

Article

Systems of Precision: Coherent Probabilities on Pre-Dynkin Systems and Coherent Previsions on Linear Subspaces [†]

Rabanus Derr ^{1,*}  and Robert C. Williamson ² 
¹ Department of Computer Science, University of Tübingen, 72076 Tübingen, Germany

² Department of Computer Science, University of Tübingen, Germany and Tübingen AI Center, 72076 Tübingen, Germany; bob.williamson@uni-tuebingen.de

* Correspondence: rabanus.derr@uni-tuebingen.de

[†] This paper is an extended version of our paper published in Proceedings of the International Symposium on Imprecise Probability: Theories and Applications, “The Set Structure of Precision” .

Abstract: In the literature on imprecise probability, little attention is paid to the fact that imprecise probabilities are precise on a set of events. We call these sets *systems of precision*. We show that, under mild assumptions, the system of precision of a lower and upper probability form a so-called (pre-)Dynkin system. Interestingly, there are several settings, ranging from machine learning on partial data over frequentist probability theory to quantum probability theory and decision making under uncertainty, in which, a priori, the probabilities are only desired to be precise on a specific underlying set system. Here, (pre-)Dynkin systems have been adopted as systems of precision, too. We show that, under extendability conditions, those pre-Dynkin systems equipped with probabilities can be embedded into algebras of sets. Surprisingly, the extendability conditions elaborated in a strand of work in quantum probability are equivalent to coherence from the imprecise probability literature. On this basis, we spell out a lattice duality which relates systems of precision to credal sets of probabilities. We conclude the presentation with a generalization of the framework to expectation-type counterparts of imprecise probabilities. The analogue of pre-Dynkin systems turns out to be (sets of) linear subspaces in the space of bounded, real-valued functions. We introduce partial expectations, natural generalizations of probabilities defined on pre-Dynkin systems. Again, coherence and extendability are equivalent. A related but more general lattice duality preserves the relation between systems of precision and credal sets of probabilities.

Keywords: pre-Dynkin system; Dynkin system; coherence; extendability; quantum probability; intersectionality; prevision; imprecise probability; partial expectation



Citation: Derr, R.; Williamson, R.C. Systems of Precision: Coherent Probabilities on Pre-Dynkin Systems and Coherent Previsions on Linear Subspace. *Entropy* **2023**, *25*, 1283. <https://doi.org/10.3390/e25091283>

Academic Editors: Andrei Khrennikov and Karl Svozil

Received: 30 May 2023

Revised: 19 July 2023

Accepted: 25 July 2023

Published: 31 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

When posing problems in probability calculus, it should be required to indicate for which events the probabilities are assumed to exist.

—Andreï Kolmogorov [1] (p. 52)

1. Introduction

Scholarship in imprecise probability largely focuses on the imprecision of probabilities. However, imprecise probability models often lead to *precise* probabilistic statements on certain events or gambles, i.e., bounded, real-valued functions. In this work, we follow a hitherto untaken route investigating the *system of precision*, i.e., the set structure on which an imprecise probability is precise. (We elaborate the exact definition of imprecise probabilities and expectation used here in Sections 3 and 6). It turns out that (pre-)Dynkin systems (as shown in Appendix B, (pre-)Dynkin systems appear under plenty of names) describe the set of events with precise probabilities (cf. Section 3). This event structure is a neglected object in the literature on imprecise probability. In particular, it constitutes a parametrized choice somewhat “orthogonal” to the standard. Roughly stated, existing approaches to imprecise probability generalize the probability measure μ_σ in a classical probability space

$(\Omega, \mathcal{F}_\sigma, \mu_\sigma)$. Following Kolmogorov's classical setup, Ω is the base set, $\mathcal{F}_\sigma$ a σ -algebra and μ_σ a countably additive probability on $\mathcal{F}_\sigma$. Approaches to imprecise probability often do not even presuppose an underlying measure space (e.g., [2]). However, they are often linked to finitely additive measure spaces $(\Omega, \mathcal{F}, \mu)$, where μ is a finitely additive probability, and $\mathcal{F}$ is an algebra of sets (sometimes called a field). We start by generalizing $\mathcal{F}_\sigma$ from a σ -algebra to a pre-Dynkin system.

This suggestion is practically motivated: what do the following scenarios have in common?

- (a) A machine learning algorithm has access to a restricted subset of attributes. It cannot jointly query all attributes simultaneously. This is called “learning on partial, aggregated information” [3]. The reasons might be manifold: for privacy preservation, “not-missing-at-random” features, restricted data base access for acceleration or multi-measurement data sets.
- (b) Quantum physical quantities, e.g., location and impulse, are (statistically) incompatible [4].
- (c) A preference ordering on a set of acts gives rise to precise beliefs on a set of events, whereas this belief is not necessarily precise for intersections of such events [5,6].

In all of these scenarios, there does not exist a precise probability over all attributes and events. Or, there is no such precise probability accessible. Two attributes might each on their own exhibit a precise probabilistic description, while a joint precise probabilistic description does not exist. On a more fundamental level, no intersectability is provided. A precise probabilistic description of two events does *not* imply that the intersection of those events possesses a precise probability. The set system for the description of the events with precise probabilities which independently turned up in the various, previously mentioned fields of research is, again, the (pre-)Dynkin system.

The question of intersectability (or “intersectionality”) is of considerable interest in the social sciences where it is used as a label to describe the problem of the *joint* effect of various individual attributes on social outcomes [7–9]. That this notion of intersectionality has something to do with set systems is clear already from the fact that the Venn diagram pictured in Figure 1 is used as an illustration both for the Wikipedia articles on hypergraphs [10] (another name for a set system [11]) and intersectionality [12].

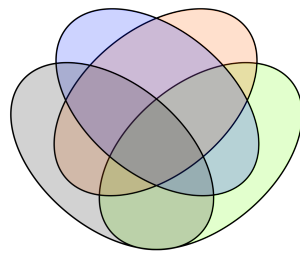


Figure 1. Exemplary illustration of Venn diagram for four sets (By Rupert Millard, CC BY-SA 3.0).

Needless to say, the concept as used in the social sciences is rich, complex, and somewhat vague, which is not necessarily held to be a weakness: “at least part of its success has been attributed to its vagueness” [13] (p. 260). Our interest is in under what circumstances precise probabilities can be ascribed to events; we speculate that such formal results may well contribute to a deeper empirical understanding of social intersectionality, without resorting to fuzzy logic [14] with its potential lack of operational definition as argued for by Cooke [15].

By rethinking the domain of probability measures, one might wonder about the origins of Kolmogorov's σ -algebra as *the* set system for events which possess probabilities. This links back to the old problem of measurability [16] (pp. 1–5). The measurability problem is the mathematical problem to assign a uniform measure to all subsets of a continuum. Giuseppe Vitali showed in 1905 that this problem is not solvable for countably additive measures [16] (p. 5) (from [17]). Hence, more restricted set systems such as the σ -algebra arose. Isaacs et al. [18] reconsidered this century-old discussion to argue for rationality of imprecise probabilities. We take their argument even further. Inside the borders of

mathematical measurability, the set of events which ought to be assigned probabilities is a modeling choice. Measurability is a modeling tool. We show that it is naturally parametrized by the set of (pre-)Dynkin systems.

All of the preceding considerations bring us to the main question of this paper: *What is the system of precision and how does it relate to an imprecise probability on “all” events?* We approach this question from three perspectives.

1. First, we show that, under mild assumptions, a pair of lower and upper probabilities assign precise probabilities, i.e., lower and upper probability coincide, to events which form a pre-Dynkin system or even a Dynkin system.
2. Second, we define probabilities on pre-Dynkin systems in accordance with the literature on quantum probability, in particular [19]. We argue that probabilities on pre-Dynkin systems, as well as their inner and outer extension, exhibit few desirable properties, e.g., subadditivity cannot be guaranteed. Hence, extendability, the ability to extend a probability from a pre-Dynkin system to a larger set structure, turns out to be crucial, as it implies coherence of the probability defined on the pre-Dynkin system. This observation links together the research from probabilities defined on weak set structures [6,19,20] to imprecise probabilities [2,21]. Furthermore, extendability guarantees the existence of a nicely behaving, so-called coherent extension. We finally show that the inner and outer extension of a probability defined on a pre-Dynkin system is always more pessimistic than its corresponding lower and upper coherent extension.
3. Last, we develop a duality theory between pre-Dynkin systems on a predefined base measure space and their respective credal sets of probabilities. The credal sets consist of all probabilities which coincide with the pre-defined measure on a pre-Dynkin system. A so-called Galois connection links together the containment structure on the set of set systems with the containment structure on the set of credal sets.

We conclude our presentation with a generalization to expectation-type counterparts of imprecise probabilities in Section 6. These are often called previsions, e.g., in [2]. Our main question thus generalizes to: *What is the system of precision and how does it relate to an imprecise expectation on “all” gambles?* In this case, by “system of precision” we mean the set of gambles on which a lower and upper expectation coincide.

1. First, we propose a generalization of a finitely additive probability defined on a pre-Dynkin system. More concretely, we define *partial expectations* which correspond to expectation functionals which are only defined on a set of linear subspaces of the space of all gambles. However, on those linear subspaces, they behave like “classical” (finitely additive) expectations.
2. Second, we show that under some properties, imprecise expectations are precise on a linear subspace of the linear space of gambles. (cf. Section 3)
3. Third, we present a natural generalization of extendability for partial expectations, which again turns out to be equivalent to coherence of the partial expectation.
4. Last, analogous to the lattice duality (A lattice is a poset with pairwise existing minimum and maximum. The duality is expressed via an antitone lattice isomorphism.) described in Section 5, we present a lattice duality for linear subsets of the space of gambles and credal sets which define coherent lower and upper previsions.

In summary, our work makes contributions in-between the research field of imprecise probabilities, probabilities defined on general set structures, and partially defined expectation functionals. Part of this work has been presented at the International Symposium on Imprecise Probabilities: Theories and Applications under the title “The Set Structure of Precision” [22]. The following version is a more complete and exhaustive presentation of this conference version. We included all omitted proofs of the conference version. We elaborated the content of Section 5. We added an entire section about the generalization to expectation-type counterparts of probabilities on pre-Dynkin systems (Section 6). We presented relations to different research areas in more detail. We shortly

discussed countably additive probabilities on Dynkin systems in the appendix, as we put emphasis on finitely additive probabilities in the main text. Before we begin the structural investigation of pre-Dynkin systems, we first introduce the used notation and fix the mathematical framework.

Notation and Technical Details

As we deal with a lot of sets, sets of sets, and rarely even sets of sets of sets in this paper, we agree on the following notation: sets are written with capital Latin or Greek letter, e.g., A or Ω . Sets of sets are denoted $\mathcal{A}$. Sets of sets of sets obtain the notation $\mathfrak{A}$. As usual, $\mathbb{R}$ is reserved for the set of real numbers, $\mathbb{N}$ for the natural numbers. The power set of a set A is written as 2^A .

In the course of this work, we require the notions of σ -algebras and algebras (of sets). An algebra is a subset of 2^Ω which contains the empty set and is closed under complement and finite union. A σ -algebra is an algebra which is closed under countable union [23] (Definition 1.1). (Our notion of an algebra should not be confused with the notion of an algebra over a field.) (Probability) measures are denoted by lowercase Greek letters, e.g., μ , ν and ψ , except for σ . Generally, we use “ σ ” to emphasize the countable nature of a mathematical object. This becomes clear when we define Dynkin systems (Definition 1). Other functions are denoted by lowercase Latin letters, e.g., f and g .

Regarding the technical setting, we roughly follow the setup of [2] (§3.6 and Appendix D). For a summary, see Table 1.

Table 1. Summary of important notations used.

$\Omega, 2^\Omega$	Base set and its power set
$[n]$	Set $\{1, \dots, n\}$
$\mathcal{D}$	Pre-Dynkin system on Ω (Definition 1)
$\mathcal{D}_\sigma$	Dynkin system on Ω (Definition 1)
$D(\mathcal{A})$	Pre-Dynkin hull of a set system $\mathcal{A} \subseteq 2^\Omega$ (Definition 1)
μ	Finitely additive probability defined on $\mathcal{D}$ (Definition 3)
μ_*, μ^*	Inner respectively outer extension (Proposition 1)
$\mu_{\mathcal{D}}, \bar{\mu}_{\mathcal{D}}$	Lower respectively upper coherent extension (Corollary 1)
$M(\mu, \mathcal{D})$	Credal set of μ on $\mathcal{D}$ (Corollary 1)
ν	Finitely additive probability defined on 2^Ω
ψ	Fixed, finitely additive probability defined on 2^Ω
χ_A	Indicator function of the set $A \subset \Omega$
Δ	Set of finitely additive probability measures on 2^Ω , set of linear previsions
$m_\psi: 2^{2^\Omega} \rightarrow 2^\Delta$	Credal set function (Definition 6)
$m_\psi^\circ: 2^\Delta \rightarrow 2^{2^\Omega}$	Dual credal set function (Definition 7)
$\overline{\text{co}}$	Convex, Weak* Closure
$B(\Omega)$	Set of real-valued, bounded functions on Ω
$\text{ba}(\Omega)$	Set of bounded, signed, finitely additive measures on 2^Ω
E	Partial Expectation (Definition 8)
$S(\Omega, \mathcal{A})$	Linear space of simple gambles on the set system $\mathcal{A}$
$B(\Omega, \mathcal{F}_\sigma)$	Linear space of bounded, $\mathcal{F}_\sigma$ -measurable functions
$\underline{\mathbb{E}}$	Coherent lower prevision (Definition 9)
$\overline{\mathbb{E}}$	Coherent upper prevision (Definition 9)
ν	Linear prevision defined on $B(\Omega)$ (equivalent to ν above)
ψ	Fixed, linear prevision defined on $B(\Omega)$ (equivalent to ψ above)
$m_\psi: 2^{B(\Omega)} \rightarrow 2^\Delta$	Generalized credal set function (Definition 11)
$m_\psi^\circ: 2^\Delta \rightarrow 2^{B(\Omega)}$	Generalized dual credal set function (Definition 12)

Let Ω be an arbitrary set. In several examples $\Omega = [n]$, where $[n]$ denotes the set $\{1, \dots, n\}$. The set $B(\Omega)$ is defined as the set of all real-valued, bounded functions on Ω .

We call those functions *gambles*. For instance, χ_A , the indicator function of $A \subseteq \Omega$, is in $B(\Omega)$. The supremum norm, $\|f\|_{\sup} := \sup_{\omega \in \Omega} |f(\omega)|$ makes $B(\Omega)$ a topological linear vector space [24]. With $\text{ba}(\Omega)$, we denote the set of all bounded, signed, finitely additive measures on 2^Ω . In fact, $\text{ba}(\Omega)$ is the topological dual space of $B(\Omega)$. So, in particular, every continuous linear functional $\nu \in B(\Omega)^*$ can be identified with a bounded, signed, finitely additive measure [24]. (As the linear functional is defined on a normed space, continuity and boundedness are equivalent.) For this reason, we use, with minor abuse of notation, the same notation for bounded, signed, finitely additive measures and for continuous linear functionals in the dual space of $B(\Omega)$, i.e., we write $\nu(f) = \int f d\nu$. The dual space $\text{ba}(\Omega)$ gets equipped with the weak* topology, i.e., the weakest topology which makes all evaluation functionals of the form $f^* \in \text{ba}(\Omega)^*$ such that $f^*(\nu) := \int f d\nu$ for some $f \in B(\Omega)$ continuous (for more details see, e.g., [25] (p. 758)). With $\Delta \subseteq \text{ba}(\Omega)$, we denote the convex, weak*-closed subset of finitely additive probability measures. The set Δ plays a major role in Walley's theory of previsions, as the measures in Δ are in one-to-one correspondence to his linear previsions [2] (Theorem 3.2.2). The operator $\overline{\text{co}}$ is the convex, weak* closure on the space $\text{ba}(\Omega)$. We further introduce the following two notations: let $\mathcal{F} \subseteq 2^\Omega$ be an algebra. Then, $S(\Omega, \mathcal{F}) \subseteq B(\Omega)$ denotes the linear subspace of simple functions on $\mathcal{F}$, i.e., scaled and added indicator functions of a finite number of disjoint sets (cf. [26] (Definition 4.2.12)). Let $\mathcal{F}_\sigma \subseteq 2^\Omega$ be a σ -algebra. Then, $B(\Omega, \mathcal{F}_\sigma) \subseteq B(\Omega)$ denotes the linear subspace of all bounded, real-valued, $\mathcal{F}_\sigma$ -measurable functions. Equipped with these notions and tools we are ready for a first preliminary question.

2. What Is a (Pre-)Dynkin System?

In this work, the main objects under consideration are pre-Dynkin systems and Dynkin systems. A (pre-)Dynkin system is a set system on Ω . It contains the empty set, is closed under complement and (countable) disjoint union. More formally:

Definition 1 ((Pre-)Dynkin system). We say $\mathcal{D} \subseteq 2^\Omega$ is a pre-Dynkin system on some set Ω if and only if all of the following conditions hold:

- (a) $\emptyset \in \mathcal{D}$,
- (b) $D \in \mathcal{D}$ implies $D^c := \Omega \setminus D \in \mathcal{D}$
- (c) $C, D \in \mathcal{D}$ with $C \cap D = \emptyset$ implies $C \cup D \in \mathcal{D}$.

We call $\mathcal{D}_\sigma \subseteq 2^\Omega$ a Dynkin system if and only if the conditions (a), (b) and

- (c') let $\{D_i\}_{i \in \mathbb{N}} \subseteq \mathcal{D}_\sigma$, if for all $i, j \in \mathbb{N}$ with $i \neq j$ it holds $D_i \cap D_j = \emptyset$ then $\bigcup_{i \in \mathbb{N}} D_i \in \mathcal{D}_\sigma$, are fulfilled.

Observe that every Dynkin system is a pre-Dynkin system. We will denote pre-Dynkin systems by the use of $\mathcal{D}$, in contrast to $\mathcal{D}_\sigma$ for Dynkin systems. This should not be confused with $D(\mathcal{A})$ for $\mathcal{A} \subseteq 2^\Omega$, which is the intersection of all pre-Dynkin systems which contain $\mathcal{A}$, i.e., the smallest pre-Dynkin system containing $\mathcal{A}$. (For $\mathcal{A} = \emptyset$ we define $D(\mathcal{A}) = \{\emptyset, \Omega\}$.) In other words, $D(\mathcal{A})$ is the *pre-Dynkin hull* generated by $\mathcal{A}$. The following short lemma will be helpful in later proofs.

Lemma 1 (Closedness under Set Difference). Let $\mathcal{D} \subseteq 2^\Omega$ be a pre-Dynkin system, if $A, B \in \mathcal{D}$ and $A \subseteq B$, then $B \setminus A \in \mathcal{D}$.

Proof. Let $A, B \in \mathcal{D}$ and $A \subseteq B$. Then, $(B \setminus A)^c = B^c \cup A$. Since $\mathcal{D}$ is closed under complement and disjoint union $B^c \cup A \in \mathcal{D}$. But, then again, the complement, $B \setminus A$, is in $\mathcal{D}$. $\square$

In classical probability theory, Dynkin systems appear as a technical object required for the measure-theoretic link between cumulative distribution functions and probability measures (cf. [23] (Proof of Lemma 1.6)). In particular, every σ -algebra, the well-known domain of probability measures, is a Dynkin system. Thus, all statements within this work

are generalizations of classical probability theoretical results. We give a short example of a pre-Dynkin system, which is not an algebra in the following. This example gets reused to illustrate forthcoming statements.

Example 1. *The smallest pre-Dynkin system, up to isomorphisms, which is not an algebra can be defined on $\Omega_4 := \{1, 2, 3, 4\}$. It is given by $\mathcal{D}_4 := \{\emptyset, 12, 34, 13, 24, \Omega_4\}$, where we write 12 as a shorthand for $\{1, 2\}$.*

Pre-Dynkin and Dynkin systems naturally arise in probability theory. For instance, the set of all subsets $A \subseteq \mathbb{N}$, such that the natural density $\mu(A) = \lim_{n \rightarrow \infty} \frac{|A \cap [n]|}{n}$ exists is a pre-Dynkin system $\mathcal{D}_{\mathbb{N}}$ but not an algebra [20]. It is sometimes called the density logic [27] and constitutes the foundation of von Mises' century-old frequential theory of probability [28] (refined and summarized in [29]). Intriguingly, this was used as an example by Kolmogorov [1] of a measure defined on a restricted set system for which it is desired to extend the measure to the power set $2^{\mathbb{N}}$ (cf. § 4); see the discussion in [30] (pp. 11–14), who observed (page 14) that “the main problem is *non-uniqueness* of an extension” and that such extended measures are impossible to verify from observed frequencies, because the relative frequencies do not converge for events in $2^{\mathbb{N}} \setminus \mathcal{D}_{\mathbb{N}}$. The non-uniqueness is naturally handled in the present paper by working with lower and upper previsions (or lower and upper probabilities) (cf. [31]).

Another class of Dynkin systems occurs in so-called marginal scenarios [32]. Marginal scenarios are settings in which marginal probability distributions for a subset of a set of random variables are given, but not the entire joint distribution. This restricted “joint measurability” of the involved random variables can be expressed via Dynkin systems [33] (Example 4.2), [34].

Pre-Dynkin systems are so helpful because they structurally align with finitely additive probability measures. The same statement holds for Dynkin systems and countably additive probabilities. If we know the probability of an event, then we know the probability of the complement; i.e., the event does not happen. If we know the probability of several events which are disjoint, then we know the probability of the union, which is just the sum. Probabilities following their standard definition go hand in hand with Dynkin systems. We see this observation manifested in many following statements.

Remarkably, (pre-)Dynkin systems appeared under a variety of names (cf. Appendix B). Fundamental to all its regular, independent occurrences in many research areas is the need for a set structure which does not allow for arbitrary intersections.

2.1. Compatibility

(Pre-)Dynkin systems are not necessarily closed under intersections. However, when the intersection of two sets (events) is contained in the (pre-)Dynkin system, we call the two events *compatible*.

Definition 2 (Compatibility). *Let A, B be elements in a pre-Dynkin system $\mathcal{D}$, then A and B are compatible if and only if $A \cap B \in \mathcal{D}$.*

This definition follows the definitions given in, e.g., [19,33,35]. (It should not be confused with the very similar, and sometimes equivalent, notion of commutativity in logical structures [36] (Definition 14) (cf. Appendix E).) Compatibility in pre-Dynkin systems is a symmetric relation, but it is not necessarily transitive. Furthermore, it is complement inherited, i.e., if A, B are compatible in a pre-Dynkin system, then so are A, B^c [4] (Lemma 3.6). Lastly, compatibility, even though expressed as intersectability, i.e., “closed under intersection”, can be equivalently expressed as unifiability, i.e., “closed under union”.

Lemma 2 (Cup gives Cap gives Cup). *Let $\mathcal{D}$ be a pre-Dynkin system and $A, B \in \mathcal{D}$. Then*

$$A \cap B \in \mathcal{D} \Leftrightarrow A \cup B \in \mathcal{D}$$

Proof. Using Lemma 1 for pre-Dynkin systems, we can quickly see that the following two decompositions give the desired equivalence:

For the “ $\Rightarrow$ ”-direction: $A \cup B = (A \setminus (A \cap B)) \cup B$. The fact $A, A \cap B, B \in \mathcal{D}$ implies $(A \setminus (A \cap B)) \cup B \in \mathcal{D}$.

For the “ $\Leftarrow$ ”-direction: $A \cap B = A \setminus ((A \cup B) \setminus B)$. The fact $A, A \cup B, B \in \mathcal{D}$ implies $A \setminus ((A \cup B) \setminus B) \in \mathcal{D}$. (A related result for Dynkin systems is given in [19] (5.1).) $\square$

Example 2. *We reconsider the set Ω_4 and pre-Dynkin system $\mathcal{D}_4$ from Example 1. The elements 12 and 34 are intersectable $12 \cap 34 = \emptyset \in \mathcal{D}_4$ and unifiable $12 \cup 34 = \Omega_4 \in \mathcal{D}_4$. The elements 12 and 13 are not intersectable $12 \cap 13 = 1 \notin \mathcal{D}_4$ and not unifiable $12 \cup 13 = 123 \notin \mathcal{D}_4$.*

The term “compatibility” underlines that closedness under intersection gets loaded with further meaning in the context of probability theory. As we define in the next section, $\mathcal{D}$ is the set of “measurable” events, i.e., events which get assigned a probability. Hence, two events, A, B , are called compatible if and only if a precise joint probabilistic description, i.e., a precise probability of $A \cap B$, exists. For a more thorough discussion of the nature of compatibility (and its cousin commutativity) we point to the literature on quantum probability, e.g., [37] (Definition 3.12), or [38].

Compatibility is not only a property of elements in pre-Dynkin systems. One can take compatibility as a primary notion; i.e., one requires the statements of Lemma 2 and [4] (Lemma 3.6) to hold. Then, a set structure which contains the empty set and the entire base set and is equipped with this notion of compatibility is a pre-Dynkin system [30] (Definition 5.1). (It is called semi-algebra in [30] (Definition 5.1).)

Interestingly, the assumption of arbitrary compatibility is fundamental to most parts of probability theory. σ -algebras, the domain of probability measures, are exactly those Dynkin systems in which all events are compatible with all others [35] (Theorem 2.1). Algebras are exactly those pre-Dynkin systems in which all events are compatible with all others. Surprisingly, it turns out that, as well, all pre-Dynkin systems can be dissected into such “blocks” of full compatibility. Every pre-Dynkin system consists of a set of maximal algebras, which we call *blocks*. In particular, maximality here stands for the following: there is no algebra contained in $\mathcal{D}$ such that some $\mathcal{A}_i$ is a strict sub-algebra of this algebra. Similar and related results can be found in [39–42].

Theorem 1 (Pre-Dynkin Systems Are Made Out of Algebras). *Let $\mathcal{D}$ be a pre-Dynkin system on Ω . Then, there is a unique family of maximal algebras $\{\mathcal{A}_i\}_{i \in I}$ such that $\mathcal{D} = \bigcup_{i \in I} \mathcal{A}_i$. We call these algebras the blocks of $\mathcal{D}$.*

Proof. For the proof, we require the definition of a compatible subset of $\mathcal{D}$. A subset $\mathcal{A} \subseteq \mathcal{D}$ is compatible if all elements are completely compatible, i.e., any finite intersection of elements in $\mathcal{A}$ is contained in $\mathcal{D}$. This is indeed a stronger requirement than pairwise compatibility (cf. Definition 2). Certainly, every subset $\mathcal{A} \subseteq \mathcal{D}$ is compatible if and only if every finite subset of $\mathcal{A}$ is compatible. Hence, compatibility is a property of so-called finite character [25] (Definition 3.46). Then, Tuckey’s lemma (e.g., [25] (Theorem 6.20.AC5)) guarantees that any compatible subset of $\mathcal{D}$ is contained in a maximal compatible subset. Since every element $D \in \mathcal{D}$ is in at least one compatible subset, e.g. $\{\emptyset, D, D^c, \Omega\} \subseteq \mathcal{D}$, the (unique) set of maximal compatible subsets $\{\mathcal{A}_i\}_{i \in I}$ covers the entire pre-Dynkin system $\mathcal{D}$. It remains to show that the maximal compatible subsets are algebras. Consider a maximal compatible subset $\mathcal{A}_i$. First, $\emptyset \in \mathcal{A}_i$ as $\emptyset$ is compatible to all sets in 2^Ω . Second, $\mathcal{A}_i$ is closed under finite intersection, otherwise there would exist a finite combination of elements $A_1, \dots, A_n \subseteq \mathcal{A}_i$ such that $A_n := \bigcap_{j=1}^n A_j \in \mathcal{D}$, but $A_n \notin \mathcal{A}_i$. Then, one can easily see that $\mathcal{A}_i \cup \{A_n\}$ would be a compatible subset which strictly contained $\mathcal{A}_i$. This is impossible,

since $\mathcal{A}_i$ is maximal. Finally, $\mathcal{A}_i$ is closed under complement. Consider $A \in \mathcal{A}_i$, we show that $\mathcal{A}_i \cup \{A^c\}$ is again a compatible subset. Let $A_1, \dots, A_n \subseteq \mathcal{A}_i$ be an arbitrary finite collection of subsets, then $A \cap \bigcap_{j=1}^n A_j \in \mathcal{D}$, hence $A^c \cap \bigcap_{j=1}^n A_j \in \mathcal{D}$ [4] (Lemma 3.6). By maximality of $\mathcal{A}_i$, we then know $A^c \in \mathcal{A}_i$. $\square$

Example 3. The pre-Dynkin system $\mathcal{D}_4$ of Example 1 consists of the algebras $\{\emptyset, 12, 34, \Omega_4\}$ and $\{\emptyset, 13, 24, \Omega_4\}$.

Theorem 1 simplifies several follow-up observations. Instead of pre-Dynkin systems, we can equivalently consider a set of algebras. However, not every union of algebras is a pre-Dynkin system. If these algebras form a compatibility structure, i.e., a set of maximal π -systems, which are non-empty set systems closed under finite intersections, then their union is a pre-Dynkin system (Definition A2 and Theorem A1 in Appendix A). Analogous results for Dynkin systems and σ -algebras exist and are given in Appendix D.1. In summary, pre-Dynkin systems are set structures which do not allow for arbitrary intersections but can be split into maximal intersectable subsets, their *blocks*.

2.2. Probabilities on Pre-Dynkin Systems

We require a notion of probability on a pre-Dynkin system to elaborate the relationship of imprecise probability and the system of precision in the following. Probabilities are classically defined on σ -algebras. We generalize this definition as e.g., stated in [23] (p. 18f) to pre-Dynkin systems.

Definition 3 (Probability Measure on a Pre-Dynkin System). Let $\mathcal{D}$ be a pre-Dynkin system. We call a function $\mu: \mathcal{D} \rightarrow [0, 1]$ a finitely additive probability measure on $\mathcal{D}$ if and only if it fulfills the following two conditions:

- (a) Normalization: $\mu(\emptyset) = 0$ and $\mu(\Omega) = 1$.
- (b) Additivity: let $A, B \in \mathcal{D}$ and $A \cap B = \emptyset$. Then, $\mu(A \cup B) = \mu(A) + \mu(B)$.

If condition (b) can be extended to countable subsets of $\mathcal{D}$, then we call μ countably additive probability measure on $\mathcal{D}$: let $I \subseteq \mathbb{N}$ and $\{A_i\}_{i \in I}$ such that $A_i \in \mathcal{D}$ for all $i \in I$ and $A_i \cap A_j = \emptyset$ for $i \neq j$, $i, j \in I$ and $\bigcup_{i \in I} A_i \in \mathcal{D}$. Then $\mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i)$.

For the sake of readability, we use “probability” and “probability measure” exchangeably. Probabilities on pre-Dynkin systems are monotone, i.e., for $A, B \in \mathcal{D}$, if $A \subseteq B$, then $\mu(A) \leq \mu(B)$. This can be seen when applying Lemma 1 and Definition 3. But, in contrast to a probability defined on a σ -algebra, a probability on a pre-Dynkin system is not necessarily modular, i.e., for $A, B \in \mathcal{D}$, $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$ [43] (p. 16). (It is, however, possible to define modular probabilities on pre-Dynkin systems. This leads to a fixed parametrization of probability functions already on simple examples [44] (p. 125).) It is that sophisticated interplay of set structure and probability function which leads us through this paper. In particular, why should we consider pre-Dynkin systems?

3. Imprecise Probabilities Are Precise on a Pre-Dynkin System

As we now demonstrate, pre-Dynkin systems are, under mild assumptions, the systems of precision. To make this formal, we solely require a normed, conjugate pair of lower and upper probability which fulfill super (resp. sub)-additivity and possibly a continuity assumption.

Theorem 2 (Imprecise Probability Induces a (Pre-)Dynkin System). Let $\ell: 2^\Omega \rightarrow [0, 1]$ and $u: 2^\Omega \rightarrow [0, 1]$ be two set functions, for which all the following properties hold:

- (a) Normalization: $u(\emptyset) = \ell(\emptyset) = 0$.
- (b) Conjugacy: $u(A) = 1 - \ell(A^c)$ for $A, A^c \in 2^\Omega$.
- (c) Subadditivity of u : for $A, B \in 2^\Omega$ such that $A \cap B = \emptyset$, then $u(A \cup B) \leq u(A) + u(B)$.
- (d) Superadditivity of ℓ : for $A, B \in 2^\Omega$ such that $A \cap B = \emptyset$, then $\ell(A \cup B) \geq \ell(A) + \ell(B)$.

Then u and ℓ define a finitely additive probability measure $\mu := u|_{\mathcal{D}} = \ell|_{\mathcal{D}}$ on the pre-Dynkin system $\mathcal{D} := \{A \in 2^\Omega : \ell(A) = u(A)\} \subseteq 2^\Omega$. If either u fulfills

(e) Continuity from below: for $A_n \in 2^\Omega$ with $A_n \subseteq A_{n+1}$ such that $\bigcup_{n=1}^\infty A_n = A \in 2^\Omega$, then $\lim_{n \rightarrow \infty} u(A_n) = u(A)$,

or ℓ fulfills

(e') Continuity from above: for $A_n \in 2^\Omega$ with $A_{n+1} \subseteq A_n$ such that $\bigcap_{n=1}^\infty A_n = A \in 2^\Omega$, then $\lim_{n \rightarrow \infty} \ell(A_n) = \ell(A)$,

then u and ℓ define a countably additive probability measure $\mu_\sigma := u|_{\mathcal{D}_\sigma} = \ell|_{\mathcal{D}_\sigma}$ on the Dynkin system $\mathcal{D}_\sigma := \{A \in 2^\Omega : \ell(A) = u(A)\} \subseteq 2^\Omega$.

Proof. We start proving the first part of the theorem. Let

$$\mathcal{D} := \{A \in 2^\Omega : \ell(A) = u(A)\}. \quad (1)$$

We show that $\mathcal{D}$ is a pre-Dynkin system. First, $\emptyset \in \mathcal{D}$ by assumption (a). Second, let $D \in \mathcal{D}$. Then, $u(D^c) = 1 - \ell(D) = 1 - u(D) = \ell(D^c)$ by the conjugacy relation. Third, let $\{A_i\}_{i \in I} \subseteq \mathcal{D}$ for finite $I \subseteq \mathbb{N}$ such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then

$$\begin{aligned} \sum_{i \in I} \ell(A_i) &\stackrel{(d)}{\leq} \ell\left(\bigcup_{i \in I} A_i\right) \\ &\stackrel{(\star)}{\leq} u\left(\bigcup_{i \in I} A_i\right) \\ &\stackrel{(c)}{\leq} \sum_{i \in I} u(A_i) \\ &\stackrel{\text{Equation (1)}}{=} \sum_{i \in I} \ell(A_i). \end{aligned}$$

For $(\star)$, observe that $\ell(A) \leq u(A)$ for all $A \in 2^\Omega$, since

$$\ell(A) + \ell(A^c) \leq \ell(A \cup A^c) = 1 = u(A \cup A^c) \leq u(A) + u(A^c),$$

and thus,

$$\begin{aligned} \ell(A) + \ell(A^c) &\leq u(A) + u(A^c) \\ \Leftrightarrow \ell(A) + 1 - u(A) &\leq u(A) + 1 - \ell(A) \\ \Leftrightarrow \ell(A) &\leq u(A). \end{aligned}$$

Concluding, we define $\mu := \ell|_{\mathcal{D}} = u|_{\mathcal{D}}$ for which it is trivial to show that it is a finitely additive probability on $\mathcal{D}$.

For the second part, we first notice that continuity from below and from above are equivalent for conjugate set functions on set systems which are closed under complement [43] (Proposition 2.3). Next, we show that subadditivity of u and continuity from below (of u) imply σ -subadditivity of u : for $\{A_i\}_{i \in I} \subseteq 2^\Omega$ such that $I \subseteq \mathbb{N}$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$ with $i, j \in I$ then $u(\bigcup_{i \in I} A_i) \leq \sum_{i \in I} u(A_i)$. In the case that I is finite, subadditivity of u is provided by assumption. For infinite I , we can construct an increasing sequence of sets, namely $B_j = \bigcup_{1 \leq i \leq j} A_i$, so that $B_j \subseteq B_{j+1}$. Furthermore, $\bigcup_{j=1}^\infty B_j = \bigcup_{i \in I} A_i$. Thus,

$$\begin{aligned} u\left(\bigcup_{i \in I} A_i\right) &= u\left(\bigcup_{j=1}^\infty B_j\right) \\ &\stackrel{(e)}{=} \lim_{j \rightarrow \infty} u(B_j) \end{aligned}$$

$$\begin{aligned}
&= \lim_{j \rightarrow \infty} u \left(\bigcup_{1 \leq i \leq j} A_i \right) \\
&\stackrel{(d)}{\leq} \lim_{j \rightarrow \infty} \sum_{1 \leq i \leq j} u(A_i) \\
&= \sum_{i \in I} u(A_i).
\end{aligned}$$

The same argument holds analogously for superadditivity and continuity from above of ℓ which is implied by continuity from below and the conjugacy relationship [43] (Proposition 2.3). In summary, the proof of the first part can then be applied again, now without the restriction that $I \subseteq \mathbb{N}$ is finite. Instead, it potentially is countable. $\square$

Example 4. Remember, $\Omega_4 = \{1, 2, 3, 4\}$. We define $\ell: 2^{\Omega_4} \rightarrow [0, 1]$ by $\ell(1) = 0$, $\ell(2) = 0.3$, $\ell(3) = 0$, $\ell(4) = 0.3$, $\ell(12) = 0.5$, $\ell(34) = 0.5$, $\ell(13) = 0.2$, $\ell(24) = 0.8$, $\ell(14) = 0.3$, $\ell(23) = 0.3$, $\ell(123) = 0.5$, $\ell(124) = 0.8$, $\ell(134) = 0.5$, $\ell(234) = 0.8$, $\ell(\Omega_4) = 1$ and $u: 2^{\Omega_4} \rightarrow [0, 1]$ by $u(A) = 1 - \ell(A^c)$. It is easy to show that ℓ and u fulfill the assumptions (a), (b), (c) and (d) in Theorem 2. The imprecise probabilities u and ℓ coincide on $\{\emptyset, 12, 34, 13, 24, \Omega_4\}$, the pre-Dynkin system described in Example 1. The example is illustrated in Figure 2. In this figure $\ell = \underline{\mu}_{\mathcal{D}_4}$ and $u = \bar{\mu}_{\mathcal{D}_4}$.

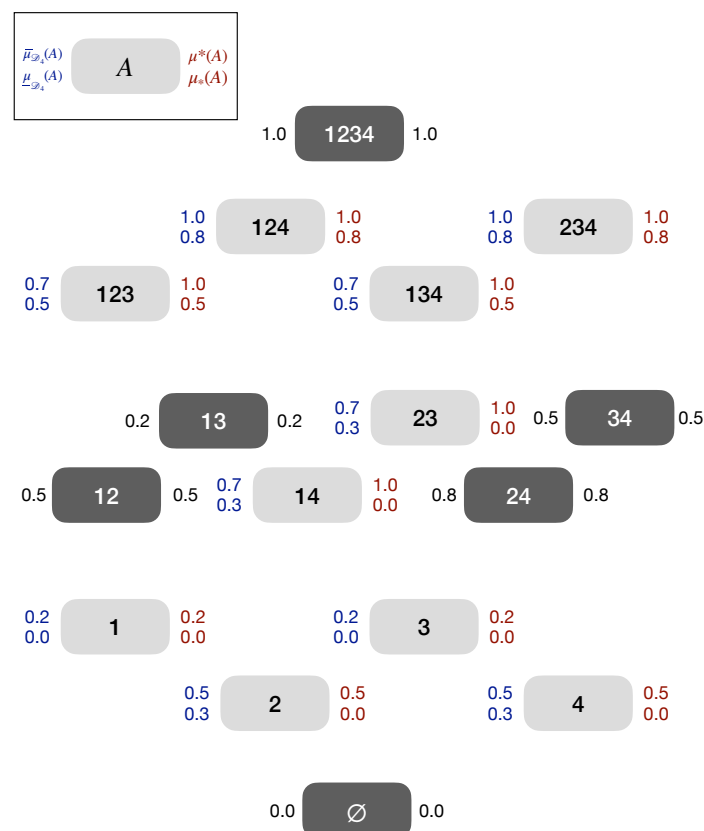


Figure 2. Illustration of the running example. The dark elements are contained in the pre-Dynkin system $\mathcal{D}$ on $\Omega = \{1, 2, 3, 4\}$. The lower and upper coherent extension, respectively, the inner and outer extension, are denoted at the sides of the elements in the set system as shown in the example in the left upper corner. Elements in $\mathcal{D}$ possess a precise probability.

In summary, imprecise probabilities are, under mild assumptions, precise on a pre-Dynkin system or even a Dynkin system. (The events in the system of precisions are called *unambiguous events* in the decision theory literature [5].) This, importantly, is also the case if the system of precision is strictly larger than the trivial pre-Dynkin systems $\{\emptyset, \Omega\}$. Exemplarily, a pair of conjugate, coherent lower and upper probability (e.g., [2] (§2.7.4)) fulfills the conditions (a)–(d). However, in several cases (e.g., distorted probability distributions), imprecise probabilities are just precise on the *system of certainty*, i.e., the events which possess 0 or 1 probability (Proposition A1). Concluding, the system of precision is a pre-Dynkin system $\mathcal{D} \subseteq 2^\Omega$. What if we first define precise, finitely additive probabilities on a pre-Dynkin system, i.e., we fix a system of precision? We can then ask for “imprecise probabilities” deduced from this probability which are defined on a larger set structure, e.g., an algebra in which the pre-Dynkin system is contained.

4. Extending Probabilities on Pre-Dynkin Systems

Precise probabilities on pre-Dynkin system naturally arise in many, distinct, applied scenarios as we argued in the Introduction (Section 1). However, we acknowledge that the definition of probabilities on pre-Dynkin systems is mathematically cumbersome. The possibilities to prove standard theorems is very limited as the approaches by Gudder [4,35], Gudder and Zerbe [45] demonstrate. However, if we consider a probability defined on a pre-Dynkin system as an imprecise probability on a larger set system with a fixed system of precision, we possibly obtain a richer, mathematical toolkit to work with. In this case the larger set system preferably is an algebra in which the pre-Dynkin system is contained. It remains to clarify how we construct the imprecise probability from the precise probability on the pre-Dynkin system.

4.1. Inner and Outer Extension

A simple but, as we show, unsatisfying solution is the use of an inner and outer measure extension. It does not rely on imposing any conditions on the probability defined on the pre-Dynkin system. We pay for this generality with the few properties that we can derive for the obtained extension.

Proposition 1 (Inner and Outer Extension [6] (Lemma 2.2)). *Let $\mathcal{D}$ be a pre-Dynkin system on Ω and μ a finitely additive probability measure on $\mathcal{D}$. The inner probability measure*

$$\mu_*(A) := \sup\{\mu(B) : A \supseteq B \in \mathcal{D}\}, \forall A \in 2^\Omega,$$

and outer probability measure

$$\mu^*(A) := \inf\{\mu(B) : A \subseteq B \in \mathcal{D}\}, \forall A \in 2^\Omega,$$

define $\mu_, \mu^* : 2^\Omega \rightarrow [0, 1]$, i.e., all of the following conditions are fulfilled:*

- (a) *Normalization:* $\mu^*(\emptyset) = 0, \mu_*(\Omega) = 1$.
- (b) *Conjugacy:* $\mu^*(A) = 1 - \mu_*(A^c), \forall A \in 2^\Omega$.
- (c) *Monotonicity:* for $A, B \in 2^\Omega$, if $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$.

Furthermore, μ_ is superadditive, for $A, B \in 2^\Omega$ if $A \cap B = \emptyset$, then $\mu_*(A \cup B) \geq \mu_*(A) + \mu_*(B)$. But μ^* is not generally subadditive.*

Example 5. For $\mathcal{D}_4$, as in Example 4, let $\mu : \mathcal{D}_4 \rightarrow [0, 1]$ be defined as $\mu(\emptyset) = 0, \mu(12) = 0.5, \mu(34) = 0.5, \mu(13) = 0.2, \mu(24) = 0.8, \mu(\Omega) = 1$. The inner and outer extension of μ on $\mathcal{D}_4$ is $\mu_*(\emptyset) = 0, \mu_*(1) = \mu_*(2) = \mu_*(3) = \mu_*(4) = 0, \mu_*(12) = 0.5, \mu_*(34) = 0.5, \mu_*(13) = 0.2, \mu_*(24) = 0.8, \mu_*(14) = 0, \mu_*(23) = 0, \mu_*(123) = 0.5, \mu_*(124) = 0.8, \mu_*(134) = 0.5, \mu_*(234) = 0.8, \mu_*(\Omega_4) = 1$ and $\mu^* = 1 - \mu_*$. The inner and outer extension are not coherent (Definition 5). In particular, the outer extension is not subadditive: $\mu^*(14) = 1 - \mu_*(23) = 1 > 0.2 + 0.5 = (1 - \mu_*(234)) + (1 - \mu_*(123)) = \mu^*(1) + \mu^*(4)$. The example is illustrated in Figure 2.

In conclusion, the inner and outer extension provides an imprecise probability, which is not necessarily coherent (cf. Definition 5) and it does not fulfill the conditions required for Theorem 2 to post hoc guarantee that the set of precision is a pre-Dynkin system. We remark that there exist normalized, conjugate, monotone superadditive but not subadditive pairs of probabilities, hence possibly inner and outer probabilities as defined here, whose system of precision is not a pre-Dynkin system (see Example 6). For this reason we now explore another, more powerful extension method.

Example 6. Let $\Omega_3 := \{1, 2, 3\}$. The probability pair defined by $\underline{\nu}(\emptyset) = \bar{\nu}(\emptyset) = 0$, $\underline{\nu}(1) = \bar{\nu}(1) = 0.2$, $\underline{\nu}(2) = \bar{\nu}(2) = 0.2$ and $\underline{\nu}(3) = 0, \bar{\nu}(3) = 0.6$ is precise on $\{\emptyset, 1, 2, 23, 13, \Omega_3\}$, which is obviously not a pre-Dynkin system.

4.2. Extendability and Its Equivalence to Coherence

In the following, we try to entirely embed pre-Dynkin systems equipped with a probability into larger algebras. Then, we extend the probability defined on the pre-Dynkin system in all possible ways to probabilities on the algebra. It turns out that this embedding is only possible under certain conditions on the probability defined on the pre-Dynkin system. We call this condition *extendability*. For the sake of generality, we focus on the extension of finitely additive probabilities from pre-Dynkin systems to algebras here. We treat countably additive probabilities, Dynkin systems and σ -algebras in Appendix D. In addition, all results until Section 4.3 can be formulated in more general terms for non-structured set systems. For the sake of simplicity, we remain within the setting of probabilities defined on pre-Dynkin systems in this work.

Extendability is the property that a probability measure defined on a pre-Dynkin system can be extended to a probability measure on an algebra containing the pre-Dynkin system. Formally:

Definition 4 (Extendability). Let $\mathcal{D}$ be a pre-Dynkin system on Ω . We call a finitely additive probability measure μ on $\mathcal{D}$ extendable to 2^Ω if and only if there is a finitely additive probability measure $\nu: 2^\Omega \rightarrow [0, 1]$ such that $\nu|_{\mathcal{D}} = \mu$.

We defined extendability with respect to the power set 2^Ω . In fact, any relativization to an arbitrary sub-algebra of 2^Ω is equivalent. A finitely additive probability defined on $\mathcal{D}$ is extendable to any sub-algebra of 2^Ω which contains $\mathcal{D}$ if and only if it is extendable to 2^Ω [26] (Theorem 3.4.4).

The definition is non-vacuous [33,46]. For instance, a probability measure on a pre-Dynkin System is not generally extendable to a measure on the generated algebra (e.g., Example 3.1 in [33]). If a probability is extendable, its extension is in general non-unique.

Extendability of probabilities on (pre-)Dynkin systems has already been part of discussions in quantum probability since 1969 [19] up to more current times [47]. It as well has been of interest in frequential probability theory [20] (Theorem 2). Several necessary and/or sufficient conditions on the structure of $\mathcal{D}$ and/or the values of μ are known [33,46,47]. We present here a sufficient and necessary condition discovered by Horn and Tarski [48] and restated in [26] (Theorem 3.2.10). In fact, Theorem 3 can be stated for a more general definition of probabilities on arbitrary set systems e.g., [26] (Theorem 3.2.10), [49] (Proposition 2.2). For the sake of simplicity, we restrict this result to pre-Dynkin systems and probabilities defined on pre-Dynkin systems.

Theorem 3 (Extendability Condition [26] (Theorem 3.2.10)). Let $\mathcal{D}$ be a pre-Dynkin system on Ω . A finitely additive probability measure μ on $\mathcal{D}$ is extendable to 2^Ω if and only if

$$\sum_{k=1}^m \chi_{B_k}(\omega) - \sum_{j=1}^n \chi_{A_j}(\omega) \geq 0, \forall \omega \in \Omega \implies \sum_{k=1}^m \mu(B_k) - \sum_{j=1}^n \mu(A_j) \geq 0 \quad (2)$$

for all finite families of sets in $\mathcal{D}$: $A_1, \dots, A_n, B_1, \dots, B_m \in \mathcal{D}$.

Example 7. For $\mathcal{D}_4$ as in Example 4 let $\mu: \mathcal{D}_4 \rightarrow [0, 1]$ be defined as $\mu(\emptyset) = 0, \mu(12) = 0.5, \mu(34) = 0.5, \mu(13) = 0.2, \mu(24) = 0.8, \mu(\Omega) = 1$. The probability μ on $\mathcal{D}_4$ meets the extendability condition.

Extendability proves to be more than a helpful mathematical property for embedding pre-Dynkin systems and their respective probabilities into algebras. Whether a probability defined on $\mathcal{D}$ can be extended to a probability on 2^Ω is directly connected to the question whether the probability measure on $\mathcal{D}$ is coherent in the sense of [2] (pp. 68, 84) or not. Coherence is a minimal consistency requirement for probabilistic descriptions which has been introduced in the fundamental work of De Finetti [50] and developed by Walley [2]. Shortly summarizing, an incoherent imprecise probability is tantamount to an irrational betting behavior, thus the name. Thus, extendability is, besides its mathematical convenience, a desirable property of probabilities in pre-Dynkin settings.

We adapt here the definition of coherence of previsions in [2] (Definition 2.5.1) to probabilities.

Definition 5 (Coherent Probability). Let $\mathcal{A} \subseteq 2^\Omega$ be an arbitrary collection of subsets. A set function $\underline{\nu}: \mathcal{A} \rightarrow [0, 1]$ is a coherent lower probability if and only if

$$\sup_{\omega \in \Omega} \sum_{i=1}^j (\chi_{A_i}(\omega) - \underline{\nu}(A_i)) - m(\chi_{A_0}(\omega) - \underline{\nu}(A_0)) \geq 0,$$

for any non-negative $n, m \in \mathbb{N}$ and any $A_0, A_1, \dots, A_n \in \mathcal{A}$. If $\mathcal{A}$ is closed under complement, the conjugate coherent upper probability is given by $\bar{\nu}(A) := 1 - \underline{\nu}(A^c)$ for all $A \in \mathcal{A}$. If furthermore $\bar{\nu}(A) = \underline{\nu}(A)$ for all $A \in \mathcal{A}$, we call $\nu := \bar{\nu}$ a coherent additive probability.

At first sight, the Horn-Tarski condition given in Theorem 3 and the coherence condition presented here already appear similar. This becomes even more apparent in Walley's reformulation of coherence for additive probabilities [2] (Theorem 2.8.7). In the following, we show that this superficial similarity is indeed based on a rigorous link. Surprisingly, Walley did not mention Horn and Tarski's work in his foundational book.

Theorem 4 (Extendability Equals Coherence). Let $\mathcal{D}$ be a pre-Dynkin system on Ω . A finitely additive probability measure μ on $\mathcal{D}$ is extendable to 2^Ω if and only if it is a coherent additive probability on $\mathcal{D}$.

Proof. If μ is a coherent additive probability on $\mathcal{D}$, then the linear extension theorem [2] (Theorem 3.4.2) applies. Hence, a coherent additive probability $\nu: 2^\Omega \rightarrow [0, 1]$ exists, such that $\nu|_{\mathcal{D}} = \mu$. In particular, ν is a finitely additive probability following Definition 3 on 2^Ω [2] (Theorem 2.8.9).

For the converse direction, we observe that if μ possess an extension following Definition 4, then such an extension is a finitely additive probability on 2^Ω following Definition 3. Hence, Walley [2] (Theorem 2.8.9) guarantees that the extension is a coherent additive probability (Definition 5). Any restriction to a subdomain $\mathcal{D} \subseteq 2^\Omega$ keeps the probability coherent and additive. $\square$

The linear extension theorem in Walley [2] (Theorem 3.4.2) used here is a generalization of de Finetti's fundamental theorem of probability [50] (Theorems 3.10.1 and 3.10.7). De Finetti's theorem is furthermore interesting, as he explicitly states that a coherent additive probability defined on an arbitrary collection of sets can be extended in a precise way (so lower and upper probability coincide) to some sets. De Finetti does not characterize this collection. Our Theorem 2, however, gives an answer to this question: the collection forms a pre-Dynkin system. It is possible to recover the statement by an argumentation using credal sets. Roughly, credal sets of coherent lower probabilities consist of all dominating coherent additive probabilities. In the case here, extendability guarantees that the

credal set of a probability defined on a pre-Dynkin system is non-empty. Furthermore, Walley [2] introduced another slightly weaker concept than coherence, called “avoiding sure loss” which is equivalent to non-empty credal sets [2] (Corollary 3.3.4). Coherence and avoiding sure loss are equivalent for probabilities defined on pre-Dynkin systems [51] (Theorem 4.12).

Theorem 4 provides a missing link between two strands of work: on the one hand, probabilities on pre-Dynkin systems and related weak set structures have been closely investigated in foundational quantum probability theory [4,19] and decision theory [5,6]. On the other hand, coherent probabilities are central to imprecise probability, in particular, the more general formulations of coherent previsions and risk measures [2,52,53].

The reader familiar with the literature on imprecise probability might well not be surprised by the equivalence of extendability and coherence. We still think that this link is indeed valuable to be spelled out explicitly here. The concept of extendability and coherence have been developed separately in two communities with different goals in focus. Coherence tries to capture “rational” betting behavior [2,50]. Extendability links to what is sometimes called “quantum weirdness”. For readers familiar with quantum probability, we sketch the relationship between extendability and the concepts of compatibility and contextuality in the following.

4.2.1. Extendability, Compatibility and Contextuality

Extendability in quantum theory tightly interacts with a series of properties and concepts which pervade discussions about the “specialness” of quantum theory in comparison to other classical physical theories: compatibility, contextuality, hidden variables and more. To be concrete, two measurements are compatible if, for any initial state, usually represented as a probability distribution (cf. [4]), there exists a joint measurement such that a fixed joint distribution for both measurement outcomes exists, whose marginals are the distribution of the single measurement [54,55]. We remark that this notion of compatibility is related, but not directly, to our Definition 2. If measurements are incompatible, then there are potentially still states such that a joint distribution of measurements exists. Only in the cases that no joint distribution of measurements exists, i.e., extendability is not provided, a measuring observer observes *contextual* behavior [55]. Translated to the language of imprecise probability, contextuality amounts to non-coherence of a probabilistic description. Compatibility, in contrast, is a structural notion. If any finitely additive probability on a pre-Dynkin system is extendable, then the pre-Dynkin system, very roughly, resembles compatible measurements. We are indeed not the first to notice intriguing links between imprecise probability and concepts therein to quantum mechanics. Benavoli and collaborators recovered the four postulates of quantum mechanics with desirability as a starting point [56] (Desirability is a relatively general framework for imprecise probability [57].)

4.2.2. Extendability and Marginal Problem

A particularly helpful special case of probabilities defined on pre-Dynkin systems are marginal scenarios. Marginal scenarios, as we already stated in Section 2, are settings in which several (classical) marginal probability distributions, for instance, for three random variables p_X, p_Y, p_Z , are given, but a joint distribution $p_{X,Y,Z}$ is not specified. We note that we do not define random variables, marginals and probability distributions rigorously here. Instead, we argue on a vague level to convey the intuition. Such marginal scenarios can be expressed via probabilities on pre-Dynkin systems as laid out in detail in [34,58] and [33] (Example 4.2). Central to marginal scenarios is the relation structure of the marginals. For instance, $p_{X,Y}, p_Z$ as well as $p_{X,Y}, p_{Y,Z}$ or $p_{X,Y}, p_{Y,Z}, p_{X,Z}$ form marginal scenarios. The so-called marginal problem for marginal scenarios now asks whether for all instantiations of the marginal scenario there exists a joint distribution, e.g., for all $p_{X,Y}, p_Z$ a joint distribution $p_{X,Y,Z}$ exists. In contrast, for some assignments $p_{X,Y}, p_{Y,Z}, p_{X,Z}$ there does not exist a corresponding $p_{X,Y,Z}$. (The attentive reader might have noticed the similarity to the notion of compatibility, for good reasons [59] (§V.B.2). Marginal scenarios are used

to represent multi-measurement settings and compatibility among the measurements.) This question was, some while ago, asked for probabilities on finite spaces [34], countably additive probabilities [58], finitely additive probabilities (cf. [60]) and recently for even more general probability models—sets of desirable gambles [61,62]. A recurring theme in all those studies is the so-called running intersection property which characterizes all solvable marginal problems. In other words, the running intersection property guarantees that *every* instantiation of the marginal distributions is *extendable*. But there exist marginal problems for which only *specific* instantiations of the marginal distributions allow for extendability.

4.3. Coherent Extension

A probability on a pre-Dynkin system $\mathcal{D}$, even when extendable, only allows for probabilistic statements on $\mathcal{D}$ itself. However, extendability guarantees that a “nice” embedding into a larger system of measurable sets exists. More specifically, extendability expressed in terms of credal sets provides a well-known tool for the worst-case extension of a probability from a pre-Dynkin system to a larger algebra.

If a finitely additive probability on a pre-Dynkin system is extendable, then we can obtain lower and upper probabilities of events which are not in the pre-Dynkin system but on a larger algebra. We follow the idea of natural extensions, e.g., as described by [2] (p. 136). In particular, [2] (Theorem 3.3.4 (b)) directly applies as long as a probability on a pre-Dynkin system is extendable.

Corollary 1 (Coherent Extension of Probability). *Let $\mathcal{D}$ be a pre-Dynkin system on Ω . For a finitely additive probability measure μ on $\mathcal{D}$ we define the credal set*

$$M(\mu, \mathcal{D}) := \{v \in \Delta : v(A) = \mu(A), \forall A \in \mathcal{D}\}.$$

If μ on $\mathcal{D}$ is extendable to 2^Ω , then $\forall A \in 2^\Omega$,

$$\underline{\mu}_{\mathcal{D}}(A) := \inf_{v \in M(\mu, \mathcal{D})} v(A), \quad \bar{\mu}_{\mathcal{D}}(A) := \sup_{v \in M(\mu, \mathcal{D})} v(A).$$

define a coherent lower respectively upper probability on 2^Ω .

Example 8. *The coherent extension of μ on $\mathcal{D}_4$ as defined in Example 7 is $\underline{\mu}_{\mathcal{D}_4} = \ell$ and $\bar{\mu}_{\mathcal{D}_4} = u$ where, ℓ and u are defined as in Example 4 (cf. [2] (p. 122)). Figure 2 illustrates the coherent extensions. Even though coherent, $\underline{\mu}_{\mathcal{D}_4}$ is neither supermodular nor submodular:*

$$\begin{aligned} \underline{\mu}_{\mathcal{D}_4}(12) + \underline{\mu}_{\mathcal{D}_4}(13) &= 0.5 + 0.2 > 0.5 + 0.0 = \underline{\mu}_{\mathcal{D}_4}(123) + \underline{\mu}_{\mathcal{D}_4}(1), \\ \underline{\mu}_{\mathcal{D}_4}(1) + \underline{\mu}_{\mathcal{D}_4}(2) &= 0.0 + 0.3 < 0.5 + 0.0 = \underline{\mu}_{\mathcal{D}_4}(12) + \underline{\mu}_{\mathcal{D}_4}(\emptyset). \end{aligned}$$

This implies that as well $\bar{\mu}_{\mathcal{D}_4}$ is neither supermodular nor submodular [43] (Proposition 2.3).

These lower and upper probabilities allow for at least two interpretations: We can assume that a precise probability on a pre-Dynkin systems $\mathcal{D} \subseteq 2^\Omega$ just reveals its values on $\mathcal{D}$, but is actually defined over 2^Ω . Then the lower and upper probability constitute lower and upper bounds of the precise “hidden probability” on 2^Ω , which is solely accessible on $\mathcal{D}$. On the other hand, we can even reject the existence of such precise “hidden probability”. Then lower and upper probability *are* the inherently imprecise probability of an event in 2^Ω but not in $\mathcal{D}$. (As remarked by Walley [2] (p. 138), De Finetti [50] surprisingly only considered the first mentioned interpretation.)

The obtained lower and upper probabilities represent the imprecise interdependencies between all events of precise probabilities. We illustrate this statement: in the variety of updating methods in imprecise probability we pick the generalized Bayes’ rule [2] (§6.4) to exemplarily compute the conditional probability of two events for the coherent extension of

a probability from a pre-Dynkin system. For $A, B \in \mathcal{D}$ such that $\mu(B) > 0$, the generalized Bayes' rule gives [2] (Theorem 6.4.2):

$$\bar{\mu}_{\mathcal{D}}(A|B) := \sup_{\nu \in M(\mu, \mathcal{D})} \frac{\nu(A \cap B)}{\nu(B)} = \frac{\sup_{\nu \in M(\mu, \mathcal{D})} \nu(A \cap B)}{\mu(B)} = \frac{\bar{\mu}_{\mathcal{D}}(A \cap B)}{\mu(B)}, \quad \forall A, B \in \mathcal{D}$$

We can easily rearrange the above as $\bar{\mu}_{\mathcal{D}}(A \cap B) = \bar{\mu}_{\mathcal{D}}(A|B)\mu(B)$. In this case the imprecision of the probability of the intersected event is *purely* controlled by the conditional probability $\bar{\mu}(A|B)$ and *not* by the marginal, which is precise. So, the imprecision captured by the lower and upper probabilities locates solely in the interdependency of the events. We remark that Dempster's rule gives the same conditional probability here [63].

4.4. Inner and Outer Extension Is More Pessimistic Than Coherent Extension

We have presented two extension methods for probabilities defined on pre-Dynkin systems. We relate the methods in the following. In the case of an extendable probability we can guarantee the following inequalities to hold.

Theorem 5 (Extension Theorem—Finitely Additive Case). *Let $\mathcal{D}$ be a pre-Dynkin system on Ω and μ a finitely additive probability on $\mathcal{D}$ which is extendable to 2^{Ω} . Then,*

$$\mu_*(A) \leq \underline{\mu}_{\mathcal{D}}(A) \leq \bar{\mu}_{\mathcal{D}}(A) \leq \mu^*(A), \quad \forall A \in 2^{\Omega}.$$

Proof. Since $\mathcal{D} \subseteq 2^{\Omega}$, we easily obtain

$$\begin{aligned} \mu_*(A) &= \sup\{\mu(B) : A \supseteq B \in \mathcal{D}\} \\ &= \sup\{\underline{\mu}_{\mathcal{D}}(B) : A \supseteq B \in \mathcal{D}\} \\ &\leq \sup\{\underline{\mu}_{\mathcal{D}}(B) : A \supseteq B \in 2^{\Omega}\} \\ &= \underline{\mu}_{\mathcal{D}}(A), \end{aligned}$$

for all $A \in 2^{\Omega}$. The other inequalities follow by the conjugacy of inner and outer measure and lower and upper coherent extension. $\square$

In words, Theorem 5 states that the inner and outer extension is more “pessimistic” than the coherent extension. We use “pessimistic” in the sense of giving a looser bound for the probabilities assigned to elements not in the pre-Dynkin system $\mathcal{D}$ but in 2^{Ω} . We remark that Walley has shown that for a probability defined on a set algebra, inner and lower coherent extension, as well as outer and upper coherent extension, coincide [2] (Theorem 3.1.5). Walley even characterizes the sets for which a unique coherent extension exists [2] (Theorem 3.1.9). In Appendix D, we demonstrate analogous inequalities for countably additive probabilities on Dynkin systems.

5. The Credal Set and Its Relation to Pre-Dynkin System Structure

In the earlier parts of the paper, we derived pre-Dynkin systems as the system of precision for relatively general imprecise probabilities. Then, we showed that, under extendability conditions, a precise probability on a pre-Dynkin system gives rise to a coherent imprecise probability on an encompassing algebra. In other words, imprecise probabilities can be “mapped” to pre-Dynkin systems and vice versa. We concretize these mappings in the following. This manifestation then reveals structure in the interplay between the systems of precision, i.e., pre-Dynkin systems, and coherent imprecise probabilities. In particular, we argue that the order structure of pre-Dynkin systems can be mapped to the space of finitely additive probabilities. This provides a (lattice) duality for coherent imprecise probabilities with precise probabilities on pre-Dynkin systems. More concretely, the duality allows for the interpolation from imprecise probabilities which are precise on

“all” events to imprecise probabilities which are precise only on the empty set and the entire set.

In the following discussion, we assume, in addition to the technicalities presented in Section 1, that a fixed finitely additive probability on 2^Ω , which we call ψ , is given. The finitely additive probability ψ with the algebra 2^Ω and the base set Ω constitute our “base measure space” analogous to the choice of a base measure space in the theory of coherent risk measures [52]. The probability ψ can be interpreted as the pivot point around which we define increasingly imprecise coherent lower and upper probabilities. The choice of ψ particularly influences the sets of measure zero, which will play a major role in characterizing so-called bipolar-closed set systems. In comparison to the previous sections, we use ψ instead of μ as “reference measure” to emphasize the difference that μ was defined on a relatively arbitrary pre-Dynkin systems $\mathcal{D}$ on Ω , while ψ is defined and fixed on the algebra 2^Ω on Ω .

5.1. Credal Set Function Maps From Pre-Dynkin Systems to Coherent Probabilities

Equipped with a reference measure ψ we define the credal set function. The name arises due to its close link to the credal set as defined in Corollary 1.

Definition 6 (Credal Set Function). *Let Δ be the set of all finitely additive probabilities on 2^Ω . For a fixed, finitely additive probability $\psi \in \Delta$, we call*

$$m_\psi: 2^{2^\Omega} \rightarrow 2^\Delta, \quad m_\psi(\mathcal{A}) := \{\nu \in \Delta: \nu(A) = \psi(A), \forall A \in \mathcal{A}\},$$

the credal set function.

Example 9. *Let $\Omega_4 = \{1, 2, 3, 4\}$ as in Example 1. With abuse of notation, we define the probability $\psi: 2^{\Omega_4} \rightarrow [0, 1]$ via its corresponding point on the simplex $\psi \in \Delta$, $\psi_1 = 0.2, \psi_2 = 0.3, \psi_3 = 0.5, \psi_4 = 0$. It follows, e.g., $m_\psi(\{12, 3\}) = \{\nu \in \Delta: \nu_1 + \nu_2 = 0.5, \nu_3 = 0.5\}$. In the subsequent examples, we implicitly assume the here defined ψ .*

We stress that although not notated explicitly, the credal set function depends upon the choice of ψ . For a fixed ψ on 2^Ω , the credal set function m_ψ maps a subset of the algebra 2^Ω to the set of all finitely additive probabilities which coincide with ψ on this subset. It should be noticed that by definition of ψ , $m_\psi(\mathcal{A}) \neq \emptyset$ for every non-empty $\mathcal{A} \subseteq 2^\Omega$, because $\psi \in m_\psi(\mathcal{A})$ for every $\mathcal{A} \subseteq 2^\Omega$.

This defined mapping now simplifies our discussion about how pre-Dynkin systems and imprecise probabilities correspond. For instance, one can easily see that the extreme case $\mathcal{A} = 2^\Omega$ corresponds to $m_\psi(\mathcal{A}) = \{\psi\}$ and $\mathcal{A} = \emptyset$ to $m_\psi(\mathcal{A}) = \Delta$. More generally, we observe the following two properties of the credal set function.

Proposition 2 (Credal Set Function is Invariant to Pre-Dynkin Hull). *Let m_ψ be the credal set function. For any $\mathcal{A} \subseteq 2^\Omega$,*

$$m_\psi(\mathcal{A}) = m_\psi(D(\mathcal{A})).$$

Proof. We need to show that

$$\{\nu \in \Delta: \nu(A) = \psi(A), A \in \mathcal{A}\} = \{\nu \in \Delta: \nu(A) = \psi(A), A \in D(\mathcal{A})\}.$$

The set inclusion of the right hand side in the left hand side is trivial. For the reverse direction, consider an element $\nu \in \Delta$ such that $\nu(A) = \psi(A)$ for $A \in \mathcal{A}$. Let

$$\mathcal{H} := \{A \in D(\mathcal{A}): \nu(A) = \psi(A)\}.$$

By Theorem 2 $\mathcal{H}$ is a pre-Dynkin system. Since $\mathcal{A} \subseteq \mathcal{H}$ we know $D(\mathcal{A}) \subseteq \mathcal{H}$. Hence, $\nu(A) = \mu(A)$ for $A \in D(\mathcal{A})$. This gives the desired inclusion. We remark that for $\mathcal{A} = \emptyset$ the equality still holds, since $D(\emptyset) = \{\emptyset, \Omega\}$. $\square$

Example 10. The credal set $m_\psi(\{12, 3\}) = \{\nu \in \Delta: \nu_1 + \nu_2 = 0.5, \nu_3 = 0.5\}$ given in Example 9 nicely illustrates Proposition 2:

$$m_\psi(\{12, 3\}) = \{\nu \in \Delta: \nu_1 + \nu_2 = 0.5, \nu_3 = 0.5, \nu_4 = 0\} = m_\psi(\{\emptyset, 12, 3, 4, 34, 123, 124, \Omega_4\}).$$

Proposition 3 (Credal Set Function Maps to Weak*-Closed Convex Sets). *Let m_ψ be the credal set function. For every non-empty $\mathcal{A} \subseteq 2^\Omega$, $m_\psi(\mathcal{A})$ is weak*-closed convex.*

Proof. The reference probability ψ is by definition coherent. Hence, for all non-empty $\mathcal{A} \subseteq 2^\Omega$, the set $m_\psi(\mathcal{A})$ is the set of all $\nu \in \Delta$ which dominate ψ . This set is, by Theorem 3.6.1 in [2], weak*-closed and convex. $\square$

In words, the credal set on some set system coincides with the credal set on its generated pre-Dynkin system. And the credal set of probabilities is always weak*-closed and convex. Proposition 2 allows us to work with credal sets of arbitrary set systems instead of the entire pre-Dynkin system. Thus, it resembles the well-known π - λ -Theorem, which is fundamental to classical probability theory [23] (Lemma A.1.3). On the other hand, this result justifies our focus on pre-Dynkin systems instead of arbitrary set systems. We do not lose generality when considering pre-Dynkin systems instead of non-structured sets of sets.

Proposition 3 guarantees that the images of the credal set function behave “nicely”. Specifically, these weak*-closed convex sets correspond to coherent previsions, i.e., generalizations of coherent probabilities as already stated by Walley [2] (Theorem 3.6.1). We elaborate this observation in Section 5.4. In conclusion, credal set functions map pre-Dynkin systems to coherent probabilities. What about the reverse mapping?

5.2. The Dual Credal Set Function

The following is a natural definition of a dual credal set function. We justify this name by Proposition 4 below.

Definition 7 (Dual Credal Set Function). *Let Δ be the set of all finitely additive probabilities on 2^Ω . Fix a finitely additive probability ψ on 2^Ω . We call*

$$m_\psi^\circ: 2^\Delta \rightarrow 2^{2^\Omega}, \quad m_\psi^\circ(Q) := \{A \in 2^\Omega: \nu(A) = \psi(A), \forall \nu \in Q\}$$

the dual credal set function.

The dual credal set function also depends upon ψ , but we do not notate this explicitly. The dual credal set function maps an arbitrary set of finitely additive probability measures on 2^Ω to the (largest) set of events on which all contained probabilities coincide. We remark that each set of finitely additive probability measures can be linked to an imprecise probability.

We suggestively called the antagonist to the credal set function the “dual credal set function”. The duality appearing here is a well-known fundamental relationship between partially ordered sets: a Galois connection. A *Galois connection* is a pair of mappings $f: X \rightarrow Y$ and $f^\circ: Y \rightarrow X$ on partially ordered sets $(Y, \leq)$ and $(X, \leq)$, which preserves order structure (cf. Corollary 2). More formally, f and f° are a Galois connection if and only if for all $x \in X, y \in Y$, $x \leq f(y) \Leftrightarrow y \leq f^\circ(x)$ [64] (§V.8). Galois connections, even though they do not form an order isomorphism, induce a lattice duality. We exploit this lattice duality to provide an order-theoretic interpolation from no compatibility at all to full compatibility. For this purpose, we first establish the Galois connection.

Proposition 4 (Galois Connection by (Dual) Credal Set Function). *The credal set function m_ψ and the dual credal set function m_ψ° form a Galois connection.*

Proof. m_ψ and m_ψ° form a Galois connection if and only if $\mathcal{A} \subseteq m_\psi^\circ(Q) \Leftrightarrow Q \subseteq m_\psi(\mathcal{A})$ [64] (§V.8). First, we show the left to right implication. We assume $\mathcal{A} \subseteq m_\psi^\circ(Q)$, i.e. every $v \in Q$ coincides with ψ on $\mathcal{A}$. Hence,

$$\begin{aligned} v \in Q &\Rightarrow v(A) = \psi(A), \forall A \in \mathcal{A} \\ &\Rightarrow v \in \{v' \in \Delta: v'(A) = \psi(A), \forall A \in \mathcal{A}\} = m_\psi(\mathcal{A}). \end{aligned}$$

In case of the right to left implication, we suppose $Q \subseteq m_\psi(\mathcal{A})$. Thus,

$$\begin{aligned} A \in \mathcal{A} &\Rightarrow v(A) = \psi(A), \forall v \in Q \\ &\Rightarrow A \in \{A' \in 2^\Omega: v(A') = \psi(A'), \forall v \in Q\} = m_\psi^\circ(Q). \end{aligned}$$

□

The mappings involved in the Galois connection are antitone, i.e., they reverse the order structure from domain to codomain. Their pairwise application is extensive, i.e., the image of an object contains the object. In summary, the following rules of calculation hold:

Corollary 2 (Rules for (Dual) Credal Set Function). *Let m_ψ be the credal set function and m_ψ° be the dual credal set function. For arbitrary $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A} \subseteq 2^\Omega$ and $Q_1, Q_2, Q \subseteq \Delta$,*

$$\begin{aligned} \mathcal{A}_1 \subseteq \mathcal{A}_2 &\Rightarrow m_\psi(\mathcal{A}_2) \subseteq m_\psi(\mathcal{A}_1), & Q_1 \subseteq Q_2 &\Rightarrow m_\psi^\circ(Q_2) \subseteq m_\psi^\circ(Q_1), & (\text{antitone}) \\ Q &\subseteq m_\psi(m_\psi^\circ(Q)), & \mathcal{A} &\subseteq m_\psi^\circ(m_\psi(\mathcal{A})), & (\text{extensive}) \\ m_\psi(\mathcal{A}) &= m_\psi(m_\psi^\circ(m_\psi(\mathcal{A}))), & m_\psi^\circ(Q) &= m_\psi^\circ(m_\psi(m_\psi^\circ(Q))), & (\text{pseudo-inverse}). \end{aligned}$$

Proof. [64] (§V.7 and V.8) □

Proposition 4 provides a tool to further investigate the dual credal set function. The reader might have noticed the similarity of the dual credal set function and the main question of Section 3: given a lower and upper probability, on which set systems do both coincide? In fact, we obtain an analogous result to Theorem 2, and again an imprecise probability is mapped to the set of events on which it is precise.

Proposition 5 (Dual Credal Set Function Maps to Pre-Dynkin Systems). *Let m_ψ° be the dual credal set function. For all non-empty $Q \subseteq \Delta$, $m_\psi^\circ(Q)$ is a pre-Dynkin system.*

Proof. We show the statement by establishing the equality $D(m_\psi^\circ(Q)) = m_\psi^\circ(Q)$. Trivially, $D(m_\psi^\circ(Q)) \supseteq m_\psi^\circ(Q)$. Furthermore,

$$\begin{aligned} D(m_\psi^\circ(Q)) &\stackrel{\text{C.2}}{\subseteq} m_\psi^\circ(m_\psi(D(m_\psi^\circ(Q)))) \\ &\stackrel{\text{P.2}}{=} m_\psi^\circ(m_\psi(m_\psi^\circ(Q))) \stackrel{\text{C.2}}{=} m_\psi^\circ(Q). \end{aligned}$$

□

Proposition 6 (Dual Credal Set Function is Invariant to Weak*-Closed Convex Hull). *Let m_ψ° be the dual credal set function. For any $Q \subseteq \Delta$, $m_\psi^\circ(Q) = m_\psi^\circ(\overline{\text{co}} Q)$.*

Proof. By Corollary 2, $m_\psi^\circ(Q) = m_\psi^\circ(m_\psi(m_\psi^\circ(Q)))$. Furthermore, $Q \subseteq m_\psi(m_\psi^\circ(Q))$. Via, Proposition 3 we obtain $\overline{\text{co}} Q \subseteq m_\psi(m_\psi^\circ(Q))$. Hence, the result follows. □

Example 11. Let $Q = \{v\}$ where we identify the probability v on 2^{Ω_4} with an element $v \in \Delta_4$. For instance, $v_1 = 0, v_2 = 0.5, v_3 = 0, v_4 = 0.5$. It is easy to see that $m_\psi^\circ(Q) = D(\{12, 3\})$ as given in Example 10.

In summary, credal set functions map pre-Dynkin systems to weak*-closed convex credal sets. Dual credal set functions map (weak*-closed convex) credal sets to pre-Dynkin systems. In addition, the two functions form a Galois connection. In fact, every Galois connection defines closure operators, i.e., extensive, monotone and idempotent maps [25] (Definition 4.5.a). The closure operators are defined as the sequential application of the credal set function and the dual credal set function to subsets of Δ or 2^Ω . In symbols: $\mathcal{A} \mapsto m_\psi^\circ(m_\psi(\mathcal{A})), Q \mapsto m_\psi(m_\psi^\circ(Q))$. In particular, these closure operators define bipolar-closed sets.

5.3. Bipolar-Closed Sets

Bipolar-closed sets are sets $\mathcal{A} \subseteq 2^\Omega$ such that $\mathcal{A} = m_\psi^\circ(m_\psi(\mathcal{A}))$, respectively, $Q \subseteq \Delta$ such that $Q = m_\psi(m_\psi^\circ(Q))$. Most importantly, the bipolar-closed sets form two antitone isomorphic lattices ordered by set inclusion [64] (Theorem V.8.20). This relationship gives us a lattice duality between set systems and credal sets of probabilities. See Figure 3 for an illustration of bipolar-closed sets and the Galois connection.

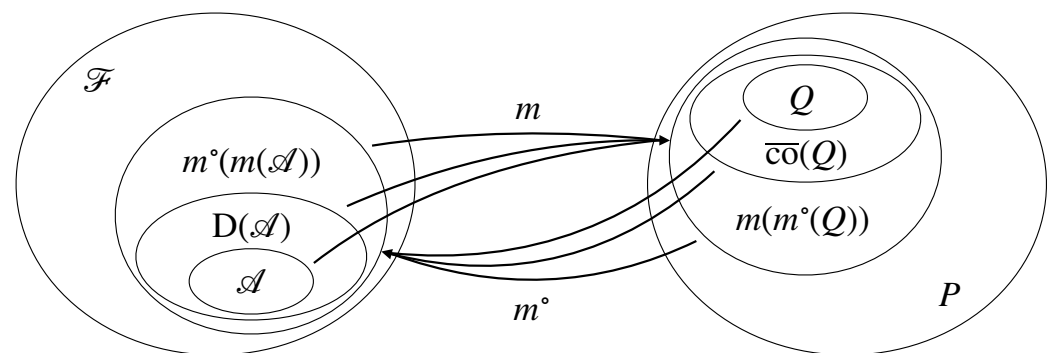


Figure 3. Galois connection between the lattice of pre-Dynkin systems and the set of credal sets. In the illustrated case, we have $m_\psi^\circ(Q) = m_\psi^\circ(m_\psi(m_\psi^\circ(\mathcal{A})))$, respectively, $m_\psi(\mathcal{A}) = m_\psi(m_\psi^\circ(m_\psi(Q)))$. The set containment on both sides follows from Proposition 2, Corollary 2 and Proposition 6.

Example 12. The pre-Dynkin system $D(\{12, 3\})$ already discussed in Example 10 and Example 11 is a bipolar-closed set. In contrast, the set $\{12, 3\}$ cannot be a bipolar-closed set, as it is not a pre-Dynkin system.

More precisely, bipolar-closed sets in the set of finitely additive probability distributions are weak*-closed convex (Proposition 3). These map to bipolar-closed subsets of 2^Ω , which are pre-Dynkin systems (Proposition 5). All of the stated properties of bipolar-closed sets are necessary. But are they sufficient?

Sufficient Conditions for Bipolar-Closed Sets

In the search for sufficient conditions for bipolar-closed sets we focus on bipolar-closed subsets of 2^Ω . We leave bipolar-closed subsets of Δ to future investigations. Already, bipolar-closed subsets of 2^Ω are surprisingly difficult to characterize. We show that bipolar-closed sets are exactly the pre-Dynkin systems if the reference probability measure has support on all elements of a finite set. To this end, let us start with a simple implication of Proposition 2 and Corollary 2.

Corollary 3. Let m_ψ be the credal set function and m_ψ° be the dual credal set function. For an arbitrary subset $\mathcal{A} \subseteq 2^\Omega$ we have

$$D(\mathcal{A}) \subseteq m_\psi^\circ(m_\psi(\mathcal{A})).$$

Proof. By Proposition 2, $m_\psi^\circ(m_\psi(\mathcal{A})) = m_\psi^\circ(m_\psi(D(\mathcal{A})))$. By Corollary 2, the statement follows. $\square$

This corollary gives rise to the follow-up question: under which circumstances does $D(\mathcal{A}) = m_\psi^\circ(m_\psi(\mathcal{A}))$? As the following theorem demonstrates, this question is closely connected to the sets of measure zero of the base probability ψ and its problems (cf. [65]).

Proposition 7 (“Closedness” under Measure Zero Sets). Let m_ψ be the credal set function and m_ψ° be the dual credal set function. Let $\mathcal{A} \subseteq 2^\Omega$. For any $A \in m_\psi^\circ(m_\psi(\mathcal{A}))$, if $\psi(A) = 0$, then all $B, C \in 2^\Omega$ such that $B \subseteq A$ and $C \supseteq A^c$ are in $m_\psi^\circ(m_\psi(\mathcal{A}))$.

Proof. Let $A \in m_\psi^\circ(m_\psi(\mathcal{A}))$ for arbitrary $\mathcal{A} \subseteq 2^\Omega$ with $\psi(A) = 0$. Consider two sets $B, C \in 2^\Omega$ such that $B \subseteq A$, respectively, $C \supseteq A^c$. Then, $\psi(B) = 0$ and $\psi(C) = 1$. Furthermore, for any $\nu \in m_\psi(m_\psi^\circ(m_\psi(\mathcal{A})))$, we have $\nu(B) = 0$ (respectively, $\nu(C) = 1$). Since $m_\psi(m_\psi^\circ(m_\psi(\mathcal{A}))) = m_\psi(\mathcal{A})$, we obtain $B, C \in m_\psi^\circ(m_\psi(\mathcal{A}))$. $\square$

A pre-Dynkin system $D(\mathcal{A})$ can only coincide with $m_\psi^\circ(m_\psi(\mathcal{A}))$ if subsets of measure zero are included. Thus, Proposition 7 provides a further necessary condition for bipolar-closed subsets of 2^Ω . Yet, it turns out that the sets of measure zero as well can give a sufficient condition for bipolar-closed sets at least in a finite setting.

Theorem 6 (Pre-Dynkin Hull is Bipolar-Closure in Finite, Discrete Setting). Let $\Omega = [n]$. Fix a finitely additive probability ψ on 2^Ω such that $\psi(A) > 0$ for every $A \in 2^\Omega \setminus \{\emptyset\}$. Let $\mathcal{D} \subseteq 2^\Omega$ be a pre-Dynkin system. Then,

$$\mathcal{D} = m_\psi^\circ(m_\psi(\mathcal{D})).$$

Proof. If $\mathcal{D} = 2^{[n]}$ the result follows directly. Thus, from now on we assume $\mathcal{D} \subsetneq 2^{[n]}$. The set inclusion $\mathcal{D} \subseteq m_\psi^\circ(m_\psi(\mathcal{D}))$ is given by Corollary 3. We show $\mathcal{D} \supseteq m_\psi^\circ(m_\psi(\mathcal{D}))$ by proving that for every $B \in 2^{[n]} \setminus \mathcal{D}$ there exists $\nu \in m_\psi(\mathcal{D})$ such that $\nu(B) \neq \psi(B)$.

Let us consider an arbitrary $B \in 2^{[n]} \setminus \mathcal{D}$. Without loss of generality (Lemma A2) we can decompose

$$B = B_{\mathcal{D}} \cup A,$$

where $B_{\mathcal{D}} \in \mathcal{D}$ and $A \notin \mathcal{D}$ is a weak atom with respect to $\mathcal{D}$ (Definition A3). Thus, we can leverage Lemma A3: there exists $\nu \in m_\psi(\mathcal{D})$ such that $\nu(A) \neq \psi(A)$. Thus,

$$\nu(B) = \nu(B_{\mathcal{D}}) + \nu(A) = \psi(B_{\mathcal{D}}) + \nu(A) \neq \psi(B_{\mathcal{D}}) + \psi(A).$$

It follows that $B \notin m_\psi^\circ(m_\psi(\mathcal{D}))$, concluding the proof. $\square$

We emphasize that the statement does not hold as long as there are non-empty subsets of Ω which get assigned zero probability.

Example 13 (Bipolar-Closure is Strictly Greater Than Pre-Dynkin Hull). Let $\Omega_3 = \{1, 2, 3\}$ and $\psi(\{1\}) = 1, \psi(\{2\}) = 0, \psi(\{3\}) = 0$ constitute a base probability space. We have $D(\{1\}) = \{\emptyset, 1, 23, \Omega_3\}$, but $m_\psi^\circ(m_\psi(\{1\})) = 2^\Omega$.

Whether this theorem can be extended to more general sets Ω is an open question. Seemingly, proofs along the line of Theorem 6 are doomed to fail, since one cannot argue via probabilities on atoms of Ω .

5.4. Interpolation From Algebra to Trivial Pre-Dynkin System

In probability theory there is a choice to be made regarding which events should get assigned probabilities [1] (p. 52). This significant choice has (mathematically) been standardized to form a (σ) -algebra (cf. standard probability space). But, already Kolmogorov, the “father” of modern probability theory, emphasized that this choice is not universal but should depend on the problem at hand. More recently, Khrennikov [66] argued that a more appropriate probabilistic modeling should appeal to weaker domains for probabilities to, for instance, represent physical observations such as quantum phenomena.

In particular, it cannot always be taken for granted that all events are compatible with all others, as implied by a (σ) -algebra (cf. Section 2.1). For instance, von Mises’ axiomatization of probability inherently reflects potential incompatible events in terms of a pre-Dynkin system [20,29]. In other words, there is a choice to be made about the system of precision. Which sets should be compatible to each other, and which should not? How do the choices of the systems of precision relate to each other?

We neglect, without loss of generality, arbitrary systems of precision and focus on pre-Dynkin systems (cf. Proposition 2). The range of choices is captured by the system of pre-Dynkin systems.

Proposition 8 (Set of Pre-Dynkin Systems is a Lattice). *The set $\mathfrak{D} := \{\mathcal{D} \subseteq 2^\Omega : \mathcal{D} = D(\mathcal{D})\}$ is ordered by set inclusion. Furthermore, $(\mathfrak{D}, \subseteq)$ is a lattice with*

$$\bigvee_{i=1}^n \mathcal{D}_i = D\left(\bigcup_{i=1}^n \mathcal{D}_i\right)$$

$$\bigwedge_{i=1}^n \mathcal{D}_i = \bigcap_{i=1}^n \mathcal{D}_i.$$

Proof. On the one hand, it is easy to show that the intersection of pre-Dynkin systems forms a pre-Dynkin system again. On other hand, the smallest pre-Dynkin system which contains a finite set of pre-Dynkin systems is by definition the pre-Dynkin system generated by the union over all elements in this finite set of pre-Dynkin systems. $\square$

Example 14. Let Ω_4 be as defined in Example 9. The minimal element in $\mathfrak{D}$ then is $\{\emptyset, \Omega_4\}$. The maximal element is 2^{Ω_4} . For the sake of brevity, we omit all further elements in $\mathfrak{D}$ and remain with the observation that $\mathcal{D}_4$ of Example 1 and $D(\{12, 3\})$ are elements of $\mathfrak{D}$.

The lattice $\mathfrak{D}$ spans a range of choices from $\mathcal{D} = 2^\Omega$, i.e., complete compatibility and only a single probability distribution in its credal set, namely $m_\psi(2^\Omega) = \{\psi\}$, to $\mathcal{D} = \{\emptyset, \Omega\}$, i.e., no compatibility and the entire space of probability distributions constitute its credal set $m_\psi(\{\emptyset, \Omega\}) = \Delta$. How “close” $\mathcal{D}$ is to the algebra 2^Ω determines how “classical” the credal set behaves. In other words, $(\mathfrak{D}, \subseteq)$ parametrizes a family of credal sets. Thus, it parametrizes coherent probabilities. The knob of compatibility can be turned from trivially nothing $(\{\emptyset, \Omega\})$, to everything (2^Ω) . How does the “amount of compatibility” of the pre-Dynkin system map to the credal sets? Or, e.g., given two pre-Dynkin systems on which a probability is defined, what is the credal set of the union of these systems?

Proposition 9 (Lattice of Dynkin Systems and Credal Sets). *The credal set function m_ψ (Definition 6) together with the lattice $(\mathfrak{D}, \subseteq)$ provides a parametrized family of credal sets for which hold $(\forall \mathcal{D}_1, \mathcal{D}_2 \in \mathfrak{D})$:*

$$m_\psi(\mathcal{D}_1 \vee \mathcal{D}_2) = m_\psi(\mathcal{D}_1 \cup \mathcal{D}_2) = m_\psi(\mathcal{D}_1) \cap m_\psi(\mathcal{D}_2)$$

$$m_\psi(\mathcal{D}_1 \wedge \mathcal{D}_2) = m_\psi(\mathcal{D}_1 \cap \mathcal{D}_2) \supseteq m_\psi(\mathcal{D}_1) \cup m_\psi(\mathcal{D}_2).$$

Proof. Concerning the first equality, we observe that for arbitrary $\mathcal{D}_1, \mathcal{D}_2 \in \mathfrak{D}$

$$m_\psi(\mathcal{D}_1 \vee \mathcal{D}_2) = m_\psi(D(\mathcal{D}_1 \vee \mathcal{D}_2)) = m_\psi(\mathcal{D}_1 \cup \mathcal{D}_2),$$

by Proposition 2. Consequently,

$$\begin{aligned} m_\psi(\mathcal{D}_1 \cup \mathcal{D}_2) &= \{v \in \Delta : v(D) = \psi(D), \forall D \in \mathcal{D}_1 \cup \mathcal{D}_2\} \\ &= \{v \in \Delta : v(D) = \psi(D), \forall D \in \mathcal{D}_1\} \cap \{v \in \Delta : v(D) = \psi(D), \forall D \in \mathcal{D}_2\} \\ &= m_\psi(\mathcal{D}_1) \cap m_\psi(\mathcal{D}_2). \end{aligned}$$

The second line follows by the definition of infimum on the lattice of pre-Dynkin systems and simple set containment: $m_\psi(\mathcal{D}_1) \subseteq m_\psi(\mathcal{D}_1 \cap \mathcal{D}_2)$ and $m_\psi(\mathcal{D}_2) \subseteq m_\psi(\mathcal{D}_1 \cap \mathcal{D}_2)$. $\square$

Unfortunately, the mentioned interpolation is slightly improper. It turns out that there are pre-Dynkin systems $\mathcal{D}_1 \neq \mathcal{D}_2$ such that $m_\psi(\mathcal{D}_1) = m_\psi(\mathcal{D}_2)$.

Example 15 (Non-Injectivity of Credal Set Function). Let $\Omega_3 = \{1, 2, 3\}$ and $\psi(\{1\}) = 1, \psi(\{2\}) = 0, \psi(\{3\}) = 0$ constitute a base probability space, cf. Example 13. Then, obviously $m_\psi(D(\{1\})) = m_\psi(2^\Omega)$, but $D(\{1\}) \neq 2^\Omega$.

The reason for this collision of credal sets is that not every pre-Dynkin system $\mathcal{D} \subseteq 2^\Omega$ is a bipolar-closed set.

Proposition 10 (Credal Set Function is Injective on Bipolar-Closed Sets). Let m_ψ be the credal set and m_ψ° be the dual credal set function. Let $\mathcal{D}_1, \mathcal{D}_2 \in \mathfrak{D}$ be pre-Dynkin systems, which are bipolar-closed. If $\mathcal{D}_1 \neq \mathcal{D}_2$, then $m_\psi(\mathcal{D}_1) \neq m_\psi(\mathcal{D}_2)$.

Proof. We prove the claim by contraposition:

$$m_\psi(\mathcal{D}_1) = m_\psi(\mathcal{D}_2) \Rightarrow m_\psi^\circ(m_\psi(\mathcal{D}_1)) = m_\psi^\circ(m_\psi(\mathcal{D}_2)) \Rightarrow \mathcal{D}_1 = \mathcal{D}_2.$$

$\square$

Hence, it is reasonable to focus on the set of bipolar-closed sets contained in 2^Ω . We define the set of interpolating pre-Dynkin systems $\mathfrak{C} := \{\mathcal{C} \subseteq 2^\Omega : \mathcal{C} = m_\psi^\circ(m_\psi(\mathcal{C}))\}$. We know that $\mathfrak{C} \subseteq \mathfrak{D}$ (Proposition 5) and $\mathfrak{C}$ is even a lattice contained in $\mathfrak{D}$.

Proposition 11 (Lattice of Bipolar-Closed Sets and Credal Sets). Let m_ψ be the credal set function and m_ψ° be the dual credal set function. The set of interpolating pre-Dynkin systems $\mathfrak{C}$ equipped with the $\subseteq$ -ordering forms a lattice:

$$\begin{aligned} \bigvee_{i=1}^n \mathcal{C}_i &= m_\psi^\circ \left(m_\psi \left(\bigcup_{i=1}^n \mathcal{C}_i \right) \right) \\ \bigwedge_{i=1}^n \mathcal{C}_i &= \bigcap_{i=1}^n \mathcal{C}_i. \end{aligned}$$

In particular, this lattice $\mathfrak{C}$ is antitone isomorphic to the lattice of bipolar-closed sets in 2^Δ . It holds (We denote the composition of functions with $\circ$.):

$$\begin{aligned} m_\psi(\mathcal{C}_1 \vee \mathcal{C}_2) &= m_\psi(\mathcal{C}_1 \cup \mathcal{C}_2) = m_\psi(\mathcal{C}_1) \cap m_\psi(\mathcal{C}_2) \\ m_\psi(\mathcal{C}_1 \wedge \mathcal{C}_2) &= m_\psi(\mathcal{C}_1 \cap \mathcal{C}_2) = (m_\psi \circ m_\psi^\circ)(m_\psi(\mathcal{C}_1) \cup m_\psi(\mathcal{C}_2)). \end{aligned}$$

Proof. By Theorem V.8.20 in [64], $\mathfrak{C}$ is a lattice and m_ψ an antitone lattice isomorphism on $\mathfrak{C}$. The equations hold by simple manipulations

$$m_\psi(\mathcal{C}_1 \vee \mathcal{C}_2) = m_\psi(m_\psi^\circ(m_\psi(\mathcal{C}_1 \cup \mathcal{C}_2))) = m_\psi(\mathcal{C}_1 \cup \mathcal{C}_2) = m_\psi(\mathcal{C}_1) \cap m_\psi(\mathcal{C}_2),$$

and

$$m_\psi(\mathcal{C}_1 \wedge \mathcal{C}_2) = m_\psi(m_\psi^\circ(m_\psi(\mathcal{C}_1 \cap \mathcal{C}_2))) = (m_\psi \circ m_\psi^\circ)(m_\psi(\mathcal{C}_1) \cup m_\psi(\mathcal{C}_2)).$$

□

Note that $\mathfrak{C}$ is not generally a sublattice of $\mathfrak{D}$. It is a lattice contained in the lattice $\mathfrak{D}$, but the closure operator for the supremum is distinct. Interestingly, the order structure which both lattices, $\mathfrak{D}$ and $\mathfrak{C}$, induce on the set of credal sets via m_ψ is identical. Every set $m_\psi(\mathcal{A})$ for arbitrary $\mathcal{A} \subseteq 2^\Omega$ is bipolar-closed (Corollary 2). Hence, the lattice of bipolar-closed sets in 2^Δ is the domain of m_ψ for elements in $\mathfrak{D}$ and $\mathfrak{C}$. In other words, the lattices $\mathfrak{D}$ and $\mathfrak{C}$ provide one and the same parametrized family of credal sets, thus one and the same parametrized family of imprecise probabilities.

In comparison to other parametrized families of imprecise probability, such as distortion risk measures [67], which heavily rely on convex analysis, the duality used here is structurally weaker. Lattice isomorphisms give a glimpse of structure to the involved dual spaces. Convex dualities as exploited in [68] are far more informative, but apparently not able to handle the structural knob which we presented in this work: the set of sets which get assigned precise probabilities. Nevertheless, a natural question arises from this lattice duality: how does this lattice duality relate to a convex duality? We leave this question open to further research. A first attempt to an answer is discussed in Appendix C, where we link the parametrized family of distorted probabilities to the pre-Dynkin system family of imprecise probabilities.

6. A More General Perspective—The Set of Gambles with Precise Expectation

To this point, we have exclusively focused on probabilities and set systems of events to which we assign probabilities. In fact, there is a more general story to be told. In the literature on imprecise probability, focus often lies on expectation-type functionals instead of probabilities and on sets of gambles (bounded functions from the base set Ω to the real numbers) instead of sets of events. One can easily see that the latter is more general and can recover the former. Indicator functions of events are gambles. An expectation-type functional evaluated on an indicator gamble of an event corresponds to a generalized probability of the event. The converse direction, i.e., recovering a unique expectation-type functional from an imprecise probability, however, is not always possible [2] (§2.7.3). In the following, we reiterate several questions which we asked in the preceding sections for probabilities and set systems.

6.1. Partial Expectations Generalize Finitely Additive Probabilities on (Pre-)Dynkin Systems

We propose the following definition of partial expectation and show afterwards that it is a natural generalization of finitely additive probabilities defined on (pre-)Dynkin systems.

Definition 8 (Partial Expectation). Let $\{L_i\}_{i \in I}$ be a non-empty family of linear subspaces of $B(\Omega)$. We call $E: \bigcup_{i \in I} L_i \rightarrow \mathbb{R}$ a partial expectation if and only if all of the following conditions are fulfilled:

- (a) for any $i \in I$ and for all $f, g \in L_i$, then $E(f + g) = E(f) + E(g)$, (Partial Linearity),
- (b) for any $i \in I$ and any $f \in L_i$, then $E(f) \geq \inf f$, (Coherence).

We remark that for this definition, we leveraged the requirements for a linear prevision (Definition 9) on a linear space given in [2] (Theorem 2.8.4). In other words, a partial

expectation is a functional which is defined on a union of linear subspaces and behaves like a “classical” (finitely additive) expectation on each of the subspaces but not necessarily on all simultaneously. It is a linear prevision when restricted to one of the subspaces L_i (cf. Definition 9).

There is a one-to-one correspondence of linear previsions and coherent additive probabilities. For every coherent additive probability ν defined on an algebra $\mathcal{A}$, there is a unique linear prevision, which we equivalently denote ν , the set of all $\mathcal{A}$ -measurable gambles, which agrees with the probability ν on the indicator gambles of the sets in $\mathcal{A}$ [2] (Theorem 3.2.2). The following Proposition exploits this correspondence. A finitely additive probability defined on a pre-Dynkin system relates one-to-one to a partial expectation which is defined on the set of linear spaces induced by the simple gambles on the blocks of the pre-Dynkin system. To this end, we introduce the following two notations: let $\mathcal{F} \subseteq 2^\Omega$ be an algebra. Then, $S(\Omega, \mathcal{F}) \subseteq B(\Omega)$ denotes the linear subspace of simple gambles on $\mathcal{F}$, i.e., scaled and added indicator gambles of a finite number of disjoint sets (cf. [26] (Definition 4.2.12)). Let $\mathcal{F}_\sigma \subseteq 2^\Omega$ be a σ -algebra. Then $B(\Omega, \mathcal{F}_\sigma) \subseteq B(\Omega)$ denotes the linear subspace of all bounded, real-valued, $\mathcal{F}_\sigma$ -measurable gambles.

Proposition 12 (Finitely Additive Probability on Pre-Dynkin System and its Partial Expectation). *Let $\mathcal{D} \subseteq 2^\Omega$ be a pre-Dynkin system. Let $\mu: \mathcal{D} \rightarrow [0, 1]$ be a finitely additive probability defined on the pre-Dynkin system with block structure $\{A_i\}_{i \in I}$. Then, μ is in one-to-one correspondence to a partial expectation $E: \bigcup_{i \in I} S(\Omega, A_i) \rightarrow \mathbb{R}$ defined on the union of linear spaces of simple gambles induced by all blocks A_i of $\mathcal{D}$.*

Proof. By Theorem 1, we can decompose the pre-Dynkin system $\mathcal{D}$ into a set of blocks $\{A_i\}_{i \in I}$. Since, $A_i \subseteq 2^\Omega$ for all $i \in I$, each of the blocks induces a the linear subspace of simple gambles $S(\Omega, A_i) \subseteq B(\Omega)$. Given a finitely additive measure $\mu: \mathcal{D} \rightarrow [0, 1]$, we now define $E: \bigcup_{i \in I} S(\Omega, A_i) \rightarrow \mathbb{R}$ by

$$E(f) := \int f d\mu|_{A_i} \text{ if } f \in S(\Omega, A_i).$$

For every $i \in I$, $E|_{S(\Omega, A_i)}$ is a linear prevision in one-to-one correspondence to the finitely additive probability $\mu|_{A_i}$ [2] (Theorem 3.2.2). Hence, conditions (a) and (b) in Definition 8 are met [2] (Theorem 2.8.4). For $f \in S(\Omega, A_i) \cap S(\Omega, A_j)$ ($i \neq j, i, j \in I$), we know that $f \in S(\Omega, A_i \cap A_j)$ (Lemma A4), hence,

$$\int f d\mu|_{A_i} = \int f d\mu|_{A_i \cap A_j} = \int f d\mu|_{A_j}.$$

Thus, E is well-defined and there is no other partial expectation which agrees with μ on the indicator gambles of the sets in $\mathcal{D}$. $\square$

The attentive reader might have noticed that we defined the partial expectation in Proposition 12 on very specific linear subspaces of $B(\Omega)$, namely the linear subspaces of simple gambles. In fact, the statement would still hold when enlarging the linear subspaces of simple gambles $S(\Omega, A_i)$ for every $i \in I$ to linear subspaces of functions which are “convergence in measure”-approximated by gambles in $S(\Omega, A_i)$. For more details, we refer the reader to [26] (Definition 4.4.5 and Corollary 4.4.9).

But, it is not the case that we can extend the definition to all sets of bounded, A_i -measurable functions, i.e., bounded functions whose pre-images of sets in the smallest algebra which contains all open sets of the real numbers are contained in A_i . For an algebra $\mathcal{A} \subseteq 2^\Omega$, the set of $\mathcal{A}$ -measurable gambles is not necessarily a linear subspace of $B(\Omega)$ [2] (p. 129).

This is different for a σ -algebra $\mathcal{A}_\sigma \subseteq 2^\Omega$. The set of bounded, $\mathcal{A}_\sigma$ -measurable functions $B(\Omega, \mathcal{A}_\sigma)$ forms a linear subspace of $B(\Omega)$ [2] (p. 129). Here, measurability is defined as the pre-image of every Borel-measurable set in $\mathbb{R}$ is in $\mathcal{A}$. In this case, the set of linear

spaces on which the partial expectation is defined is given by all bounded, measurable functions on the σ -blocks.

Proposition 13 (Finitely Additive Probability on Dynkin Systems and its Partial Expectation). *Let $\mathcal{D}_\sigma \subseteq 2^\Omega$ be a Dynkin system on the base set Ω . Let $\mu: \mathcal{D}_\sigma \rightarrow [0, 1]$ be a finitely additive probability defined on the Dynkin system. Then μ is in one-to-one correspondence to a partial expectation $E: \bigcup_{i \in I} B(\Omega, \mathcal{A}_i) \rightarrow \mathbb{R}$ defined on the union of linear spaces of measurable gambles induced by all σ -blocks $\mathcal{A}_i$ of $\mathcal{D}_\sigma$.*

Proof. By Theorem A2, we can decompose the Dynkin system $\mathcal{D}_\sigma$ into a set of σ -blocks $\{\mathcal{A}_i\}_{i \in I}$. Since, $\mathcal{A}_i \subseteq 2^\Omega$ for all $i \in I$, each of the σ -blocks induce a linear subspace of $B(\Omega)$, which we denote as $B(\Omega, \mathcal{A}_i)$. Given a finitely additive measure $\mu: \mathcal{D}_\sigma \rightarrow [0, 1]$, we now define $E: \bigcup_{i \in I} B(\Omega, \mathcal{A}_i) \rightarrow \mathbb{R}$ by

$$E(f) := \int f d\mu|_{\mathcal{A}_i} \text{ if } f \in B(\Omega, \mathcal{A}_i).$$

For every $i \in I$, $E|_{B(\Omega, \mathcal{A}_i)}$ is a linear prevision in one-to-one correspondence to the finitely additive probability $\mu|_{\mathcal{A}_i}$ [2] (Theorem 3.2.2). Hence, conditions (a) and (b) in Definition 8 are met [2] (Theorem 2.8.4). For $f \in B(\Omega, \mathcal{A}_i) \cap B(\Omega, \mathcal{A}_j)$ ($i \neq j, i, j \in I$), we know that $f \in B(\Omega, \mathcal{A}_i \cap \mathcal{A}_j)$ (Lemma A5); hence,

$$\int f d\mu|_{\mathcal{A}_i} = \int f d\mu|_{\mathcal{A}_i \cap \mathcal{A}_j} = \int f d\mu|_{\mathcal{A}_j}.$$

Thus, E is well-defined and there is no other partial expectation which agrees with μ on the indicator gambles of the sets in $\mathcal{D}$. $\square$

It remains to emphasize that there are partial expectations defined on families of linear subspaces which are not induced by finitely additive probabilities on Dynkin systems. A simple example is given by a linear space which does not contain the constant gamble corresponding to the indicator gamble of the set Ω . Hence, the definition of a partial expectation is indeed a generalization of the definition of a finitely additive probability on a pre-Dynkin system.

Under the name “partially specified probabilities”, Lehrer [69] introduced a closely related notion to our partial expectation. Lehrer, however, assumed that there is by definition an underlying probability distribution over the entire base set (or better said, a σ -algebra on the base set). Hence, his partially specified probabilities are by definition extendable (see Definition 10), a fact, which he implicitly exploited by re-defining the natural extension following [2] (Lemma 3.1.3 (e)) of partially specified probabilities [70] (§3.2). Lehrer did not ask for the structure of the set of gambles with precise expectations, nor did he draw any connection to Walley’s work, nor did he link his “partially specified probabilities” to finitely additive probabilities on pre-Dynkin systems.

6.2. System of Precision—The Space of Gambles with Precise Expectations

In Section 3, we have shown that imprecise probabilities are precise on (pre-)Dynkin systems. The natural analogue of this *set structure of precision* is the *space of gambles with precise expectation*, which actually forms a linear subspace. The space of gambles with precise expectation is known as a set of ambiguity-free or unambiguous gambles in the finance and decision theory literature [71].

Theorem 7. (Imprecise Expectations Are Precise on a Linear Subspace of Precise Gambles) *Let $B(\Omega)$ be the linear space of bounded, real-valued functions on Ω . Let $L: B(\Omega) \rightarrow \mathbb{R}$ and $U: B(\Omega) \rightarrow \mathbb{R}$ be two functionals, for which all the following properties hold:*

- (a) Normalization: $L(\chi_\Omega) = U(\chi_\Omega) = 1$.
- (b) Conjugacy: $U(f) = -L(-f)$ for $f \in B(\Omega)$.

- (c) Subadditivity of U : for $f, g \in \mathcal{B}(\Omega)$, we have $U(f + g) \leq U(f) + U(g)$.
- (d) Superadditivity of L : for $f, g \in \mathcal{B}(\Omega)$, we have $L(f + g) \geq L(f) + L(g)$.
- (e) Positive Homogeneity: for $\alpha \in [0, \infty)$ and $f \in \mathcal{B}(\Omega)$, we have $L(\alpha f) = \alpha L(f)$ and $U(\alpha f) = \alpha U(f)$.

Then L and U coincide on the linear space $\mathcal{S} := \{f \in \mathcal{B}(\Omega) : L(f) = U(f)\} \subseteq \mathcal{B}(\Omega)$, the space of gambles with precise expectation, which contains all constant gambles.

Proof. We define

$$\mathcal{S} := \{f \in \mathcal{B}(\Omega) : L(f) = U(f)\}, \quad (3)$$

and show that $\mathcal{S}$ forms a linear subspace of $\mathcal{B}(\Omega)$. First, let $f, g \in \mathcal{S}$, then

$$L(f) + L(g) \stackrel{(d)}{\leq} L(f + g) \stackrel{(\star)}{\leq} U(f + g) \stackrel{(c)}{\leq} U(f) + U(g) \stackrel{\text{Eq.3}}{=} L(f) + L(g).$$

For $(\star)$ observe that $L(f) \leq U(f)$ for all $f \in \mathcal{B}(\Omega)$, since

$$L(f) + L(-f) \leq L(0) \stackrel{(a),(e)}{=} 0 \stackrel{(a),(e)}{=} U(0) \leq U(f) + U(-f),$$

we have

$$L(f) + L(-f) \leq U(f) + U(-f) \Leftrightarrow L(f) - U(f) \leq U(f) - L(f) \Leftrightarrow L(f) \leq U(f).$$

Second, let $f \in \mathcal{S}$ and $\alpha \in \mathbb{R}$. If $\alpha \geq 0$, then

$$L(\alpha f) \stackrel{(e)}{=} \alpha L(f) \stackrel{\text{Eq.3}}{=} \alpha U(f) \stackrel{(e)}{=} U(\alpha f).$$

Otherwise,

$$L(\alpha f) \stackrel{(b)}{=} -U(-\alpha f) \stackrel{(e)}{=} \alpha U(f) \stackrel{\text{Eq.3}}{=} \alpha L(f) \stackrel{(b)}{=} -\alpha U(-f) \stackrel{(e)}{=} U(\alpha f).$$

Third, $\mathcal{S}$ contains all constant gambles by (a), (b) and (e). Hence, we have shown that $\mathcal{S}$ forms a linear subspace of $\mathcal{B}(\Omega)$ which contains all constant gambles. $\square$

The choice of properties for the lower and upper expectation functional is not arbitrary. We tried to resemble the properties involved in the analogous statement for lower and upper probabilities (Theorem 2). One can easily check that a lower and upper expectation with the given properties (a)–(d) forms a lower and upper probability as required in Theorem 2 if the expectation is restricted to indicator gambles. However, we added property (e), positive homogeneity.

Without the property of positive homogeneity, the resulting set of gambles with precise expectations would not form a proper linear subspace, as then one can only guarantee closedness of $\mathcal{S}$ under rational multiplication. The condition of positive homogeneity “fills up” the gaps with all real-scaled functions. Instead of positive homogeneity one can as well demand a continuity assumption of the lower and upper functional L and U , e.g., [2] (Property (l) Theorem 2.6.1).

It is known that the system of precision of coherent lower and upper probability is a linear space [72]. The derivation here slightly generalizes this to not necessarily coherent lower and upper functionals L and U . We emphasize that coherent previsions (see Definition 9) fulfill all of the demanded properties [2] (Theorem 2.6.1).

In Theorem 2, we did not only derive the structure of the system of precision for events, we as well showed that the restriction of the lower and upper probability to this system give us a probability defined on a (pre-)Dynkin system. In the statement here, we can define a similar restriction $P: \mathcal{S} \rightarrow \mathbb{R}$ with $P := L|_{\mathcal{S}} = U|_{\mathcal{S}}$. However, P is not necessarily coherent and not necessarily a partial expectation. In other words, we cannot

directly recover a partial expectation on subdomain of $B(\Omega)$ induced by a lower and upper expectation functional. For probabilities, this recovery was possible.

But, we can sidestep this restriction. Any pair of lower and upper expectations, as we defined them here, induce a unique lower and upper probability. The resulting finitely additive probability on the set structure of precision gives rise to a partial expectation (Proposition 12) on a set of linear subspaces contained in the space of gambles with precise expectation S of L and U .

A converse construction, however, is not possible. There is no unique lower and upper expectation functional with the given properties associated with a lower and upper probability fulfilling the axioms of Theorem 2 [2] (§2.7.3). Concluding, lower and upper expectation as defined here are not the “perfect” analogues of lower and upper probabilities.

This as well explains the mismatch between systems of precision for probabilities and expectations. A lower and upper expectation fulfills the properties of a lower and upper probability but not vice versa. Hence, only weaker statements about the system of precision are possible for probabilities. As a result, the analogue of the set structure of precision, a pre-Dynkin system, is the space of gambles with precise expectations, a *single* linear subspace. In Definition 8, however, we equated pre-Dynkin systems with *sets* of linear subspaces. In this case, a one-to-one correspondence between a finitely additive probability on a pre-Dynkin system and a generalized expectation, concretely a partial expectation, can be established. Hence, the analogy of pre-Dynkin systems and linear subspaces of gambles depends on the correspondence of probability and expectation.

6.3. Generalized Extendability Is Equivalent to Coherence

Partial expectations are, as we have shown, a natural generalization of finitely additive probabilities on pre-Dynkin systems. Hence, it is not far-fetched to ask for definitions of coherence and extendability again, now in the more general context. It turns out that the same story can be re-told on a more general scale: The definition of coherent probabilities (Definition 5) is in fact just the reduction of the following definition of a coherent prevision to indicator gambles.

Definition 9 (Coherent Prevision [2] (Definition 2.5.1)). *Let $L \subseteq B(\Omega)$ be an arbitrary subset of the linear space of bounded functions. A functional $\underline{\mathbb{E}}: L \rightarrow \mathbb{R}$ is a coherent lower prevision if and only if*

$$\sup_{\omega \in \Omega} \sum_{i=1}^j (f_i(\omega) - \underline{\mathbb{E}}(f_i)) - m(f_0(\omega) - \underline{\mathbb{E}}(f_0)) \geq 0,$$

for non-negative $n, m \in \mathbb{N}$ and $f_0, f_1, \dots, f_n \in L$. If $L = -L$, the conjugate coherent upper prevision is given by $\overline{\mathbb{E}}(f) := -\underline{\mathbb{E}}(-f)$ for all $f \in L$. If, furthermore, $\underline{\mathbb{E}}(f) = \overline{\mathbb{E}}(f)$ for all $f \in L$, we call $\nu := \underline{\mathbb{E}} = \overline{\mathbb{E}}$ a linear prevision.

This definition of coherent previsions is substantiated by consistency of gamblers regarding their betting behavior on gambles with uncertain outcome, e.g., [2] (§2.3.1). Importantly, the rather opaque but general definition of coherence can be simplified greatly for coherent previsions defined on linear subspaces of $B(\Omega)$. Theorem 2.5.5 in [2] shows that coherence for lower previsions on linear subspaces can be expressed as superadditivity, positive homogeneity, and accepting sure gains (see [2] (Definition 2.3.3)).

Having introduced the notion of a coherent prevision, we now envisage the link between partial expectations and coherent previsions. We introduced extendability for finitely additive probabilities on pre-Dynkin systems as a useful property. It guarantees that the probability can “nicely” be embedded into “larger” finitely additive probability which is defined on an encompassing algebra. Hence, the analogue for partial expectations is straightforward.

Definition 10 (General Extendability). A partial expectation $E: \bigcup_{i \in I} L_i \rightarrow \mathbb{R}$ is extendable if and only if there exists a partial expectation $E': B(\Omega) \rightarrow \mathbb{R}$ such that $E'|_{\bigcup_{i \in I} L_i} = E$.

Interestingly, the extendability condition provided in Theorem 3 has a (more general) cousin adapted to the setting of gambles instead of events.

Proposition 14 (Extendability Condition for Previsions). (cf. [60] (Theorem 6.1).) Let $\{L_i\}_{i \in I}$ be a non-empty family of linear subspaces of $B(\Omega)$. A partial expectation $E: \bigcup_{i \in I} L_i \rightarrow \mathbb{R}$ is extendable if and only if for every finite collection of functions $f_1, \dots, f_n \in \{L_i\}_{i \in I}$,

$$\sum_{i=1}^n f_i \geq 0 \implies \sum_{i=1}^n E(f_i) \geq 0.$$

Proof. It seems that Theorem 6.1 [60] is equivalent to our statement. However, there is a subtlety which we want to argue here is indeed irrelevant. Extendability of a partial expectation requires the existence of a positive, *normed*, linear functional on $B(\Omega)$, whose restriction on the according linear subspaces coincides with the partial expectation. Theorem 6.1 in [60] only guarantees that a positive, linear functional exists. But, normedness of such functional is automatically given if $\chi_\Omega \in L_i$ for some $i \in I$. Otherwise, we extend the partial expectation E to $E': \{\alpha \chi_\Omega: \alpha \in \mathbb{R}\} \cup \bigcup_{i \in I} L_i \rightarrow \mathbb{R}$ such that

$$E'(f) := \begin{cases} \alpha & \text{if } f \in \{\alpha \chi_\Omega: \alpha \in \mathbb{R}\} \\ E(f) & \text{otherwise.} \end{cases}$$

Then again, Theorem 6.1 [60] applies. $\square$

Against the background that extendability and coherence define the same concept for finitely additive probabilities on pre-Dynkin systems, the resulting equivalence of extendability and coherence for partial expectations is of little surprise.

Proposition 15 (Extendability is Equivalent to Coherence). Let $\{L_i\}_{i \in I}$ be a non-empty family of linear subspaces of $B(\Omega)$. The partial expectation $E: \bigcup_{i \in I} L_i \rightarrow \mathbb{R}$ is extendable if and only if E is a linear prevision, i.e., is coherent.

Proof. If E is a linear prevision on $\bigcup_{i \in I} L_i$, then there is a linear prevision $E': B(\Omega) \rightarrow \mathbb{R}$ such that $E'|_{\bigcup_{i \in I} L_i} = E$ [2] (Theorem 3.4.2). Conversely, if E is an extendable partial expectation, then its extension is obviously a linear prevision, and hence it is coherent. The restriction of a coherent linear prevision to any subset of gambles is coherent (and linear). $\square$

6.4. A Duality Theory for Previsions and Families of Linear Subspaces

In Section 5, we step by step spelled out an order relationship between the set structure of precision and credal sets, a model for (coherent) imprecise probabilities. Naturally the presented generalization begs the question whether a related relationship between credal sets and the spaces of gambles with precise expectations exists. We answer affirmatively. We redefine the credal and dual credal set function and shortly discuss its analogous properties. Again, we require a “reference measure”. In this case, it is a fixed linear prevision ψ on the space of all gambles $B(\Omega)$, which is indeed in one-to-one correspondence to a finitely additive probability measure on 2^Ω .

Definition 11 (Generalized Credal Set Function). Let Δ be the set of linear previsions on the Banach space $B(\Omega)$. For a fixed linear prevision $\psi \in \Delta$, we call

$$m_\psi: 2^{B(\Omega)} \rightarrow 2^\Delta, \quad m_\psi(\mathcal{G}) := \{\nu \in \Delta: \nu(g) = \psi(g), \forall g \in \mathcal{G}\},$$

the generalized credal set function.

Definition 12 (Generalized Dual Credal Set Function). Let Δ be the set of linear previsions on the Banach space $B(\Omega)$. For a fixed linear prevision $\psi \in \Delta$, we call

$$m_\psi^\circ: 2^\Delta \rightarrow 2^{B(\Omega)}, \quad m_\psi^\circ(\mathcal{Q}) := \{g \in B(\Omega) : \nu(g) = \psi(g), \forall \nu \in \mathcal{Q}\},$$

the generalized dual credal set function.

Why can we call those functions “generalized”? Simply because any system of sets is equivalently represented as its set of indicator gambles which span their own linear space of simple gambles, i.e., linear combinations of indicator gambles.

The generalized credal set function maps, as the credal set function in Definition 6, to weak*-closed, convex subsets of Δ . The generalized dual credal set function, however, reveals a first subtlety. It maps to linear subspaces of $B(\Omega)$. The dual credal set function following Definition 7 mapped to pre-Dynkin systems. In Proposition 12 and Proposition 13, families of linear subspaces were the analogues of (pre-)Dynkin systems. Here, a single linear subspace is the analogue of a pre-Dynkin system. For a first step towards an explanation of this asymmetry, see Section 6.2. Finally, the pair of functions constitute a Galois connection.

Proposition 16 (Properties of Generalized (Dual) Credal Set Function). Let m_ψ be a generalized credal set function and m_ψ° be a generalized dual credal set function. All the following properties hold:

- (a) The generalized credal set function m_ψ maps to weak*-closed, convex sets.
- (b) The generalized dual credal set function m_ψ° maps to a linear subspace.
- (c) The generalized credal set function m_ψ and generalized dual credal set function m_ψ° form a Galois connection.

Proof. (a) We have fixed ψ to a linear prevision. Hence, it is coherent. For any $\mathcal{G} \subseteq B(\Omega)$, $m_\psi(\mathcal{G})$ is the set of all linear previsions which dominate ψ on $\mathcal{G}$. Theorem 3.6.1 in [2] then states that this set is weak*-closed and convex.

(b) Let $\mathcal{Q} \subseteq \Delta$.

Additivity Let $f, g \in m_\psi^\circ(\mathcal{Q})$. Then, for all $\nu \in \Delta$,

$$\nu(f + g) = \nu(f) + \nu(g) = \psi(f) + \psi(g) = \psi(f + g),$$

$$\text{i.e., } f + g \in m_\psi^\circ(\mathcal{Q}).$$

Homogeneity Let $f \in m_\psi^\circ(\mathcal{Q})$ and $\alpha \in \mathbb{R}$. Then, for all $\nu \in \Delta$,

$$\nu(\alpha f) = \alpha \nu(f) = \alpha \psi(f) = \psi(\alpha f),$$

i.e., $\alpha f \in m_\psi^\circ(\mathcal{Q})$. For homogeneity we need the easy fact that a linear prevision is not only positive homogeneous but generally homogeneous. For this consider a linear prevision ν and any gamble $f \in B(\Omega)$ with $\alpha < 0$, then $\nu(\alpha f) = -\nu(-\alpha f) = \alpha \nu(f)$.

- (c) The two functions constitute a Galois connection (cf. Proposition 4), $\mathcal{G} \subseteq m_\psi^\circ(\mathcal{Q}) \Leftrightarrow \mathcal{Q} \subseteq m_\psi(\mathcal{G})$. To this end, we show the left to right implication,

$$\nu \in \mathcal{Q} \Rightarrow \nu(g) = \psi(g), \forall g \in \mathcal{G} \Rightarrow \nu \in m_\psi(\mathcal{G}),$$

and the right to left implication,

$$g \in \mathcal{G} \Rightarrow \nu(g) = \psi(g), \forall \nu \in \mathcal{Q} \Rightarrow g \in m_\psi^\circ(\mathcal{Q}).$$

This concludes the proof. $\square$

Galois connections possess a series of helpful properties [64] (§V.7 and V.8). For instance, they give rise to a bipolar-closure operator. A non-empty subset $\mathcal{Q} \subseteq \Delta$ is bipolar-closed if and only if $\mathcal{Q} = m_\psi(m_\psi^\circ(\mathcal{Q}))$. A non-empty subset $\mathcal{G} \subseteq B(\Omega)$ is bipolar-closed if and only if $\mathcal{G} = m_\psi^\circ(m_\psi(\mathcal{G}))$. Furthermore, Proposition 16 provides necessary conditions for bipolar-closed sets. For instance, a bipolar-closed set $\mathcal{G} \subseteq B(\Omega)$ is a linear subspace. But is every such linear subspace a bipolar-closed set? No.

Example 16. Let $\Omega_2 := \{1, 2\}$, then $B(\Omega_2) = \{\alpha_1\chi_{\{1\}} + \alpha_2\chi_{\{2\}} : \alpha_1, \alpha_2 \in \mathbb{R}\}$. Hence, linear functionals on $B(\Omega_2)$ are defined via their behavior on the basis. Let $\psi(\chi_{\{1\}}) = 1$. Then, $\{\alpha_1\chi_{\{1\}} : \alpha_1 \in \mathbb{R}\} \subseteq B(\Omega_2)$ is a linear subspace, but it is, as one can easily check, not bipolar-closed, because the demand for normalization of any linear prevision ν which coincides with ψ on $\chi_{\{1\}}$ requires $\nu(\chi_{\{2\}}) = 0$.

Again, it seems to be more intricate than expected to characterize bipolar-closed sets. For pre-Dynkin systems and finitely additive measures we already collected some first hints that sets of measure zero play an important role in the characterization of bipolar-closed sets. In the case of linear subspaces and linear previsions, we observe a similar “combinatorial restriction”. In order to improve understanding, let us replace 2^Δ by $2^{\text{ba}(\Omega)}$ in Definition 11, Definition 12 and Proposition 16 (The proposition still holds.), which is equivalent to stating that linear previsions are not necessarily normalized, nor positive. Then, by leveraging the Hahn–Banach-type Theorem 1.5.14 in [26], one can easily see that linearity of a subset $\mathcal{G} \subseteq B(\Omega)$ is not only a necessary, but as well a sufficient condition for bipolar-closedness for those modified “credal set functions”. Thus, the restriction to actual linear previsions makes the characterization of bipolar-closed sets more complex. A compelling, more exhaustive answer still waits to be found.

Analogous to the discussion in Section 5.4, it is possible to provide a lattice duality and interpolation scheme via the generalized (dual) credal set functions. Instead of the lattice of pre-Dynkin systems $(\mathfrak{D}, \subseteq)$, the interpolation is directed by the lattice of linear subspaces $(\mathcal{L}, \subseteq)$ of $B(\Omega)$. As commonly known, the lattice of linear subspaces has the two operations $L_1 \wedge L_2 := L_1 \cap L_2$ and $L_1 \vee L_2 := \text{lin}(L_1 \cup L_2)$. We denote the linear span with lin . Its minimal element is the trivial zero vector linear subspace $\{0\}$. Its maximal element is the entire space of all gambles $B(\Omega)$. Due to higher generality of the here-presented (dual) credal set function, the interpolation provided is more fine-grained than for the previously given interpolation by pre-Dynkin systems. The following set containment (trivially) holds:

$$\{m_\psi(\mathcal{D}) : \mathcal{D} \in \mathfrak{D}\} \subseteq \{m_\psi(L) : L \in \mathcal{L}\},$$

where $m_\psi(\mathcal{D}) = m_\psi(\{\chi_D : D \in \mathcal{D}\})$. In other words, the lattice of pre-Dynkin systems is “contained” in the lattice of linear subspaces. However, as for pre-Dynkin systems the interpolation via linear subspaces is improper. The reason for this is again that not every linear subspace is bipolar-closed. By restriction to linear, bipolar-closed subspaces one can clean up the setup. For details we refer to Section 5.4. We do not make explicit the detailed reiteration of the same argument here.

In summary, we confirmed our findings of Section 5 extended to previsions. The generalized dual lattice setup underlines the structural consistency between the system of precision and its corresponding imprecise probability.

7. Conclusions and Open Questions

In this paper, we have explicated relations between the systems of precision and imprecise probabilities (respectively expectations). First, we have shown that the system of precision forms a pre-Dynkin system (respectively, a linear subspace). This structural insight raises a series of follow-up questions: How does the system of precision of a coherent prevision relate to the set of desirable gambles of this prevision? How does the

preference ordering change the set structure of precision for the corresponding beliefs? What is the role of coherence with respect to the system of precision?

Second, we defined finitely additive probabilities on pre-Dynkin systems. The equivalence of extendability and coherence of such probabilities strengthens the link between quantum probability and imprecise probability. We speculate that further insights can be obtained by exploiting this relationship. In addition, the generalization of finitely additive probabilities on pre-Dynkin systems to partial expectations directly opens the door to machine learning applications. In robust machine learning, the expected risk minimization framework is extended to more general expectation functionals. Partial expectation can, possibly after more computational investigations, deliver the desired robustness against dependencies in specific domains, such as privacy preservation, “not-missing-at-random” features, restricted data base access or multi-measurement data.

Finally, we developed a duality theory of systems of precision and imprecise probabilities (respectively expectations). A Galois connection defines a parametrized family of imprecise probabilities which follow an order structure provided by the lattice of pre-Dynkin systems (respectively the lattice of linear subspaces).

In modern statistics, especially in machine learning, probabilistic statements are increasingly tailored to individuals. Individual probabilistic statements, however, require justification. One can interpret probabilities on pre-Dynkin systems as probabilities which do not allow for such statements in the first place. One could perceive this fact as a weakness. We, in contrast, embrace its strength, when for ethical, legislative or other reasons, individualistic assignments are harmful, unjustifiable, forbidden or not desirable. We provide a first, rough interpolation scheme via the lattice duality. It demonstrates the space of adjustability of probabilistic assumptions in real-world scenarios. The involved pre-Dynkin systems are mathematical definitions of levels of group resolution. The question of how to choose such set-systems is related to the questions of intersectionality.

Several fundamental, technical questions remain open: how does the lattice duality imposed by pre-Dynkin systems or linear spaces relate to other dualities, such as convex duality, exploited in the field of imprecise probability. Can one easily characterize the bipolar-closed sets? Why is there no clear analogy between pre-Dynkin systems and linear subspaces?

We leave this collection of intriguing questions open to future work contributing to an understanding of the system of precision and the imprecise probability model.

Author Contributions: Conceptualization, R.D. and R.C.W.; Formal analysis, R.D.; Funding acquisition, R.C.W.; Investigation, R.D.; Methodology, R.D.; Project administration, R.C.W.; Supervision, R.C.W.; Validation, R.C.W.; Visualization, R.D.; Writing—original draft, R.D.; Writing—review & editing, R.C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC number 2064/1—Project number 390727645.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the International Max Planck Research School for Intelligent System (IMPRS-IS) for supporting Rabanus Derr. Many thanks to Christian Fröhlich for helpful discussions and feedback. We thank all anonymous reviewers on a first version for the detailed and helpful feedback. In particular, with their help, the statements and proofs of Theorem 4 and Theorem 5 have been cleaned and shortened substantially.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Lemmas and Proofs

Appendix A.1. Compatibility Structure

We have shown in Theorem 1 that every pre-Dynkin system can be decomposed into a union of algebras. A trivial follow-up question remains to be answered: is every union of algebras a pre-Dynkin system? The general answer is no.

Example A1 (Union of Algebras is not Necessarily Pre-Dynkin System). Let $\Omega = \{1, 2, 3\}$. Then, $\mathcal{A}_1 := \{\emptyset, 1, 23, \Omega\}$ and $\mathcal{A}_2 := \{\emptyset, 12, 3, \Omega\}$ are algebras on Ω . However, $\mathcal{A}_1 \cup \mathcal{A}_2$ do not form a pre-Dynkin system, as e.g., $1 \cup 3 = 13 \notin \mathcal{A}_1 \cup \mathcal{A}_2$.

However, we can give sufficient conditions under which the union over a set of algebras form a pre-Dynkin system. To this end, we have to introduce the so-called compatibility structure, which is made out of π -systems.

Definition A1 (π -System). Let Ω be an arbitrary base set. A π -system $\mathcal{I}$ is a non-empty subset of 2^Ω such that for arbitrary $A_1, \dots, A_n \in \mathcal{I}$, $\bigcap_{1 \leq i \leq n} A_i \in \mathcal{I}$.

Definition A2 (Compatibility Structure). A set of π -systems $\{\mathcal{I}_i\}_{i \in I}$ is called a compatibility structure of $\mathcal{A}$ if and only if the following two conditions hold:

- (a) $\bigcup_{i \in I} \mathcal{I}_i = \mathcal{A}$
- (b) For every π -system $\mathcal{I} \subseteq \mathcal{A}$ there exist $i \in I$ such that $\mathcal{I} \subseteq \mathcal{I}_i$.

A “compatibility structure” is a system of set systems. We call it compatibility structure, because every finite intersection of elements in a π -system is compatible with every other finite intersection of elements from the same π -system. The π -systems in the compatibility structure are not necessarily disjoint. Given a set of π -systems, the union of these π -systems is not necessarily a compatibility structure. In fact, there are possible incompatibilities between the π -systems (e.g., Example A1). It turns out that being a compatibility structure is a sufficient condition for a set of algebras to form a pre-Dynkin system.

Theorem A1 (Union of Algebras is Pre-Dynkin System if Compatibility Structure). Let $\{\mathcal{A}_i\}_{i \in I}$ be a family of algebras. If $\{\mathcal{A}_i\}_{i \in I}$ is a compatibility structure, then $\mathcal{A}_\cup := \bigcup_{i \in I} \mathcal{A}_i$ is a pre-Dynkin system.

Proof. We show that $\mathcal{A}_\cup$ is a pre-Dynkin system if it is the union of the compatibility structure. First, it contains the empty set. Second, if $A \in \mathcal{A}_\cup$ there is an i such that $A \in \mathcal{A}_i$. Then $A^c \in \mathcal{A}_i$, thus $A^c \in \mathcal{A}_\cup$. Third, let $\{A_j\}_{j \in [n]}$ be a subset of $\mathcal{A}_\cup$ with $n \in \mathbb{N}$ such that $A_k \neq \emptyset$, $A_l \neq \emptyset$ and $A_k \cap A_l = \emptyset$ for all $k \neq l$; $k, l \in [n]$. Thus, $\{A_j\}_{j \in [n]} \cup \{\emptyset\}$ is closed under finite intersection. Hence, it is a π -system. By definition of a compatibility structure, there is an $i \in I$ such that $\{A_j\}_{j \in [n]} \cup \{\emptyset\} \subseteq \mathcal{A}_i$. Since $\mathcal{A}_i$ is an algebra, it is closed under disjoint union. Thus, $\bigcup_{j \in [n]} A_j \in \mathcal{A}_i \subseteq \mathcal{A}_\cup$. $\square$

Appendix A.2. Supremum of a Chain of Algebras Is an Algebra

The following lemma is used to prove that every pre-Dynkin system can be dissected into algebras (Theorem 1).

Lemma A1 (Supremum of a Chain of Algebras is an Algebra). Let $(\{\mathcal{A}_i\}_{i \in I}, \subseteq)$ be a non-empty chain of algebras. Then it has an upper bounding algebra $\mathcal{A}_{\sup}$.

Proof. Let $\mathcal{A}_{\sup} = \bigcup_{i \in I} \mathcal{A}_i$. Trivially, $\mathcal{A}_i \subseteq \mathcal{A}_{\sup}$ for every $i \in I$. It remains to show that $\mathcal{A}_{\sup}$ is an algebra. It is non-empty by construction. Consider any finite subset

$\{A_1, \dots, A_n\} \subseteq \mathcal{A}_{\text{sup}}$. Without loss of generality, there exist $j \in I$ such that $\{A_1, \dots, A_n\} \subseteq \mathcal{A}_j$. Thus,

$$\bigcup_{i=1}^n A_i \in \mathcal{A}_j \subseteq \mathcal{A}_{\text{sup}}.$$

Furthermore, for every $A \in \mathcal{A}_{\text{sup}}$, there exist $j \in I$ such that $A \in \mathcal{A}_j$. Thus,

$$A^c \in \mathcal{A}_j \subseteq \mathcal{A}_{\text{sup}},$$

which concludes the proof. $\square$

Appendix A.3. Sufficient Conditions for Bipolar-Closed Sets

In order to give sufficient conditions for bipolar-closed sets we introduce the following notion.

Definition A3 (Weak Atom with Respect to Pre-Dynkin System). *A weak atom with respect to a pre-Dynkin system $\mathcal{D}$ on Ω is a set $A \subseteq \Omega$ such that every $B \in \mathcal{D}$ with $B \subseteq A$ is either $B = \emptyset$ or $B = A$.*

Weak atoms always exist, for instance the empty set is a weak atom with respect to any pre-Dynkin system $\mathcal{D}$ on any non-empty set Ω .

Lemma A2 (Decomposition into Pre-Dynkin Set and Atom). *Let $\mathcal{D} \subseteq 2^\Omega$ be a pre-Dynkin system. Every element $B \in 2^\Omega \setminus \mathcal{D}$ can be expressed as a disjoint union of an element $B_{\mathcal{D}} \in \mathcal{D}$ and a weak atom $A \in 2^\Omega \setminus \mathcal{D}$ with respect to $\mathcal{D}$.*

Proof. We define $B_{\mathcal{D}} \subseteq B$ as a maximal set $B_{\mathcal{D}} \in \mathcal{D}$ such that for all $H \in \mathcal{D}$ with $B_{\mathcal{D}} \subseteq H \subseteq B$ we have $H = B_{\mathcal{D}}$. By Zorn's lemma, such an element always exists. To see this, consider the poset of all subsets of B which are in $\mathcal{D}$ ordered by set inclusion. Then, we decompose $B \setminus B_{\mathcal{D}} = A$. Clearly, $A \neq \emptyset$. Furthermore, $A \notin \mathcal{D}$ and there is no $H \in \mathcal{D}$ such that $H \subseteq A$ except of the empty set, otherwise $B_{\mathcal{D}}$ would not have been maximal in our sense. Both statements follow from the closedness of $\mathcal{D}$ under disjoint union. Thus, A is a weak atom with respect to $\mathcal{D}$. $\square$

With these definitions at hand we can show the following crucial lemma.

Lemma A3 (Weak Atoms Are Not in the Bipolar-Closed Sets – Finite, Discrete Setting). *Let $\Omega = [n]$. We fix a finitely additive probability ψ on 2^Ω such that $\psi(F) > 0$ for every $F \in 2^\Omega \setminus \{\emptyset\}$. Let $\mathcal{D} \subseteq 2^\Omega$ be a pre-Dynkin system. Then, for every weak atom A with respect to $\mathcal{D}$ such that $A \notin \mathcal{D}$, we have*

$$A \notin m_\psi^\circ(m_\psi(\mathcal{D})),$$

where m_ψ and m_ψ° are defined following Definitions 6 and 7, respectively.

Proof. We show the Lemma by constructing, for every weak atom $A \subseteq \Omega$ with respect to $\mathcal{D}$ such that $A \notin \mathcal{D}$, a probability measure on 2^Ω which coincides with ψ on the pre-Dynkin system $\mathcal{D}$ but differs to ψ on the event A . Observe, weak atoms $A \notin \mathcal{D}$ exist as long as $\mathcal{D} \neq 2^\Omega$ (by Lemma A2). Second, the weak atom A contains an element $i \in A$ such that $\{i\} \notin \mathcal{D}$. Otherwise, $A \in \mathcal{D}$.

So, we fix an arbitrary weak atom A with respect to $\mathcal{D}$ such that $A \notin \mathcal{D}$. Due to the finiteness of $[n]$ we can represent every measure $\nu \in \Delta$ as $\nu_1, \dots, \nu_n$ with the constraints $0 \leq \nu_k \leq 1$ and $\sum_{k=1}^n \nu_k = 1$. With this in mind, we construct a probability $\nu \in \Delta$ such that $\nu(B) = \psi(B)$ for all $B \in \mathcal{D}$ but $\nu(A) \neq \psi(A)$.

To this end, we define ν on all elements in $[n]$. We choose

$$\nu_i = \psi_i + \epsilon$$

for the $i \in A$ specified earlier, with $\epsilon := \min_{k \in [n]} \psi_k$. By assumption, $\epsilon > 0$.

By Theorem 1 we know that we can decompose $\mathcal{D}$ into sub-algebras. In fact, we can decompose $\mathcal{D}$ into a finite set of sub-algebras as $\mathcal{D}$ is finite itself. Furthermore, set algebras on finite sets are build upon a finite set of atoms [26] (Remark 1.1.17.(2)). Those atoms form a partition of the finite set; hence, each set algebra on a finite set is in one-to-one correspondence to a partition (cf. [26] (Proposition 1.1.18)). Hence, $\mathcal{D}$ determines a family of partitions corresponding to its blocks. In detail, $\mathcal{D} = \bigcup_{o \in O} \mathcal{A}_o$ for set algebras $\mathcal{A}_o$ and finite O (Theorem 1). Then, we define $\{\mathcal{B}_o\}_{o \in O}$ as the corresponding set of partitions. In particular, $D(\mathcal{B}_o) = \mathcal{A}_o$. Thus, $\mathcal{D} = \bigcup_{o \in O} D(\mathcal{B}_o) = D(\bigcup_{o \in O} \mathcal{B}_o)$, because

$$\bigcup_{o \in O} D(\mathcal{B}_o) \subseteq D\left(\bigcup_{o \in O} \mathcal{B}_o\right) \subseteq D(\mathcal{D}) = \mathcal{D} = \bigcup_{o \in O} \mathcal{A}_o = \bigcup_{o \in O} D(\mathcal{B}_o).$$

For every partition $\mathcal{B}_o$ we have one $B_o \in \mathcal{B}_o$ such that $i \in B_o$. Obviously, $B_o \cap A \neq \emptyset$. But as well, $B_o \cap A^c \neq \emptyset$. Otherwise, $B_o \subseteq A$, thus $B_o = \emptyset$, because A is a weak atom which is not contained in the pre-Dynkin system $\mathcal{D}$. Thus, there is at least $j_o \in B_o$ with $j_o \notin A$. For one such j_o , we define the probability

$$\nu_{j_o} := \psi_{j_o} - \frac{\epsilon}{|\{j_o : o \in O\}|}.$$

By choice of ϵ and $1 \leq |\{j_o : o \in O\}| \leq \infty$ we have $\nu_{j_o} \geq 0$. Observe, we divide ϵ by the number of atoms on which we decrease the probability to guarantee normalization of ν . Finally, for all $l \in [n] \setminus (\{j_o : o \in O\} \cup A)$, we set $\nu_l = \psi_l$. This assignment leads to the following conclusions: $\nu(B_o) = \psi(B_o)$ for every $o \in O$. More generally, for all $B \in \bigcup_{o \in O} \mathcal{B}_o$, we have $\nu(B) = \psi(B)$.

The probability distribution ν is uniquely determined by the probability on the partitions, because $\mathcal{D} = D(\bigcup_{o \in O} \mathcal{B}_o)$ (see above) and the π - λ -Theorem [23] (Lemma A.1.3). So, $\nu \in m_\psi(\mathcal{D})$. But, $\nu(A) = \nu(\{i\}) + \nu(A \setminus \{i\}) = \psi(\{i\}) + \psi(A \setminus \{i\}) + \epsilon \neq \psi(A)$. Thus, $A \notin m_\psi^\circ(m_\psi(\mathcal{D}))$. $\square$

Appendix A.4. Intersection of Linear Subspaces

Lemma A4 (Intersection of Simple Gamble Spaces). *Let $\mathcal{A}, \mathcal{B} \subseteq 2^\Omega$ be two algebras. Then,*

$$S(\Omega, \mathcal{A}) \cap S(\Omega, \mathcal{B}) = S(\Omega, \mathcal{A} \cap \mathcal{B}).$$

Proof. It is clear that $S(\Omega, \mathcal{A} \cap \mathcal{B}) \subseteq S(\Omega, \mathcal{A})$ and $S(\Omega, \mathcal{A} \cap \mathcal{B}) \subseteq S(\Omega, \mathcal{B})$, which makes the right to left set inclusion obvious. For the left to right inclusion, observe that a function $f \in S(\Omega, \mathcal{A}) \cap S(\Omega, \mathcal{B})$ is a linear combination of indicator gambles of disjoint events $A_1, \dots, A_n \in \mathcal{A}$, $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and a linear combination of indicator gambles of disjoint events $B_1, \dots, B_m \in \mathcal{B}$, $\beta_1, \dots, \beta_m \in \mathbb{R}$. Without loss of generality, we can demand that all $\alpha_1, \dots, \alpha_n$ (respectively $\beta_1, \dots, \beta_m$) have to be pairwise distinct, because we can replace indicator gambles scaled with the same factor by the appropriately scaled indicator function of the disjoint union of the events. Then, there is $i \in [n]$ for every $j \in [m]$ such that $\alpha_i = \beta_j$ and vice versa. More importantly, there is $i \in [n]$ for every $j \in [m]$ such that $A_i = B_j$ and vice versa. Hence, $\{A_1, \dots, A_n\} = \{B_1, \dots, B_m\} \in \mathcal{A} \cap \mathcal{B}$. $\square$

Lemma A5 (Intersection of Measurable Gamble Spaces). *Let $\mathcal{A}_\sigma, \mathcal{B}_\sigma \subseteq 2^\Omega$ be two σ -algebras. Then,*

$$B(\Omega, \mathcal{A}_\sigma) \cap B(\Omega, \mathcal{B}_\sigma) = B(\Omega, \mathcal{A}_\sigma \cap \mathcal{B}_\sigma).$$

Proof. It is clear that $B(\Omega, \mathcal{A}_\sigma \cap \mathcal{B}_\sigma) \subseteq B(\Omega, \mathcal{A}_\sigma)$ and $B(\Omega, \mathcal{A}_\sigma \cap \mathcal{B}_\sigma) \subseteq B(\Omega, \mathcal{B}_\sigma)$, which makes the right to left set inclusion obvious. For the left to right inclusion, observe that for a function $f \in B(\Omega, \mathcal{A}_\sigma) \cap B(\Omega, \mathcal{B}_\sigma)$, the preimage of any Borel-measurable set in $\mathbb{R}$ is contained in $\mathcal{A}_\sigma$ and $\mathcal{B}_\sigma$. $\square$

Appendix B. Names of Pre-Dynkin Systems and Dynkin Systems

See Table A1 for a list of names for pre-Dynkin systems. Table A2 summarizes a list of names for Dynkin systems.

Table A1. A summary of names for pre-Dynkin systems found in the literature.

pre-Dynkin system [20]
additive-class [26] (p. 2)
concrete logic [46,73]
partial field [74]
quantum-mechanical algebra [75]
semi-algebra [30] (p. 13)
set-representable orthomodular poset [76]

Table A2. A summary of names for Dynkin systems found in the literature.

Dynkin system [77]
d-system [23] (p. 193)
λ -class [78] (p. 7)
quantum-mechanical σ -algebra [75]
σ -class [33]

Appendix C. Credal Sets of Pre-Dynkin System Probabilities—Credal Sets of Distorted Probabilities

To the best of the authors' knowledge, there has not been any attempt to parametrize a family of imprecise probabilities via the set of induced precise probabilities. In fact, a much better known class of imprecise probabilities is parametrized via distortion functions [67]. We use distorted probability functions as they regularly occur as examples of imprecise probabilities [2,67,68]. In particular, there is a one-to-one correspondence of distorted probabilities as defined in the following and so-called spectral risk measures, an important class of coherent upper previsions often used in economics and finance [68].

Definition A4 (Credal Set of Distorted Probability). Let $\gamma: [0, 1] \rightarrow [0, 1]$ be a concave, increasing function with $\gamma(0) = 0$ and $\gamma(1) = 1$. Let ψ denotes a finitely additive probability on an algebra 2^Ω on Ω . We overload the definition of a credal set

$$M(\psi, \gamma) := \{v \in \Delta: v(F) \leq \gamma(\psi(F)), \forall F \in 2^\Omega\}.$$

How does the credal set of probabilities for distorted probabilities relate to the credal set of probabilities for a probability defined on a Dynkin system?

Lemma A6 (Distortion Lemma). Let $\gamma: [0, 1] \rightarrow [0, 1]$ be a concave, increasing function with $\gamma(0) = 0$ and $\gamma(1) = 1$. If γ is not the identity function, then $\gamma(x) > x$ for all $x \in (0, 1)$.

Proof. As γ is concave, it is, in particular, quasi-concave. Thus, $x \mapsto \frac{\gamma(x)}{x}$ is decreasing [79] (Definition 10.1.1). This gives the following inequalities:

$$\frac{\gamma(x)}{x} \geq \frac{\gamma(x')}{x'} \geq \frac{\gamma(1)}{1} = 1,$$

for $0 < x \leq x'$. This implies, if there were $x \in (0, 1)$ such that $\gamma(x) = x$, then γ would be the identity function. We excluded this by assumption, and thus it follows $\gamma(x) > x$ for all $x \in (0, 1)$. $\square$

In the following Proposition, we show that the set of events on which all measures of a credal set $M(\psi, \gamma)$ coincide forms a pre-Dynkin system. Actually, it is the *system of certainty*, i.e., the set of all events which get assigned either the value 0 or the value 1. We reuse the notation of the dual credal set function m_ψ° which, as we noted earlier, maps a set of probabilities to the set structure on which those probabilities coincide.

Proposition A1 (Events of Precise Probability for Distorted Probabilities). *Let $\gamma: [0, 1] \rightarrow [0, 1]$ be a concave, increasing function with $\gamma(0) = 0$ and $\gamma(1) = 1$, which is not the identity function. Let ψ denote a finitely additive probability on an algebra 2^Ω on Ω . Let $M(\psi, \gamma)$ be the credal set (Definition A4) and m_ψ° the dual credal set function (Definition 7). Let $\mathcal{F}_0 := \{F \in 2^\Omega: \psi(F) = 0\}$ denote the set of all measure zero sets and $\mathcal{F}_{01} := \{F \in 2^\Omega: \psi(F) = 0 \text{ or } \psi(F) = 1\}$ denote the set of all measure zero or one sets. Then,*

$$m_\psi^\circ(M(\psi, \gamma)) = D(\mathcal{F}_0) = \mathcal{F}_{01}.$$

Proof. We show the equality of all sets via a circular set containment.

(a) $\mathcal{F}_{01} \subseteq D(\mathcal{F}_0)$

If $F \in 2^\Omega$ such that $\psi(F) = 0$, then clearly $F \in \mathcal{F}_0 \subseteq D(\mathcal{F}_0)$. If $F \in 2^\Omega$ such that $\psi(F) = 1$, then $\psi(F^c) = 0$. Thus, $F^c \in \mathcal{F}_0$ because pre-Dynkin systems are closed under complement $F \in D(\mathcal{F}_0)$.

(b) $D(\mathcal{F}_0) \subseteq m_\psi^\circ(M(\psi, \gamma))$

First, we write out

$$m_\psi^\circ(M(\psi, \gamma)) = \{F \in 2^\Omega: \nu(F) = \psi(F), \forall \nu \in \{\nu' \in \Delta: \nu'(F) \leq \gamma(\psi(F)), \forall F \in 2^\Omega\}\}$$

Since $\gamma(0) = 0$, it is easy to see that $\mathcal{F}_0 \subseteq m_\psi^\circ(M(\psi, \gamma))$. By Proposition 5, we know that $m_\psi^\circ(M(\psi, \gamma))$ is a pre-Dynkin system. Thus, $D(\mathcal{F}_0) \subseteq m_\psi^\circ(M(\psi, \gamma))$.

(c) $m_\psi^\circ(M(\psi, \gamma)) \subseteq \mathcal{F}_{01}$

We show this set inclusion via contraposition. If $F \in 2^\Omega$ has measure $\psi(F) \in (0, 1)$, then $F \notin m_\psi^\circ(M(\psi, \gamma))$. For this we have to argue that there is a measure $\nu_F \in M(\psi, \gamma)$ for every $F \in 2^\Omega$ with $\psi(F) \in (0, 1)$ such that $\nu_F(F) \neq \psi(F)$.

Observe that $\gamma \circ \psi$ defines a normalized, monotone, submodular set function on 2^Ω [43] (p. 17). Furthermore, any normalized, monotone, submodular set function induces a translation equivariant, monotone, positively homogeneous and sub-additive functional $L_{\gamma \circ \psi}$ on all $f \in B(\Omega)$ such that $L_{\gamma \circ \psi}(\chi_F) = (\gamma \circ \psi)(F)$ for all $F \in 2^\Omega$ [80] (p. 260), [2] (p. 130), [43] (Proposition 5.1, Theorem 6.3). Hence, $L_{\gamma \circ \psi}$ is a coherent upper prevision [2] (p. 65). Thus, Walley's extreme point theorem applies [2] (Theorem 3.6.2 (c)). (Even though this theorem is stated in terms of coherent lower previsions, it applies to coherent upper previsions, too. The weak*-compactness of the credal set $M(\psi, \gamma)$, which is given by the coherence of $L_{\gamma \circ \psi}$ and [2] (Theorem 3.6.1), is crucial.) For any function $f \in B(\Omega)$, in particular any χ_F with $F \in 2^\Omega$, there is a linear prevision $f \mapsto \langle f, \nu \rangle$ with $\nu \in \Delta$ on $B(\Omega)$ dominated by $L_{\gamma \circ \psi}$ such that $\langle \chi_F, \nu \rangle = L_{\gamma \circ \psi}(\chi_F)$. More concretely, for any $F \in 2^\Omega$ there is a $\nu_F \in M(\psi, \gamma)$ such that $\nu_F(F) = (\gamma \circ \psi)(F)$. If $\psi(F) \in (0, 1)$, then Lemma A6 applies and $\nu_F(F) = (\gamma \circ \psi)(F) = \gamma(\psi(F)) > \psi(F)$ gives the desired inequality. In conclusion, there is no $F \in 2^\Omega$ with measure $\psi(F) \in (0, 1)$ such that $F \in m_\psi^\circ(M(\psi, \gamma))$. This implication finalizes the proof.

$\square$

We call the set $\mathcal{F}_{01}$ a *system of certainty* for ψ . Hence, the system of precision of a distorted probability is the system of certainty. Since the proposition clarifies the relation

from distorted imprecise probability to probability on Dynkin systems, the reverse question immediately follows: when is $M(\psi|_{\mathcal{D}}, \mathcal{D}) \subseteq M(\psi, \gamma)$? In other words, given an extendable probability defined on a pre-Dynkin system, what is a distortion function of an extension of this probability such that the credal set of the former is contained in the credal set of the latter. The simple example below gives an instantiation of this problem for which it is easy to find a solution. However, the problem is harder for more general cases.

Example A2. Let $\Omega = \{1, 2\}$ and $\psi(1) = \psi(2) = 0.5$. Let $\mathcal{D} = \{\emptyset, \Omega\}$. Then, for the admittedly extreme distortion function $\gamma(0) = 0$, otherwise $\gamma(a) = 1$, for $a \in (0, 1]$ the credal set $M(\psi, \gamma) = \Delta$ is the set of all probability measures. Hence, $M(\psi|_{\mathcal{D}}, \mathcal{D}) = \Delta \subseteq M(\psi, \gamma)$.

Appendix D. Dynkin Systems and Countably Additive Probability

In the main text, we presented the majority of the results for the general case: pre-Dynkin systems and finitely additive probabilities. We did so, as we emphasize clearly here, not for the sake of mathematical generality. There are probabilistic problems which demand for the use of finitely additive probabilities, e.g., von Mises frequentistic notion of probability [20,29]. As pre-Dynkin systems and finitely additive probabilities walk hand in hand, so too do Dynkin systems and countably additive probabilities. It is possible to strengthen some results when we assume Dynkin systems and countably additive probabilities. Furthermore, countably additive probabilities are more familiar to students of probability theory. Finitely additive probabilities still eke out an exotic living [26,81,82]. Nevertheless, they are central in large parts of literature on imprecise probability [2,21].

Appendix D.1. Dynkin Systems

Many of the following results are analogous to the statements for pre-Dynkin systems. The subsequent theorem is analogous to Theorem 1. We remark, however, that we cannot use the same proof-technique as in Theorem 1 because the generalization of Lemma A1 to σ -algebras does not hold. Hence, we use a different technique to prove Theorem A2 for Dynkin systems and σ -algebras.

Theorem A2 (Dynkin Systems are made out of σ -Algebras). *Let $\mathcal{D}_\sigma$ be a Dynkin system on an arbitrary set Ω . Then, there is a unique family of maximal σ -algebras $\{\mathcal{A}_i\}_{i \in I}$ such that $\mathcal{D}_\sigma = \bigcup_{i \in I} \mathcal{A}_i$. We call these σ -algebras the σ -blocks of $\mathcal{D}_\sigma$.*

Proof. Since Dynkin systems are pre-Dynkin systems, Theorem 1 guarantees that $\mathcal{D}_\sigma$ is constituted of a set of algebras $\{\mathcal{A}_i\}_{i \in I}$. Each algebra $\mathcal{A}_i$ is closed under finite intersection. Thus, it is a “compatible collection” following the terms of Gudder [35]. It follows that $\mathcal{A}_i$ is contained in a sub- σ -algebra of $\mathcal{D}_\sigma$ by Theorem 2.1 in [35]. As any sub- σ -algebra is an algebra and $\mathcal{A}_i$ is maximal, i.e., there is no algebra contained in $\mathcal{D}$ such that $\mathcal{A}_i$ is a strict sub-algebra of this algebra, $\mathcal{A}_i$ itself is a σ -algebra. Hence, $\mathcal{A}_i$ are the σ -blocks of $\mathcal{D}_\sigma$. $\square$

Not every union of σ -algebras is a Dynkin system (cf. Appendix A.1). But, if a union of σ -algebras forms a compatibility structure, then it is a Dynkin system.

Theorem A3 (Union of σ -Algebras is Dynkin system if Compatibility Structure). *Let $\{\mathcal{A}_i\}_{i \in I}$ be a family of σ -algebras on an arbitrary Ω . If $\{\mathcal{A}_i\}_{i \in I}$ is a compatibility structure, then $\mathcal{A}_\cup := \bigcup_{i \in I} \mathcal{A}_i$ is a Dynkin system.*

Proof. We show that $\mathcal{A}_\cup$ is a Dynkin system if it is the union of the compatibility structure. First, it contains the empty set. Second, if $A \in \mathcal{A}_\cup$ there is an i such that $A \in \mathcal{A}_i$. Then $A^c \in \mathcal{A}_i$, thus $A^c \in \mathcal{A}_\cup$. Third, let $\{A_j\}_{j \in J}$ be a subset of $\mathcal{A}_\cup$ with $J \subseteq \mathbb{N}$ such that $A_k \neq \emptyset$, $A_l \neq \emptyset$ and $A_k \cap A_l = \emptyset$ for all $k \neq l; k, l \in J$. Thus, $\{A_j\}_{j \in J} \cup \{\emptyset\}$ is closed under finite intersection. It is thus a π -system. By definition of a compatibility structure there is an $i \in I$ such that $\{A_j\}_{j \in J} \cup \{\emptyset\} \subseteq \mathcal{A}_i$. Since $\mathcal{A}_i$ is a σ -algebra, it is closed under countable disjoint union. Thus, $\bigcup_{j \in J} A_j \in \mathcal{A}_i \subseteq \mathcal{A}_\cup$. $\square$

Appendix D.2. Technical Setup

In order to work on firm ground when introducing countably additive probabilities, we change our basic technical setup. We summarize the used notations in Table A3.

Table A3. Summary of used notations in Appendix D.

Ω	Polish Space
$\mathcal{D}_\sigma$	Dynkin system on Ω (Definition 1)
μ_σ	Countably additive probability defined on $\mathcal{D}_\sigma$ (Definition 3)
$\sigma(\mathcal{A})$	σ -algebra hull of set system $\mathcal{A} \subseteq 2^\Omega$
$\mathcal{F}_\sigma$	Borel- σ -algebra on Ω
$\text{ca}(\Omega, \mathcal{F}_\sigma)$	Set of bounded, signed, countably additive measures on $\mathcal{F}_\sigma$
$\Delta_\sigma(\Omega, \mathcal{F}_\sigma)$ respectively Δ_σ	Set of countably additive probability measures on $\mathcal{F}_\sigma$
$M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)$	σ -Credal set of μ_σ on $\mathcal{D}_\sigma$ (Proposition A3)
$\underline{\mu}_{\mathcal{D}_\sigma}, \bar{\mu}_{\mathcal{D}_\sigma}$	Lower respectively upper coherent σ -extension (Proposition A4)

First, we assume now that Ω is a Polish space, that is, a separable completely metrizable topological space [80] (p. 20). We denote by $\mathcal{F}_\sigma$ the Borel- σ -algebra with respect to the given topology. With $\text{ca}(\Omega, \mathcal{F}_\sigma)$, we denote the space of all finite, signed countably additive measures on $(\Omega, \mathcal{F}_\sigma)$ (cf. [80] (p. 20)).

We define $\Delta_\sigma \subseteq \text{ca}(\Omega, \mathcal{F}_\sigma)$ as the set of all probability measures on $(\Omega, \mathcal{F}_\sigma)$. Every measure in Δ_σ is regular [80] (p. 20). The set $\text{ca}(\Omega, \mathcal{F}_\sigma)$ is equipped with the total variation distance $d_{TV}(\nu, \nu') := \sup_{A \in \mathcal{F}_\sigma} |\nu(A) - \nu'(A)|$ [83] (p. 366). We emphasize that this is a different topology in comparison to the topology used above. The weak* topology is weaker than the topology induced by the total variation distance.

Lastly, we denote the smallest σ -algebra which contains $\mathcal{A}$ with $\sigma(\mathcal{A})$. We say that $\sigma(\mathcal{A})$ is the σ -algebra hull of $\mathcal{A}$.

Appendix D.3. Dynkin Probability Spaces

The definition of probability measures on pre-Dynkin systems directly applies to Dynkin systems, too. Analogous to Kolmogorov's probability space, we can now leverage Definition 3 for a definition of a Dynkin probability space.

Definition A5 (Dynkin Probability Space [19] (p. 296)). *The triple $(\Omega, \mathcal{D}_\sigma, \mu_\sigma)$ is called a Dynkin probability space if and only if (a) Ω is a non-empty base space, (b) $\mathcal{D}_\sigma$ is a Dynkin system on Ω and (c) μ_σ is a countably additive probability measure on $\mathcal{D}_\sigma$ following Definition 3.*

Dynkin probability spaces have been called “quantum probability spaces” in quantum probability theory [19]. Dynkin probability spaces generalize Kolmogorov's probability spaces. If $\mathcal{D}_\sigma$ were a σ -algebra, then the Dynkin probability space would become a classical probability space. Theorem A2 provides another interesting link between Dynkin and classical probability spaces.

Proposition A2. *Every Dynkin probability space $(\Omega, \mathcal{D}_\sigma, \mu_\sigma)$ defines a collection of classical probability space $\{(\Omega, \mathcal{A}_i, \mu_i)\}_{i \in I}$ where $\mathcal{D}_\sigma = \bigcup_{i \in I} \mathcal{A}_i$ and $\mu_i := \mu_\sigma|_{\mathcal{A}_i}$ are consistent, i.e., $\mu_i(A) = \mu_j(A)$ for any $A \in \mathcal{A}_i \cap \mathcal{A}_j$ and $i, j \in I$.*

Such “multi-Kolmogorov” probability spaces have similarly been formalized in [37] (p. 32) or [34] (p. 154). The multiplicity of probability spaces has been interpreted as a collection of contexts. Each context possesses its own classical probability space. (In quantum physics, we obtain classical behavior in single context and quantum behavior across context. In machine learning, each marginal scenario gets equipped an own probability space, i.e., an own context (cf. [3,32]).) Again, as for pre-Dynkin systems and finitely additive prob-

abilities, the question of embedding Dynkin probability spaces into classical probability spaces arises.

Appendix D.4. Conditions for Extendability for Countably Additive Probabilities

Definition A6 (σ -Extendability). Let $\mathcal{F}_\sigma$ be the Borel- σ -algebra on a Polish space Ω and $\mathcal{D}_\sigma$ a Dynkin system contained in this σ -algebra. With Δ_σ we denote the set of all countably additive probability measures on $(\Omega, \mathcal{F}_\sigma)$. We call a countably additive probability μ_σ on $\mathcal{D}_\sigma$ σ -extendable if and only if there is a countably additive probability measure $\nu \in \Delta_\sigma$ such that $\nu|_{\mathcal{D}_\sigma} = \mu_\sigma$.

We call this extendability “ σ ” to emphasize the difference to Definition 4 and the properties of countably additive probability measures. Extendability of Dynkin probability spaces, as well as for finitely additive probabilities on pre-Dynkin systems, has already been part of discussions in quantum probability since 1969 [19] up to more current times [47]. Most of the results, e.g., [33,46,47], are only stated in terms of finitely additive probability measure and/or apply only on a structurally restricted class of Dynkin systems. A theorem due to Maharam [60] in the literature on marginal probabilities can be tweaked to give a universal sufficient and necessary criterion for σ -extendability.

Theorem A4 (σ -Extendability Condition for Countably Additive Probability). Let $\mathcal{F}_\sigma$ be the Borel- σ -algebra on a Polish space Ω and $\mathcal{D}_\sigma$ a Dynkin system contained in this σ -algebra such that $\sigma(\mathcal{D}_\sigma) = \mathcal{F}_\sigma$. Let μ_σ be a countably additive probability on $\mathcal{D}_\sigma$. For each σ -block $\mathcal{A}_i$ of $\mathcal{D}_\sigma$ we assume that $\mu_i := \mu_\sigma|_{\mathcal{A}_i}$ is inner regular, i.e., if $\mu(G) = \sup\{\mu(K) : K \subseteq G, K \text{ compact and measurable}\}$ [25] (p. 808). The probability μ_σ is σ -extendable if and only if

$$\sum_{j=1}^n f_j(\omega) \geq 0, \forall \omega \in \Omega \implies \sum_{j=1}^n \int_{\Omega} f_j(\omega) d\mu_{i_j}(\omega) \geq 0$$

for all finite families of measurable simple gambles $\{f_j\}_{j \in [n]} \subseteq S(\Omega, \mathcal{D}_\sigma)$.

Proof. Proposition A2 separates the Dynkin probability space into a set of Kolmogorov probability spaces. Then, Theorem 8.1 in [60] applies. As technical requirements, we emphasize that Ω is a Hausdorff space and all restrictions μ_i are inner regular. Thus, there exists a probability measure on the algebra generated by $\mathcal{D}_\sigma$. Since the probability measure is countably additive, Caratheodory’s Extension Theorem [23] (Theorem 1.7) states that it can be uniquely extended to the Borel- σ -algebra $\mathcal{F}_\sigma = \sigma(\mathcal{D}_\sigma)$. $\square$

Similar and related results can be found in [34,48,58]. In particular, every so-called marginal scenario can be represented as a countably additive probability on a Dynkin system (cf. [34], [33] (Example 4.2)). Marginal scenarios often arise in practical setups. They are defined as settings in which for a collection of random variables only specific partial joint distributions of the random variables are given. For those scenarios there are purely combinatorial conditions on the Dynkin system sufficient for the extendability [34,58].

Appendix D.5. Credal Set of Countably Additive Probabilities on Dynkin Systems

The credal set of probabilities which we defined in Corollary 1 contains finitely additive probabilities, some of which might not be countably additive. For this reason we redefine the credal set for countably additive probabilities. The credal set is the set of all countably additive probabilities which coincide with the reference probability μ_σ on $\mathcal{D}_\sigma$.

Proposition A3 (σ -Credal Set for Probabilities on Dynkin Systems). Let $\mathcal{F}_\sigma$ be the Borel- σ -algebra on a Polish space Ω and $\mathcal{D}_\sigma$ a Dynkin system contained in this σ -algebra. Let μ_σ be a

countably additive probability on $\mathcal{D}_\sigma$ and Δ_σ the set of all countably additive probabilities on $\mathcal{F}_\sigma$. We call

$$M_\sigma(\mu_\sigma, \mathcal{D}_\sigma) := \{v \in \Delta_\sigma : v(A) = \mu_\sigma(A), \forall A \in \mathcal{D}_\sigma\},$$

the σ -credal set of μ_σ on $\mathcal{D}_\sigma$. If $M_\sigma(\mu_\sigma, \mathcal{D}_\sigma) \neq \emptyset$, then it is closed and convex.

Proof. Convexity Let $\{v_i\}_{i \in [n]} \subseteq M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)$ and $\alpha_i \in [0, 1]$ for all $i \in [n]$ with $\sum_{i=1}^n \alpha_i = 1$, then $\sum_{i=1}^n \alpha_i v_i \in M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)$, because

$$\sum_{i=1}^n \alpha_i v_i(A) = \sum_{i=1}^n \alpha_i \mu_\sigma(A) = \mu_\sigma(A), \quad \forall A \in \mathcal{D}_\sigma.$$

Closedness We assumed $\text{ca}(\Omega, \mathcal{F}_\sigma)$ to be equipped with the total variation distance. Hence, of a set $Q \subseteq \text{ca}(\Omega, \mathcal{F}_\sigma)$ can be identified via the convergence of sequences in Q [84] (Lemma 21.2).

Suppose $v_1, v_2, \dots \in M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)$ is a sequence of probability measures such that $\lim_{n \rightarrow \infty} v_n \stackrel{TV}{=} v$. This directly implies set-wise convergence,

$$v(A) = \lim_{n \rightarrow \infty} v_n(A) = \lim_{n \rightarrow \infty} \mu_\sigma(A) = \mu_\sigma(A), \quad \forall A \in \mathcal{D}_\sigma.$$

It follows that $v \in M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)$.

□

In fact, one can even guarantee weak closure of the σ -credal set, since it is convex [85] (Theorem II.A.4). Rather obviously, the following corollary holds.

Corollary A1 (σ -Extendability and σ -Credal Set). *Let $\mathcal{F}_\sigma$ be the Borel- σ -algebra on a Polish space Ω and $\mathcal{D}_\sigma$ a Dynkin system contained in this σ -algebra. The countably additive probability μ_σ on $\mathcal{D}_\sigma$ is σ -extendable if and only if $M_\sigma(\mu_\sigma, \mathcal{D}_\sigma) \neq \emptyset$.*

Thus, we can redefine the lower and upper coherent extension if the Dynkin probability space is continuously extendable. This extension is then derived from a set of countably additive probability measures.

Proposition A4 (Coherent σ -Extension of Probability). *Let $\mathcal{F}_\sigma$ be the Borel- σ -algebra on a Polish space Ω and $\mathcal{D}_\sigma$ a Dynkin system contained in this σ -algebra. Assume the countably additive probability μ_σ on $\mathcal{D}_\sigma$ is σ -extendable. Then,*

$$\underline{\mu}_{\mathcal{D}_\sigma}(A) := \inf_{v \in M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)} v(A), \quad \bar{\mu}_{\mathcal{D}_\sigma}(A) := \sup_{v \in M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)} v(A), \quad \forall A \in \mathcal{F}_\sigma,$$

define a coherent lower (respectively, upper) probability on $\mathcal{F}_\sigma$ in the sense of [2].

Proof. First, we notice the every countably additive probability is also finitely additive. Via Theorem 4 and application of Theorem 3.3.4 (b) in [2], we directly obtain the result. □

The coherent extension is one of two methods of measure extension presented in this paper. Again, we ask how the coherent σ -extension relates to the inner and outer measure construction.

Corollary A2 (Extension Theorem—Countably Additive Case). *Let $\mathcal{F}_\sigma$ be the Borel- σ -algebra on a Polish space Ω and $\mathcal{D}_\sigma$ a Dynkin system contained in this σ -algebra. Suppose the countably additive probability μ_σ on $\mathcal{D}_\sigma$ is σ -extendable. Then,*

$$\mu_*(A) \leq \underline{\mu}_{\mathcal{D}_\sigma}(A) \leq \bar{\mu}_{\mathcal{D}_\sigma}(A) \leq \mu^*(A), \quad \forall A \in \mathcal{F}_\sigma.$$

Proof. Since $M_\sigma(\mu_\sigma, \mathcal{D}_\sigma) \subseteq M(\mu_\sigma, \mathcal{D}_\sigma)$, clearly

$$\inf_{\nu \in M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)} \nu(A) \geq \inf_{\nu \in M(\mu_\sigma, \mathcal{D}_\sigma)} \nu(A), \quad \forall A \in \mathcal{F}_\sigma,$$

$$\sup_{\nu \in M_\sigma(\mu_\sigma, \mathcal{D}_\sigma)} \nu(A) \leq \sup_{\nu \in M(\mu_\sigma, \mathcal{D}_\sigma)} \nu(A), \quad \forall A \in \mathcal{F}_\sigma.$$

Hence, the analogous Theorem 5 gives the result. $\square$

Appendix E. From Set Systems to Logical Structures and Back

In the year 1936, Stone demonstrated a celebrated representation result of logical algebras [86]. In particular, he showed that any Boolean algebra can be equivalently represented by an algebra of sets and vice versa. This representation results opened the door to many generalizations of the domain of probabilities (see [36] for a nice summary). For instance, Boolean algebras can be replaced by weaker logical structures, such as orthomodular lattices. Interestingly, modular ortholattices and orthomodular lattices have been investigated as representational structures for quantum theoretic descriptions and measurements [87,88]. Furthermore, it turned out that an analogue to Stone’s representation holds for some orthomodular lattices. Some orthomodular lattices are isomorphic to pre-Dynkin systems [74]. On the other hand, every (pre-)Dynkin system is isomorphic to an (σ -)orthomodular poset, a order-theoretic generalization of an orthomodular lattice (cf. [41]). In summary, the investigation of probabilities on general logical structures parallels our work presented here. We, however, have stuck to set structures as the domain of probabilities.

References

1. Kolmogorov, A.N. The general theory of measure and probability calculus. In *Collected Works of the Mathematical Section, Communist Academy, Section for Natural and Exact Sciences*; Springer, Berlin/Heidelberg, Germany, 1929; Volume 1, pp. 8–21. (In Russian). Translated to English in A.N. Shirayev (Editor), *Selected Works of A.N. Kolmogorov, Volume II Probability and Mathematical Statistics*, pages 48–59. 1992.
2. Walley, P. *Statistical Reasoning with Imprecise Probabilities*; Chapman and Hall: London, UK, 1991.
3. Eban, E.; Mezuman, E.; Globerson, A. Discrete Chebyshev classifiers. In *Proceedings of the International Conference on Machine Learning*. PMLR, Beijing, China, 21–26 June 2014; pp. 1233–1241.
4. Gudder, S.P. *Stochastic Methods in Quantum Mechanics*; North Holland: Amsterdam, The Netherlands, 1979.
5. Epstein, L.G.; Zhang, J. Subjective probabilities on subjectively unambiguous events. *Econometrica* **2001**, *69*, 265–306. [\[CrossRef\]](#)
6. Zhang, J. Subjective ambiguity, expected utility and Choquet expected utility. *Econ. Theory* **2002**, *20*, 159–181. [\[CrossRef\]](#)
7. Cole, E.R. Intersectionality and research in psychology. *Am. Psychol.* **2009**, *64*, 170. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Shields, S.A. Gender: an intersectionality perspective. *Sex Roles* **2008**, *59*, 309–311. [\[CrossRef\]](#)
9. Weldon, S.L. Intersectionality. In *Politics, Gender, and Concepts: Theory and Methodology*; Goertz, G., Mazur, A.G., Eds.; Cambridge University Press: Cambridge, UK, 2008; pp. 193–218.
10. Anonymous. Hypergraph. *Wikipedia* **2023**. Available online: <https://en.wikipedia.org/wiki/Hypergraph> (accessed on 4 August 2023).
11. Berge, C. *Hypergraphs: Combinatorics of Finite Sets*; North-Holland: Amsterdam, The Netherlands, 1989.
12. Anonymous. Intersectionality. *Wikipedia* **2023**. Available online: <https://en.wikipedia.org/wiki/Intersectionality> (accessed on 4 August 2023).
13. Hancock, A.M. Empirical intersectionality: A tale of two approaches. *UC Irvine Law Rev.* **2013**, *3*, 259–296.
14. Hancock, A.M. When multiplication doesn’t equal quick addition: Examining intersectionality as a research paradigm. *Perspect. Politics* **2007**, *5*, 63–79. [\[CrossRef\]](#)
15. Cooke, R. The anatomy of the squizzel: The role of operational definitions in representing uncertainty. *Reliab. Eng. Syst. Saf.* **2004**, *85*, 313–319. [\[CrossRef\]](#)
16. Elstrodt, J. *Maß- und Integrationstheorie*; Springer: Berlin/Heidelberg, Germany, 2018.
17. Vitali, G. *Sul problema della misura dei Gruppi di punti di una retta: Nota*; Tipografia Gamberini e Parmeggiani: Bologna, Italy, 1905.
18. Isaacs, Y.; Hájek, A.; Hawthorne, J. Non-measurability, imprecise credences, and imprecise chances. *Mind* **2022**, *131*, 892–916. [\[CrossRef\]](#)
19. Gudder, S.P. Quantum probability spaces. *Proc. Am. Math. Soc.* **1969**, *21*, 296–302. [\[CrossRef\]](#)
20. Schurz, G.; Leitgeb, H. Finitistic and frequentistic approximation of probability measures with or without σ -additivity. *Stud. Log.* **2008**, *89*, 257–283. [\[CrossRef\]](#)

21. Augustin, T.; Coolen, F.P.; De Cooman, G.; Troffaes, M.C. *Introduction to Imprecise Probabilities*; John Wiley & Sons: New York, NY, USA, 2014.
22. Derr, R.; Williamson, R.C. The set structure of precision. In Proceedings of the International Symposium on Imprecise Probability: Theories and Applications, Oviedo, Spain, 11–14 July 2023.
23. Williams, D. *Probability with Martingales*; Cambridge University Press: Cambridge, UK, 1991.
24. Hildebrandt, T.H. On bounded linear functional operations. *Trans. Am. Math. Soc.* **1934**, *36*, 868–875. [\[CrossRef\]](#)
25. Schechter, E. *Handbook of Analysis and Its Foundations*; Academic Press: Cambridge, MA, USA, 1997.
26. Rao, K.B.; Rao, M.B. *Theory of Charges: A Study of Finitely Additive Measures*; Academic Press: Cambridge, MA, USA, 1983.
27. Pták, P. Concrete quantum logics. *Int. J. Theor. Phys.* **2000**, *39*, 827–837. [\[CrossRef\]](#)
28. von Mises, R. Grundlagen der Wahrscheinlichkeitsrechnung. *Math. Z.* **1919**, *5*, 52–99. [\[CrossRef\]](#)
29. von Mises, R.; Geiringer, H. *Mathematical Theory of Probability and Statistics*; Academic Press: Cambridge, MA, USA, 1964.
30. Khrennikov, A.Y. *Interpretations of Probability*; Second Revised and Extended Edition; Walter de Gruyter: Berlin, Germany, 2009.
31. Frohlich, C.; Derr, R.; Williamson, R.C. Towards a strictly frequentist theory of imprecise probability. In Proceedings of the International Symposium on Imprecise Probability: Theories and Applications, Oviedo, Spain, 11–14 July 2023.
32. Cuadras, C.M.; Fortiana, J.; Rodriguez-Lallena, J.A. *Distributions with Given Marginals and Statistical Modelling*; Springer: Berlin/Heidelberg, Germany, 2002.
33. Gudder, S.P. An extension of classical measure theory. *SIAM Rev.* **1984**, *26*, 71–89. [\[CrossRef\]](#)
34. Vorob'ev, N.N. Consistent families of measures and their extensions. *Theory Probab. Appl.* **1962**, *7*, 147–163. [\[CrossRef\]](#)
35. Gudder, S.P. Generalized measure theory. *Found. Phys.* **1973**, *3*, 399–411. [\[CrossRef\]](#)
36. Narens, L. An introduction to lattice based probability theories. *J. Math. Psychol.* **2016**, *74*, 66–81. [\[CrossRef\]](#)
37. Khrennikov, A.Y. *Contextual Approach to Quantum Formalism*; Springer Science & Business Media: Berlin, Germany, 2009.
38. Rivas, Á. On the role of joint probability distributions of incompatible observables in Bell and Kochen–Specker Theorems. *Ann. Phys.* **2019**, *411*, 167939. [\[CrossRef\]](#)
39. Katriňák, T.; Neubrunn, T. On certain generalized probability domains. *Matematický Časopis* **1973**, *23*, 209–215.
40. Šipoš, J. Subalgebras and sublogics of σ -logics. *Math. Slovaca* **1978**, *28*, 3–9.
41. Brabec, J. Compatibility in orthomodular posets. *Časopis pro Pěstování Matematiky* **1979**, *104*, 149–153. [\[CrossRef\]](#)
42. Vallander, S.S. The structure of separable Dynkin algebras. *Vestn. St. Petersburg Univ. Math.* **2016**, *49*, 219–223. [\[CrossRef\]](#)
43. Denneberg, D. *Non-Additive Measure and Integral*; Springer Science & Business Media: Berlin, Germany, 1994.
44. Navara, M.; Pták, P. Considering uncertainty and dependence in Boolean, quantum and fuzzy logics. *Kybernetika* **1998**, *34*, 121–134.
45. Gudder, S.P.; Zerbe, J.E. Generalized monotone convergence and Radon–Nikodym theorems. *J. Math. Phys.* **1981**, *22*, 2553–2561. [\[CrossRef\]](#)
46. De Simone, A.; Navara, M.; Pták, P. Extending states on finite concrete logics. *Int. J. Theor. Phys.* **2007**, *46*, 2046–2052. [\[CrossRef\]](#)
47. De Simone, A.; Pták, P. Measures on circle coarse-grained systems of sets. *Positivity* **2010**, *14*, 247–256. [\[CrossRef\]](#)
48. Horn, A.; Tarski, A. Measures in Boolean algebras. *Trans. Am. Math. Soc.* **1948**, *64*, 467–497. [\[CrossRef\]](#)
49. de Simone, A.; Pták, P. Extending coarse-grained measures. *Bull. Pol. Acad. Sci. Math.* **2006**, *54*, 1–11. [\[CrossRef\]](#)
50. De Finetti, B. *Theory of Probability: A Critical Introductory Treatment*; John Wiley & Sons: New York, NY, USA, 1974; Reprinted in 2017.
51. Troffaes, M.C.; De Cooman, G. *Lower Previsions*; John Wiley & Sons: New York, NY, USA, 2014.
52. Delbaen, F. Coherent risk measures on general probability spaces. In *Advances in Finance and Stochastics*; Sandmann, K., Schönbucher, P.J., Eds.; Springer: Berlin, Germany, 2002; pp. 1–37.
53. Pelessoni, R.; Vicig, P. Imprecise previsions for risk measurement. *Int. J. Uncertain. Fuzziness Based Syst.* **2003**, *11*, 393–412. [\[CrossRef\]](#)
54. Busch, P.; Heinosaari, T.; Schultz, J.; Stevens, N. Comparing the degrees of incompatibility inherent in probabilistic physical theories. *Europhys. Lett.* **2012**, *103*, 10002. [\[CrossRef\]](#)
55. Xu, Z.P.; Cabello, A. Necessary and sufficient condition for contextuality from incompatibility. *Phys. Rev. A* **2018**, *99*, 020103. [\[CrossRef\]](#)
56. Benavoli, A.; Facchini, A.; Zaffalon, M. Quantum mechanics: The Bayesian theory generalized to the space of Hermitian matrices. *Phys. Rev. A* **2016**, *94*, 042106. [\[CrossRef\]](#)
57. Walley, P. Towards a unified theory of imprecise probability. *Int. J. Approx. Reason* **2000**, *24*, 125–148. [\[CrossRef\]](#)
58. Kellerer, H.G. Verteilungsfunktionen mit gegebenen Marginalverteilungen. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **1964**, *3*, 247–270. [\[CrossRef\]](#)
59. Budroni, C.; Cabello, A.; Gühne, O.; Kleinmann, M.; Larsson, J.Å. Kochen–Specker contextuality. *Rev. Mod. Phys.* **2022**, *94*, 045007. [\[CrossRef\]](#)
60. Maharam, D. Consistent extension of linear functionals and of probability measures. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability: Held at the Statistical Laboratory, University of California*; University of California Press: Berkeley, CA, USA, 1972; Volume 6, pp. 127–147.

61. Miranda, E.; Zaffalon, M. Compatibility, Coherence and the RIP. In *Uncertainty Modelling in Data Science, SMPs2018, Proceedings of the International Conference Series on Soft Methods in Probability and Statistics, Compiègne, France, 17–21 September 2018*; Advances in Intelligent Systems and Computing Book Series; Destercke, S.; Denoeux, T., Gil, M.A., Grzegorzewski, P., Hryniewicz, O., Eds.; Springer: Berlin, Germany, 2018; Volume 832, pp. 166–174.
62. Casanova, A.; Kohlas, J.; Zaffalon, M. Information algebras in the theory of imprecise probabilities, an extension. *Int. J. Approx. Reason.* **2022**, *150*, 311–336. [\[CrossRef\]](#)
63. Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.* **1967**, *38*, 325–339. [\[CrossRef\]](#)
64. Birkhoff, G. *Lattice Theory*; American Mathematical Society: Providence, RI, USA, 1940.
65. Rota, G.C. Twelve problems in probability no one likes to bring up. In *Algebraic Combinatorics and Computer Science*; Springer: Berlin, Germany, 2001; pp. 57–93.
66. Khrennikov, A.Y. Bell could become the Copernicus of probability. *Open Syst. Inf. Dyn.* **2016**, *23*, 1650008. [\[CrossRef\]](#)
67. Wirch, J.L.; Hardy, M.R. Distortion risk measures: coherence and stochastic dominance. In *Proceedings of the International Congress on Insurance: Mathematics and Economics, State College, PA, USA, 2001*; pp. 15–17.
68. Fröhlich, C.; Williamson, R.C. Risk measures and upper probabilities: Coherence and stratification. *arXiv* **2022**, arXiv:2206.03183.
69. Lehrer, E. Partially specified probabilities: Decisions and games. *Am. Econ. J. Microecon.* **2012**, *4*, 70–100. [\[CrossRef\]](#)
70. Lehrer, E. *Coherent Risk Measures Induced by Partially Specified Probabilities*; Technical Report; Tel Aviv University: Tel Aviv, Israel, 2007.
71. De Castro, L.I.; Chateaufort, A. Ambiguity aversion and trade. *Econ. Theory* **2011**, *48*, 243–273. [\[CrossRef\]](#)
72. Seidenfeld, T. Dietrich College of Humanities and Social Sciences, Carnegie Mellon University, Pittsburgh, PA, USA. Personal Communication at ISIPTA, 2023.
73. Ovchinnikov, P.G. Measures on finite concrete logics. *Proc. Am. Math. Soc.* **1999**, *127*, 1957–1966. [\[CrossRef\]](#)
74. Godowski, R. Varieties of orthomodular lattices with a strongly full set of states. *Demonstr. Math.* **1981**, *14*, 725–734. [\[CrossRef\]](#)
75. Suppes, P. The probabilistic argument for a non-classical logic of quantum mechanics. *Philos. Sci.* **1966**, *33*, 14–21. [\[CrossRef\]](#)
76. Pták, P. Some nearly Boolean orthomodular posets. *Proc. Am. Math. Soc.* **1998**, *126*, 2039–2046. [\[CrossRef\]](#)
77. Jun, S.J.; Pinkse, J.; Xu, H. Tighter bounds in triangular systems. *J. Econom.* **2011**, *161*, 122–128. [\[CrossRef\]](#)
78. Chow, Y.S.; Teicher, H. *Probability Theory: Independence, Interchangeability, Martingales*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 1988.
79. Rubshtein, B.Z.A.; Grabarnik, G.Y.; Muratov, M.A.; Pashkova, Y.S. *Foundations of Symmetric Spaces of Measurable Functions*; Springer: Berlin/Heidelberg, Germany, 2016.
80. Huber, P.J. *Robust Statistics*; John Wiley & Sons: New York, NY, USA, 1981.
81. Kadane, J.B.; Schervish, M.J.; Seidenfeld, T.; Goel, P.K.; Zellner, A. Statistical implications of finitely additive probability. In *Bayesian Inference and Decision Techniques*; Elsevier Science Publisher: Amsterdam, The Netherlands, 1986; pp. 69–76; Reprinted in *Rethinking the Foundations of Statistics*, pages 211–232. 1999.
82. Chichilnisky, G. The foundations of probability with black swans. *J. Probab. Stat.* **2010**, *2010*, 838240. [\[CrossRef\]](#)
83. Aliprantis, C.D.; Border, K.C. *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd ed.; Springer: Berlin/Heidelberg, Germany, 2006.
84. Munkres, J. *Topology*, 2nd ed.; Pearson Education Limited: London, UK, 2014.
85. Wojtaszczyk, P. *Banach Spaces for Analysts*; Cambridge University Press: Cambridge, UK, 1991.
86. Stone, M.H. The theory of representations for Boolean algebras. *Trans. Am. Math. Soc.* **1936**, *40*, 37–111.
87. Birkhoff, G.; von Neumann, J. The logic of quantum mechanics. *Ann. Math.* **1936**, *37*, 1–26. [\[CrossRef\]](#)
88. Husimi, K. Studies on the foundation of quantum mechanics. I. *Proc. Phys. Math. Soc. Jpn. 3rd Ser.* **1937**, *19*, 766–789.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.