G(J)}$ for which $\mathrm{Mn}_{G(J)}(f) = h$. Define now, for all $i \in G$, $J^*_i = J_i \cup N_i$ and let $\tilde{f}$ be the cylindrical extension of f to $\mathcal{X}_{G(J^*)}$. Choose $\pi \in \mathcal{P}_{G(J^*)}$ such that, for all $i \in G$, $\pi_i(J_i) = N_i$. Then clearly, $\pi^t(\tilde{f})$ can be identified with a gamble $f' \in \mathcal{G}(\mathcal{X}_{G(N)})$, meaning that $\pi^t(\tilde{f}) = \tilde{f}'$, with $\tilde{f}'$ the cylindrical extension of f' to $\mathcal{X}_{G(J^*)}$. We now have that

$$h = \mathrm{Mn}_{G(J)}(f) = \mathrm{Mn}_{G(J^*)}(\tilde{f}) = \mathrm{Mn}_{G(J^*)}\big(\pi^t \tilde{f}\big) = \mathrm{Mn}_{G(J^*)}\big(\tilde{f}'\big) = \mathrm{Mn}_{G(N)}\big(f'\big),$$

where the third equality follows from the definition of $\mathrm{Mn}_{G(J^*)}$, and the second and last equalities from Lemma 26. Furthermore, since $\mathcal{D}_{G(J)}$ and $\mathcal{D}_{G(J^*)}$ are related through marginalisation, $f \in \mathcal{D}_{G(J)}$ implies that $\tilde{f} \in \mathcal{D}_{G(J^*)}$. Since $\mathcal{D}_{G(\mathbb{N})}$ is partially exchangeable, we know from Proposition 18 that $\mathcal{D}_{G(J^*)}$ is partially exchangeable as well. Using ID1, we now infer from $\tilde{f} \in \mathcal{D}_{G(J^*)}$ that $\pi^t \tilde{f} = \tilde{f} - (\tilde{f} - \pi^t \tilde{f}) \in \mathcal{D}_{G(J^*)}$ or, equivalently, that $\tilde{f}' \in \mathcal{D}_{G(J^*)}$. Since $\mathcal{D}_{G(N)}$ and $\mathcal{D}_{G(J^*)}$ are related through marginalisation, this in turn implies that $f' \in \mathcal{D}_{G(N)}$. Hence, we find that, indeed, $h = \mathrm{Mn}_{G(N)}(f') \in \mathrm{Mn}_{G(N)}(\mathcal{D}_{G(N)})$.

The proof for the converse inclusion is completely analogous. It suffices to interchange J and N in the reasoning above. □

Proof of Theorem 23. Consider any lower prevision $\underline{P}_{G(\mathbb{N})}$ on $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and recall that for every choice of $J_i \Subset \mathbb{N}$ with $i \in G$, $\underline{P}_{G(J)}$ is the corresponding marginal lower prevision on $\mathcal{G}(\mathcal{X}_{G(J)})$, as given by Eq. (7).

We start by proving the converse implication (sufficiency). Assume that there is some coherent lower prevision $\underline{R}$ on $\mathcal{V}(\Sigma_G)$ such that for every choice of $J_i \Subset \mathbb{N}$, with $i \in G$, and every $f \in \mathcal{G}(\mathcal{X}_{G(J)})$, $\underline{P}_{G(J)}(f) = \underline{R}(\mathrm{Mn}_{G(J)}(f))$. We will show that each of these $\underline{P}_{G(J)}$ is coherent and partially exchangeable, which, due to Propositions 5 and 19, will then imply that $\underline{P}_{G(\mathbb{N})}$ is coherent and partially exchangeable. So choose $J_i \Subset \mathbb{N}$ with $i \in G$. Then since $\mathrm{Mn}_{G(J)}$ is a linear operator, $\underline{P}_{G(J)}$ clearly satisfies coherence properties P2 and P3 because $\underline{R}$ does. Furthermore, since $\min \mathrm{Mn}_{G(J)}(f) \geq \min f$ for all $f \in \mathcal{G}(\mathcal{X}_{G(J)})$, $\underline{P}_{G(J)}$ satisfies P1 because $\underline{R}$ does. Hence, $\underline{P}_{G(J)}$ is coherent. To show that it is also partially exchangeable, consider any $f \in \mathcal{I}^{par}_{G(J)}$. By Proposition 17 and Definition 4, it suffices to show that $\underline{P}_{G(J)}(f) \geq 0$ or, equivalently, that $\underline{R}(\mathrm{Mn}_{G(J)}(f)) \geq 0$. Since $\mathrm{Mn}_{G(J)}(f) = 0$ by Proposition 10 and the definition of $\mathrm{Mn}_{G(J)}$, this follows trivially from the coherence of $\underline{R}$.

To prove the uniqueness of $\underline{R}$, consider any $n_G \in \mathbb{N}^G$ and any $h \in \mathcal{V}^{n_G}(\Sigma_G)$. Then, as shown near the end of our proof for Theorem 13, $\mathrm{Hy}_{G(N)}(b_h^{n_G} \circ T_{G(N)}) = b_h^{n_G}$. Since we also know that $\mathrm{CoMn}^{n_G}(b_h^{n_G}) = h$, we find that $\mathrm{Mn}_{G(N)}(b_h^{n_G} \circ T_{G(N)}) = \mathrm{CoMn}^{n_G}(\mathrm{Hy}_{G(N)}(b_h^{n_G} \circ T_{G(N)})) = h$ and, consequently, that

$$\underline{R}(h) = \underline{R}\big(\mathrm{Mn}_{G(N)}\big(b_h^{n_G} \circ T_{G(N)}\big)\big) = \underline{P}_{G(N)}\big(b_h^{n_G} \circ T_{G(N)}\big).$$

To complete the proof, we now turn to the direct implication (necessity). Assume that $\underline{P}_{G(\mathbb{N})}$ is coherent and partially exchangeable. Then we infer from Propositions 2 and 17 that there is some set $\mathcal{D}_{G(\mathbb{N})} \subseteq \overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ of desirable gambles on $\mathcal{X}_{G(\mathbb{N})}$ that is coherent relative to $\overline{\mathcal{G}}(\mathcal{X}_{G(\mathbb{N})})$ and partially exchangeable, and for which $\underline{P}_{G(\mathbb{N})} = \underline{P}_{\mathcal{D}_{G(\mathbb{N})}}$. Due to Theorem 22, this in turn implies the existence of a Bernstein coherent set $\mathcal{H}$ of polynomial gambles on Σ_G such that for every choice of $J_i \Subset \mathbb{N}$ with $i \in G$, $\mathcal{D}_{G(J)} = \mathrm{Mn}^{-1}_{G(J)}(\mathcal{H}_{n_G})$, with $\mathcal{H}_{n_G} := \mathcal{H} \cap \mathcal{V}^{n_G}(\Sigma_G)$. We now let $\underline{R} := \underline{P}_{\mathcal{H}}$, with $\underline{P}_{\mathcal{H}}$ given by Eq. (1). Then, by Lemma 28, $\underline{R}$ is a coherent lower prevision on $\mathcal{V}(\Sigma_G)$. Furthermore, for every choice of $J_i \Subset \mathbb{N}$ with $i \in G$, we find for all $f \in \mathcal{G}(\mathcal{X}_{G(J)})$ that

$$\begin{aligned}\underline{P}_{G(J)}(f) &= \sup\{\alpha \in \mathbb{R}\colon f - \alpha \in \mathcal{D}_{G(J)}\} = \sup\big\{\alpha \in \mathbb{R}\colon f - \alpha \in \mathrm{Mn}^{-1}_{G(J)}(\mathcal{H}_{n_G})\big\}\\ &= \sup\big\{\alpha \in \mathbb{R}\colon \mathrm{Mn}_{G(J)}(f - \alpha) \in \mathcal{H}_{n_G}\big\}\\ &= \sup\big\{\alpha \in \mathbb{R}\colon \mathrm{Mn}_{G(J)}(f) - \alpha \in \mathcal{H}\big\} = \underline{R}\big(\mathrm{Mn}_{G(J)}(f)\big). \qquad \square\end{aligned}$$

Lemma 28. *For any Bernstein coherent set $\mathcal{H}$ of polynomial gambles on Σ_G, the corresponding lower prevision $\underline{P}_{\mathcal{H}}$ on $\mathcal{V}(\Sigma_G)$, as defined through Eq. (1), is coherent.*

Proof. Due to Eq. (1), and since $\mathcal{H}$ satisfies B3 and B4, $\underline{P}_{\mathcal{H}}$ clearly satisfies P2 and P3. To prove that $\underline{P}_{\mathcal{H}}$ satisfies P1, we need to show, for all $h \in \mathcal{V}(\Sigma_G)$, that $\underline{P}_{\mathcal{H}}(h) \geq \inf h$ or, equivalently, due to Eq. (1), that for all $\epsilon > 0$, $h^* := h - \inf h + \epsilon \in \mathcal{H}$. So consider any $\epsilon > 0$ and any $h \in \mathcal{V}(\Sigma_G)$, meaning that there is some $n_G \in \mathbb{N}^G$ such that $h \in \mathcal{V}^{n_G}(\Sigma_G)$. Let $h' := h - \inf h$ and let p' be the polynomial on $\times_{i \in G} \mathbb{R}^{\mathcal{X}_i}$ that is given by $p' := \sum_{m \in \mathcal{N}^{n_G}} b_{h'}^{n_G}(m) B_m$. Then clearly, $p'(\theta) = h'(\theta) \geq 0$ for all $\theta \in \Sigma_G$ and therefore, by an argument similar to that used near the end of the proof of Proposition 14, $p'(\theta) \geq 0$ for all $\theta \in \times_{i \in G} \mathbb{R}^{\mathcal{X}_i}_{\geq 0}$. Hence, if we let $p := p' + \epsilon (1/g \sum_{i \in G} \sum_{x \in \mathcal{X}_i} (\theta_i)_x)^{\sum_{i \in G} n_i}$, then $p(\theta) > 0$ for all $\theta \in \times_{i \in G} \mathbb{R}^{\mathcal{X}_i}_{\geq 0} \setminus \{0\}$. Because p is also homogeneous, we can apply Pólya's result [54, Theorem 5.5.1] to infer that there is some $k \in \mathbb{N}$ such that all of the coefficients of $(\sum_{i \in G} \sum_{x \in \mathcal{X}_i} (\theta_i)_x)^k p$ and therefore also those of $p^* := (1/g \sum_{i\in G} \sum_{x\in\mathcal{X}_i} (\theta_i)_x)^k p$ are non-negative, implying that p^* is a positive linear combination of Bernstein polynomials [clearly it is not possible for all the coefficients to be zero]. Since, for all $\theta \in \Sigma_G$, $1/g \sum_{i\in G} \sum_{x\in\mathcal{X}_i} (\theta_i)_x = 1$ and $h'(\theta) = p'(\theta)$, and therefore also $p^*(\theta) = p(\theta) = h'(\theta) + \epsilon = h^*(\theta)$, we find that h^* is a positive linear combination of Bernstein gambles. Hence, since $\mathcal{H}$ satisfies B2, B3 and B4, indeed $h^* \in \mathcal{H}$. $\square$

References

[1] B. de Finetti, Colloque Consacré à la Théorie des Probabilités, VI Partie, vol. 793, Herman & Cie, Paris, 1938, pp. 5–18, Translated as Chapter 9 of [44].

[2] G. Link, Representation theorems of the de Finetti type for (partially) symmetric probability measures, in: R.C. Jeffrey (Ed.), Studies in Inductive Logic and Probability, vol. 2, University of California Press, 1980, pp. 207–231.

[3] P. Diaconis, D. Freedman, de Finetti's generalizations of exchangeability, in: R.C. Jeffrey (Ed.), Studies in Inductive Logic and Probability, vol. 2, University of California Press, 1980, pp. 233–249.

[4] P. Diaconis, D. Freedman, de Finetti's theorem for Markov chains, Ann. Probab. 8 (1980) 115–130.

[5] P. Diaconis, D. Freedman, Partial exchangeability and sufficiency, in: G.J.K. Roy (Ed.), Proceedings of the Indian Statistical Institute Golden Jubilee International Conference on Statistics: Applications and New Directions, Indian Statistical Institute, Calcutta, 1984, pp. 205–236.

[6] D.J. Aldous, Representations for partially exchangeable arrays of random variables, J. Multivar. Anal. (1981) 581–598.

[7] D.N. Hoover, Row–column exchangeability and a generalized model for probability, in: G. Koch, F. Spizzichino (Eds.), Exchangeability in Probability and Statistics, North-Holland, 1982, pp. 281–291.

[8] B. de Finetti, Theory of Probability: A Critical Introductory Treatment, John Wiley & Sons, Chichester, 1974–1975, English translation of [55], two volumes.

[9] C.A.B. Smith, Consistency in statistical inference and decision, J. R. Stat. Soc. A 23 (1961) 1–37.

[10] P.M. Williams, Notes on conditional previsions, Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975.

[11] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.

[12] S. Moral, Epistemic irrelevance on sets of desirable gambles, Ann. Math. Artif. Intell. 45 (2005) 197–214.

[13] G. de Cooman, F. Hermans, Imprecise probability trees: bridging two theories of imprecise probability, Artif. Intell. 172 (2008) 1400–1427.

[14] M. Zaffalon, E. Miranda, Notes on desirability and conditional lower previsions, Ann. Math. Artif. Intell. 66 (2010) 251–309.

[15] G. de Cooman, E. Quaeghebeur, Exchangeability and sets of desirable gambles, Int. J. Approx. Reason. 53 (2012) 363–395, Special issue in honour of Henry E. Kyburg, Jr..

[16] I. Couso, S. Moral, Sets of desirable gambles: conditioning, representation, and precise probabilities, Int. J. Approx. Reason. 52 (2011) 1034–1055.

[17] G. de Cooman, E. Miranda, Irrelevant and independent natural extension for sets of desirable gambles, J. Artif. Intell. Res. 45 (2012) 601–640.

[18] J. De Bock, G. de Cooman, Credal networks under epistemic irrelevance using sets of desirable gambles, in: Proceedings of the Eighth International Symposium on Imprecise Probability: Theories and Applications, 2013, pp. 99–108.

[19] P. Walley, Towards a unified theory of imprecise probability, Int. J. Approx. Reason. 24 (2000) 125–148.

[20] G. de Cooman, E. Miranda, Symmetry of models versus models of symmetry, in: W.L. Harper, G.R. Wheeler (Eds.), Probability and Inference: Essays in Honor of Henry E. Kyburg, Jr., King's College Publications, 2007, pp. 67–149.

[21] G. de Cooman, E. Quaeghebeur, E. Miranda, Exchangeable lower previsions, Bernoulli 15 (2009) 721–735.

[22] E. Quaeghebeur, G. de Cooman, F. Hermans, Accept & reject statement-based uncertainty models, arXiv:1208.4462v4 [math.PR], 2014, submitted for publication.

[23] F. Hermans, An operational approach to graphical uncertainty modelling, Ph.D. thesis, Ghent University, 2012, http://hdl.handle.net/1854/LU-2153963.

[24] L.E. Dubins, L.J. Savage, Inequalities for Stochastic Processes—How to Gamble If You Must, McGraw-Hill, New York, 1965, Reprinted by Dover Publications in 1976.

[25] F.G. Cozman, Sets of probability distributions, independence, and convexity, Synthese 186 (2012) 577–600.

[26] A.I. Dale, Some results in partial exchangeability, Studia Sci. Math. Hung. 19 (1984) 233–244.

[27] J.M. Bernardo, A.F.M. Smith, Bayesian Theory, John Wiley & Sons, Chichester, 1994.

[28] G. de Cooman, E. Miranda, The F. Riesz representation theorem and finite additivity, in: D. Dubois, M.A. Lubiano, H. Prade, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Soft Methods for Handling Variability and Imprecision (Proceedings of SMPS 2008), 2008, pp. 243–252.

[29] E. Miranda, G. de Cooman, E. Quaeghebeur, The Hausdorff moment problem under finite additivity, J. Theor. Probab. 20 (2007) 663–693.

[30] E. Miranda, G. de Cooman, E. Quaeghebeur, Finitely additive extensions of distribution functions and moment sequences: the coherent lower prevision approach, Int. J. Approx. Reason. 48 (2008) 132–155.

[31] S. Ferson, V. Kreinovich, L. Ginzburg, D.S. Myers, K. Sentz, Constructing probability boxes and Dempster–Shafer structures, Technical report SAND2002-4015, Sandia National Laboratories, 2003.

[32] G. de Cooman, J. De Bock, M.A. Diniz, Coherent predictive inference under exchangeability with imprecise probabilities, J. Artif. Intell. Res. (2014), in press.

[33] L.E. Dubins, Finitely additive conditional probabilities, conglomerability and disintegrations, Ann. Probab. 3 (1975) 88–99.

[34] G. Coletti, R. Scozzafava, Probabilistic Logic in a Coherent Setting, Trends Log., vol. 15, Kluwer, Dordrecht, 2002.

[35] F.G. Cozman, T. Seidenfeld, Independence for full conditional measures and their graphoid properties, in: Foundations of the Formal Sciences VI, Reasoning About Probabilities and Probabilistic Reasoning, College Publications, London, 2009, pp. 1–29.

Please cite this article in press as: J. De Bock et al., Representation theorems for partially exchangeable random variables, Fuzzy Sets and Systems (2014), http://dx.doi.org/10.1016/j.fss.2014.10.027

[36] P. Walley, Inferences from multinomial data: learning about a bag of marbles, J. R. Stat. Soc. B 58 (1996) 3–57, with discussion.
[37] P. Walley, J.-M. Bernard, Imprecise probabilistic prediction for categorical data, Technical report CAF-9901, Laboratoire Cognition et Activitées Finalisées, Université de Paris 8, 1999.
[38] P. Walley, A bounded derivative model for prior ignorance about a real-valued parameter, Scand. J. Stat. 24 (1997) 463–483.
[39] A. Benavoli, M. Zaffalon, A model of prior ignorance for inferences in the one-parameter exponential family, J. Stat. Plan. Inference 142 (2012) 1960–1979.
[40] F. Mangili, A. Benavoli, New prior near-ignorance models on the simplex, in: F. Cozman, T. Denœux, S. Destercke, T. Seidenfeld (Eds.), Proceedings of the Eighth International Symposium on Imprecise Probability: Theories and Applications, ISIPTA'13, SIPTA, 2013, pp. 213–222.
[41] G. de Cooman, J. De Bock, M.A. Diniz, Predictive inference under exchangeability, and the Imprecise Dirichlet Multinomial Model, in: Proceedings of the 12th Brazilian Meeting on Bayesian Statistics, 2014, in press.
[42] J.-M. Bernard, An introduction to the imprecise Dirichlet model for multinomial data, Int. J. Approx. Reason. 39 (2005) 123–150.
[43] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 2006.
[44] B. de Finetti, Probability, Induction and Statistics, Wiley, London, 1972.
[45] A. Bruno, Sugli eventi parzialmente scambiabili, G. Ist. Ital. Attuari 27 (1964) 174–196, Translated as Chapter 10 of [44].
[46] J. Martin, Bayesian Decision Processes and Markov chains, Wiley, New York, 1967.
[47] S. Petrone, A. Raftery, A note on the Dirichlet process prior in Bayesian nonparametric inference with partial exchangeability, Stat. Probab. Lett. 36 (2000) 69–83.
[48] A. Gugliemi, E. Melilli, Approximating de Finetti's measures for partially exchangeable sequences, Stat. Probab. Lett. 48 (2000) 309–315.
[49] S. Walker, P. Muliere, A bivariate Dirichlet process, Stat. Probab. Lett. 64 (2003) 1–7.
[50] A. Lijoi, B. Nipoti, I. Prünster, Bayesian inference with dependent normalized completely random measures, Bernoulli 20 (3) (2014) 1260–1291.
[51] R.B. Nelsen, J.J.Q. Molina, J.A.R. Lallena, M.U. Flores, Best-possible bounds on sets of bivariate distribution functions, J. Multivar. Anal. 90 (2004) 348–358.
[52] R. Pelessoni, P. Vicig, I. Montes, E. Miranda, Imprecise copulas and bivariate stochastic orders, in: Proceedings of the EUROFUSE Workshop on Uncertainty and Imprecision Modelling in Decision Making, 2013, pp. 217–224.
[53] N. Alon, Combinatorial nullstellensatz, Comb. Probab. Comput. 8 (1999) 7–29.
[54] M. Marshall, Positive Polynomials and Sums of Squares, Math. Surv. Monogr., American Mathematical Society, 2008.
[55] B. de Finetti, Teoria delle Probabilità, Einaudi, Turin, 1970.