Artificial Intelligence 275 (2019) 457–486

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

The logic of qualitative probability

James P. Delgrande[a,*], Bryan Renne[b], Joshua Sack[c]

[a] School of Computing Science, Simon Fraser University, Burnaby, B.C. V5A 1S6, Canada
[b] Vancouver, B.C., Canada
[c] Department of Mathematics and Statistics, California State University Long Beach, Long Beach, CA 90840, USA

ARTICLE INFO

Article history:
Received 19 March 2018
Received in revised form 21 May 2019
Accepted 4 July 2019
Available online 10 July 2019

Keywords:
Knowledge representation
Qualitative probability
Modal logic

ABSTRACT

In this paper we present a theory of qualitative probability. The usual approach of earlier work was to specify a binary operator $\preceq$ on formulas with $\phi \preceq \psi$ having the intended interpretation that the event expressed by ϕ is no more probable than that expressed by ψ. We generalise these approaches by extending the domain of the operator $\preceq$ from the set of events to the set of finite sequences of events. If Φ and Ψ are finite sequences of events, $\Phi \preceq \Psi$ has the intended interpretation that the combined probabilities of the elements of Φ are no greater than those of Ψ. A sound and complete axiomatisation for this operator over finite outcome sets is given. We argue that our approach is more perspicuous and intuitive than previous accounts. As well, we show that the approach is sufficiently expressive to capture the results of axiomatic probability theory and to encode rational linear inequalities. We also prove that our approach generalises the two major accounts for finite outcome sets.

1. Introduction

Much of Artificial Intelligence (AI) in one fashion or another deals with uncertain information. Classical probability theory provides the standard formal framework for expressing and reasoning with uncertain information: propositions are assigned a probability, perhaps within a given context or conditioned by given evidence; these numbers may then be related according to the familiar rules of probability. As well, there has been interest in qualitative probabilistic reasoning, where one may assert that some event is more probable than another without specifying the exact numerical probabilities of the events in question. In many cases this latter approach offers a pragmatic, intuitive, and practical counterpoint to classical probability theory, both in commonsense reasoning in particular, and in Artificial Intelligence in general.

Classical probability theory provides a quantitative framework for reasoning under uncertainty; it has had many significant, notable applications across diverse areas of AI. However, in some cases classical probability may be too fine-grained or demanding; and in many situations, determining exact numerical probabilities may be difficult or impossible. This could be because a full theory of a domain of application is lacking, as is the case, for example, in much of the biological and medical sciences. Or it may simply be too difficult, or not worth the effort, to determine probabilities when a comparative measure will do. For example, it is usually not worth expending a great deal of energy in deciding whether or not to take an umbrella. Furthermore, reasoning with exact probabilities can be complex, and typically relies on certain (in)dependence assumptions. And often one simply wants to compare the likelihood of two events without having to give exact probabili-

* Corresponding author.
E-mail addresses: jim@cs.sfu.ca (J.P. Delgrande), bryan@renne.org (B. Renne), joshua.sack@csulb.edu (J. Sack).

https://doi.org/10.1016/j.artint.2019.07.002
0004-3702/

ties. Thus for example, without using exact probabilities, someone might believe that her second-choice candidate is more likely to win than her first-choice candidate, and cast her vote accordingly. A final, crucially-important, reason for being able to deal with qualitative information in AI is that a general knowledge-based system will simply have no choice: a knowledge-based system must be able to reason with the information that it is given, and such information will often be qualitative rather than quantitative. This in turn reflects the fact that humans appear to most often use qualitative rather than quantitative expressions of information.

In a qualitative approach, one can assert that some event is more probable than another without specifying exact numerical probabilities. Such an approach avoids the above difficulties, in that one is not obliged to determine specific probabilities. Moreover, particularly in conversation or in commonsense reasoning, assertions of qualitative probability will often convey information at the "right" level of detail. Thus, in saying that if it rains the picnic will probably be cancelled, the essential information regarding the occurrence (or not) of the picnic is conveyed.

The division in AI (and perhaps science as a whole) between qualitative and quantitative approaches has been explored by, among others, Henry Kyburg (e.g. [22]) between what he calls the *probabilist* and the *logicist* way of thinking about the world. The former might make hedged claims, as with probabilistic reasoning, while the latter may make categorical claims made in a hedged way, as developed in approaches to nonmonotonic reasoning. So a full account of qualitative probability may shed light on the relation between quantitative notions of uncertainty on the one hand, and categorical claims based on likelihood on the other.

Our goal in this paper is to develop and explore the foundations of qualitative probability theory, beginning from first principles. We do this by generalizing from comparisons of likelihoods of pairs of formulas to comparisons of the combination of likelihoods of pairs of sequences of formulas; combinations are performed using an operator we call "summation" and comparison is using a relation called "at most".[1] Qualitative probability assertions are of course not arbitrary, but rather certain entailment relations hold between assertions. For example, we assert that "at most" is transitive: if the sum of likelihoods of events in Φ is no greater than those of Ψ, and the sum of likelihoods of events in Ψ is no greater than those of Ξ, then the sum of likelihoods of events in Φ is no greater than those of Ξ. Another example is that "sum" respects commutativity: if the sum of the likelihoods of A and B is no greater than the likelihood of C, then the sum of likelihoods of B and A is no greater than that of C. The central issue is to provide a satisfactory formal characterisation of qualitative probability or, more precisely, to specify the principles that a binary operator $\preceq$ must satisfy in order to exactly capture the intended interpretation "is no greater than" and that the combination operator $\oplus$ satisfies "summation". Specifically, the problem is to give conditions on the operators $\preceq$ and $\oplus$ so that, for a given consistent set of assertions, there is guaranteed to be a (quantitative) probability assignment that is compatible with $\preceq$ and $\oplus$ in this set of assertions. There has been substantial previous work on this topic, without the involvement of summation. However, it has been a surprisingly difficult and subtle problem to provide a characterisation of qualitative probability that is both complete (with respect to the quantitative probability interpretation) and intuitive. We argue that previous work is not wholly satisfactory in this regard.

In Section 3 we give a detailed overview of previous work but, to set the stage, we briefly summarise this work here. Work in qualitative probability goes back to de Finetti [5,6] who gave a number of principles that he conjectured were sufficient to capture this notion. Kraft et al. [20] showed that these principles were not sufficient, and added a condition to de Finetti's to obtain a necessary and sufficient set. A simpler version of their result was given by Scott [28]. Building on Scott's work, Segerberg [29] provided an axiomatisation of an operator $\preceq$ that was sound and complete for the probability interpretation. Gärdenfors [10] provided a simplified account for finite outcome sets. A drawback to these approaches is that the condition identified by Kraft et al. is unwieldy and non-finite. In the case of Segerberg's and Gärdenfors's axiomatisations, this condition is represented by infinitely many axiom schemata whose size grows exponentially and is, again, nonobvious. In a somewhat different vein, Fagin et al. [9] provide a quantitative (as opposed to qualitative) approach to reasoning about probability. Their approach is expressed at a much higher level, and assumes the existence of integers, as well as addition and multiplication.

We address these problems and generalise previous approaches by extending the domain of the operator $\preceq$ from the set of formulas to the set of finite *sequences* of formulas. If Φ and Ψ are finite sequences of formulas, $\Phi \preceq \Psi$ has the intended interpretation that the summed probabilities of the elements of Φ is not greater than the summed probabilities of the elements of Ψ.[2] Our goal is to develop a theory of qualitative probability that is complete, foundational, and perspicuous. That is, by *complete* we mean that the theory expresses the set of conditions identified in [20]; by *foundational*, the theory makes minimal assumptions in the axiomatic account; and by *perspicuous*, the resulting axiomatisation is intuitive, clear,

[1] The terminology "summation" and "at most" do not necessarily presuppose numbers or quantities. For example, the summation of non-numeric elements is common in group theory and ring theory, and a non-numeric element being no greater than another is permitted by order theory. However, the formal semantics of our qualitative probability logic interprets such expressions quantitatively, and our axiomatisation shows how this quantitative interpretation can be captured qualitatively. This mirrors the way the formal semantics of earlier qualitative probability logics interprets the language quantitatively. We return to this point at the end of Section 3.1 once relevant background material has been presented.

[2] There is another way of viewing how we generalise earlier work. In Segerberg's and Gärdenfors's approaches, the comparison of summations of likelihoods of pairwise inconsistent formulas can be expressed using disjunction; this follows from the additivity condition of probability. In other words, the $\oplus$ operator coincides with disjunction $\bigvee$ (expressible in their language) when the formulas are pairwise inconsistent (our theory proves this: see Theorem 4.9.1), and our $\oplus$ generalizes the summation of likelihoods to all sequences regardless of inconsistency.

and readily understandable. We provide a sound and complete axiomatisation for our notion of qualitative probabilistic comparison over finite outcome sets. We argue that this approach is simpler, more perspicuous, and more intuitive than previous accounts. Unlike Segerberg's and Gärdenfors's axiomatisations, ours is schematically finite and it avoids the use of an exponentially-large scheme. Unlike Fagin et al., our approach is qualitative and we do not employ the machinery of arithmetic in expressing our proof theory. Further, our approach is sufficiently expressive to capture results of axiomatic probability theory. For example, the relation $P(\phi) + P(\psi) = P(\phi \vee \psi) + P(\phi \wedge \psi)$ is a theorem of our system, expressed as $\phi \oplus \psi \approx (\phi \vee \psi) \oplus (\phi \wedge \psi)$. Last, our framework generalises the approaches for finite outcome sets due to Gärdenfors and Fagin et al., and captures a modest restriction of Segerberg's logic.

The next section introduces the concepts from modal logic and probability used in this paper, while the third section reviews earlier work on qualitative probability and related notions. Section 4 describes our logic LQP, including soundness and completeness results, and key derivations obtained in the system; this also includes a discussion of expressing quantitative notions in what is arguably a wholly qualitative theory, culminating in valid rational linear inequalities. Section 5 briefly compares our approach with other work, while the last section is a brief conclusion. There are two appendices; the first contains further details on related work, while the second contains proofs. This paper is a substantially expanded version of [7].

2. Mathematical preliminaries

In the paper we make use of basic concepts from modal logic and probability theory. For background on modal logic, the reader may consult any one of the excellent textbooks such as [2,4,19]; the reader need only be acquainted with the syntax and Kripke semantics of the modal logics KD and S5, although passing reference will be made to classical systems of modal logic [4] in Section 3.2. As for probability theory, most of the paper will be concerned with finite spaces; however, in a couple of places the reader will need to have some basic understanding of how probabilities are defined over infinite spaces using σ-algebras. For further information, we refer the reader to any basic text on probability theory. We also refer the reader to [15,25,26] for information both on probability theory and on its role in reasoning about uncertainty in AI.

2.1. Modal logic

In a modal logic, a proposition is true or false at a *possible world*. In the standard formulation of a possible worlds semantics, often referred to as *Kripke semantics*, a model is a triple $M = \langle W, R, P\rangle$ where W is a set (of possible worlds); R is a binary relation between possible worlds, specifying which worlds are *accessible* from which other worlds; and P is an assignment of truth values to atomic sentences at possible worlds. The language of a modal logic is that of propositional logic augmented by a unary operator $\Box$.

More formally, let $\mathcal{P}$ be a non-empty set of *propositional atoms*. A *model* is a triple $M = \langle W, R, P\rangle$ where:

1. W is a set (of possible worlds),
2. $R \subseteq W \times W$ is an *accessibility relation*, and
3. $P : W \to 2^{\mathcal{P}}$.

Thus $R(w, w')$ just if w' is accessible from w or, according to w, w' is a possible world. P specifies for each world which propositional atoms are true at that world. The semantics is given in terms of a *pointed model*, (M, w) where w is a possible world in M. Truth conditions are as follows, where for convenience we drop the parentheses from a pointed model:

1. $M, w \models p$ iff $p \in P(w)$ where $p \in \mathcal{P}$.
2. $M, w \models \neg\phi$ iff $M, w \not\models \phi$.
3. $M, w \models \phi \vee \psi$ iff $M, w \models \phi$ or $M, w \models \psi$.
4. $M, w \models \Box\phi$ iff for every $w' \in W$ such that $R(w, w')$, we have $M, w' \models \phi$.

$M, w \models \phi$ asserts that ϕ is true at world w in model M while $M, w \models \Box\phi$ asserts that ϕ is true at every world w' such that Rw, w'.

The modal logic corresponding to the above semantics is called K; it is defined by the following axiom schemes and rules of inference:

(PC) All tautologies of classical propositional logic
(K) $\Box(\phi \supset \psi) \supset (\Box\phi \supset \Box\psi)$
(MP) From $\phi \supset \psi$ and ϕ, infer ψ
(Nec) From ϕ, infer $\Box\phi$

In the sequel, we will deal with the modal logic KD, in which the accessibility relation is *serial*, that is, where for any $w \in W$, there is w' such that $R(w, w')$. This is captured by the axiom:

(D) $\Box\phi \supset \neg\Box\neg\phi$.

Last, for modal logic S5, the accessibility relation is an equivalence relation, that is, it is reflexive, symmetric, and transitive. S5 has been used to model introspective reasoners in which the formula $\Box\phi$ has intended interpretation "the agent knows that ϕ is true". The axiomatisation is that of K augmented with the following axioms:

(T) $\Box\phi \supset \phi$
(4) $\Box\phi \supset \Box\Box\phi$
(5) $\neg\Box\phi \supset \Box\neg\Box\phi$

2.2. Probability

To define a basic probabilistic system, we begin with a nonempty set Ω called the *sample space*. Each member of the sample space is called an *outcome*. The sample space represents all possible ways an experiment or occurrence might turn out. For example, the sample space might be $\{1, 2, 3, \dots, 6\}$, with each outcome representing a possible throw of a six-sided die. We call a set of outcomes an *event* or a *proposition*. A proposition is taken to be true if and only if, after the experiment or occurrence, the actual outcome is among those that make up the proposition in question. Thus, for the outcome set $\Omega = \{1, 2, 3, \dots, 6\}$ representing a throw of a six-sided die, the proposition $\{2, 4, 6\}$ represents the situation where the result of the throw is an even number.

Probability can be thought of as the expectation some agent has as to the likelihood of various propositions. This likelihood is measured as a real number in the range $[0, 1]$, with 0 indicating the agent considers the proposition impossible and 1 indicating the agent considers the proposition a certainty. Thus it is impossible for a throw of the six-sided die to show a seven (i.e., a throw of seven has probability 0). But it is certain that a throw will yield a whole number between 1 and 6, inclusive (i.e., a throw of 1 to 6 has probability 1, thereby excluding the possibility the die lands on an edge, does not ever land, or some such other strange outcome).

In the simplest case, probabilities are given by a function, called a *probability measure*, P, that assigns to a proposition E a real number $P(E) \in [0, 1]$ subject to *Kolmogorov's axioms*:

1. $0 \le P(E)$
2. $P(\Omega) = 1$
3. if $E_0, E_1, E_2, \dots$ is a pairwise disjoint collection of propositions, then

$$P(\textstyle\bigcup_{i\in\omega} E_i) = \sum_{i\in\omega} P(E_i) \ .$$

That is, propositions are assigned non-negative probabilities; the outcome must be among those in the sample space; and the probability of a proposition can be obtained by partitioning the proposition in to non-overlapping pieces and adding up the probabilities of each piece.

Relating this to modal logics, a set of possible worlds can be thought of as the set of outcomes. A proposition (or event) then can be specified by a subset of these possible worlds. In the finite case every set of possible worlds can be characterised by a formula, and a probability can be assigned to every set of possible worlds.

However, there may be cases where one does not want probabilities assigned to every set of possible worlds, or indeed it may be that probabilities cannot be assigned (as in the infinite case) so that other desirable properties hold. To this end, a probability measure can be more generally defined with respect to a *σ-algebra*, where a σ-algebra of subsets of Ω is a set of subsets of Ω that contains Ω and is closed under complement relative to Ω and under countable unions.[3]

Then, a probability measure on a σ-algebra $F \subseteq 2^{\Omega}$ is a function of type $F \to [0, 1]$ satisfying the following generalised *axioms of Kolmogorov*:

1. $0 \le P(E)$ for each event $E \in F$
2. $P(\Omega) = 1$
3. if $\langle E_i \rangle_{i\in\omega}$ is a pairwise disjoint countable sequence of events in F, then

$$P(\textstyle\bigcup_{i\in\omega} E_i) = \sum_{i\in\omega} P(E_i) \ .$$

3. Background

3.1. Qualitative probability

Consider the problem of specifying a relation $\preceq$ between formulas,[4] where $\phi \preceq \psi$ has the intended interpretation that ϕ is not more probable than ψ. That is, the problem is to provide conditions on $\preceq$ such that for a set of such assertions there is guaranteed to be a *realizing* probability measure $P(\cdot)$ on formulas. This means that for all formulas ϕ and ψ,

[3] Contrast this with the notion of an *algebra* which is a set of subset of Ω such that it first, contains Ω and, second, is closed under complement and *finite* unions.

[4] We henceforth talk of *formulas* rather than *events* when referring to sentences of some logic.

$$\phi \preceq \psi \text{ iff } P(\phi) \leq P(\psi).$$

De Finetti [5,6] conjectured that the following conditions were necessary and sufficient, where 0 is some inconsistent formula: For each ϕ, ψ, and γ,

1. $0 \preceq \psi$
2. $\phi \preceq \psi$ and $\psi \preceq \gamma$ implies $\phi \preceq \gamma$
3. $\phi \preceq \psi$ or $\psi \preceq \phi$
4. If $\phi \wedge \gamma$ and $\psi \wedge \gamma$ are each inconsistent, then: $\phi \preceq \psi$ iff $\phi \vee \gamma \preceq \psi \vee \gamma$.

While these conditions are clearly sound, Kraft et al. [20] showed that they are not complete, in that there are orderings on propositions that satisfy de Finetti's conditions but for which there is no realizing probability measure. For our purposes, their counterexample is most easily phrased in terms of possible worlds. Consider the set of possible worlds

$$W = \{w_1, w_2, w_3, w_4, w_5\}.$$

A subset of W can be thought of as representing a proposition. Consider the relations:

$$\begin{array}{rclcrcl} \{w_3\} & \preceq & \{w_1, w_2\} & \qquad & \{w_1, w_5\} & \preceq & \{w_2, w_3\} \\ \{w_2, w_4\} & \preceq & \{w_1, w_3\} & & \{w_1, w_2, w_3\} & \preceq & \{w_4, w_5\}. \end{array}$$

[20] show that these relations can be extended to an ordering on all subsets of W that satisfies de Finetti's conditions but for which there is no corresponding probability measure. (In the counterexample, an assignment of probability of .2 to each world is easily seen to be inconsistent. [20] show that *every* such assignment of probabilities is inconsistent.) They also provide a criterion so as to ensure a realizing probability measure always exists.

Scott [28] reformulated and simplified these results in an algebraic form. Segerberg [29] developed a logic of qualitative probability that made use of Scott's results; this logic had a binary operator $\preceq$ and a unary modal operator $\Box$ of necessitation. Gärdenfors [10] (also [11]) subsequently simplified Segerberg's approach by restricting to finite sets and defining necessitation as probabilistic certainty: $\Box\phi \doteq (1 \preceq \phi)$.

Segerberg's and Gärdenfors's axiomatisations both use the following abbreviation schema:

$$\phi_1, \ldots \phi_m \mathbb{E} \psi_1, \ldots \psi_m \doteq \Box(C_0 \vee \cdots \vee C_m) \tag{1}$$

where $m \geq 1$ and for $0 \leq i \leq m$, the formula C_i is the disjunction of all conjunctions

$$e_1\phi_1 \wedge \cdots \wedge e_m\phi_m \wedge f_1\psi_1 \wedge \cdots \wedge f_m\psi_m$$

where exactly i of the e's and i of the f's are the negation sign, and the rest are the empty string. The overall import is that a disjunct C_i in (1) asserts that exactly i of the ϕ's and i of the ψ's are false; and so $\phi_1, \ldots \phi_m \mathbb{E} \psi_1, \ldots \psi_m$, which we write as $(\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m$, asserts that exactly the same number of ϕ's are true as are the ψ's. Then each of their logics contains the following schema, encoding the Kraft et al. [20] condition[5]:

(A4) $((\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m \wedge \bigwedge_{k=1}^{m-1} (\phi_k \preceq \psi_k)) \supset (\psi_m \preceq \phi_m)$ for all $m \geq 1$

Gärdenfors gives an axiomatisation for his logic QP of *qualitative probability* that encodes the de Finetti principles along with the schema (A4) in the context of propositional logic. Segerberg's logic PK is more general and his axiomatisation, while largely analogous to Gärdenfors's, is more elaborate. These logics are shown to be sound and complete with respect to a possible worlds model in which a probability measure is associated with sets of possible worlds.

Gärdenfors's and Segerberg's axiomatisations are not ideal for several reasons. First, if an axiomatisation is intended to clearly lay out underlying principles for deductions in a logic, the key axiom schema (A4) would fail this criterion, in that it is opaque and non-perspicuous. (Segerberg calls it "formidable".) Second, the disjunction represented in the $\mathbb{E}$ definition grows exponentially with m.[6] Third, (A4) specifies *infinitely many* axiom schemas, one for each positive integer m; consequently, the above axiomatisation for QP is not schematically finite.

A different (and quite separate) approach to reasoning about probability is the theory $\mathsf{AX}_{\mathrm{meas}}$ of *quantitative* probability of [9]. Their language permits Boolean combinations of linear inequalities of the form

$$c \preceq a_1 w(\phi_1) + \cdots + a_n w(\phi_n),$$

where c and the a_i's are integers and the ϕ_i's do not contain $\preceq$'s or $w(\cdot)$'s. That is, their language does not allow nesting of inequalities. The expression $w(\phi)$ is mapped in the semantics to a real number called the "weight" of ϕ. This ends up being the probability of event ϕ. A generalisation of this theory is presented in [8]. This latter formalism includes multiple agents and an additional S5 modal operator for knowledge; as well nestings of the operator $\preceq$ are admitted. Given that this system diverges from our interests at hand, in later discussions we focus on the theory $\mathsf{AX}_{\mathrm{meas}}$ of [9].

In Section 5, we compare our approach with the logics QP, PK, and $\mathsf{AX}_{\mathrm{meas}}$.

[5] Gärdenfors calls the schema (A4(m)).

[6] More precisely, $(\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m$ contains $\sum_{i=0}^m \binom{m}{i}$ disjuncts, which is bounded below by 2^m and above by $(2^m)^2$.

Qualitative vs. quantitative probability What distinguishes qualitative from quantitative probability (truth valued) logics is that qualitative probability logics do not employ quantities or arithmetic operations in the syntax, and the informal reading of the qualitative probability formulas do not require a quantitative interpretation. For example, the quantitative probability logic of [9] makes numbers explicit in the language. On the other hand, the comparison operator $\preceq$ of qualitative probability logic need not imply that the likelihoods being compared are numeric. As well, while we use the suggestive term "summation", it does not presuppose numbers as operands; as we noted earlier, the summation of non-numeric elements is common in group theory and ring theory. However, both qualitative and quantitative probability logics share a formal semantics involving real (quantitative) probability. In the case of a qualitative probability logic, the axiomatisation thus is shown (via the soundness and completeness results) to faithfully reflect probabilistic principles, in that for a consistent set of assertions, there is guaranteed to be a corresponding realising probability distribution.

3.2. Other approaches

In this section we briefly summarise work that is not directly related to a general notion of qualitative probability, but that nonetheless can be seen as addressing similar concerns.

To begin, there has been work addressing a modal notion of *probable*, where the formula $P\phi$ has the intended interpretation that ϕ is probably true. *Probable* in these cases could mean that the probability of ϕ is at least .5, or is greater than .5, or is greater than some fixed parameter $c > .5$. Herzig [17] investigates such a modal operator along with other modal concepts for belief and action. (As well [17] contains a good overview of earlier work on modal probability.) The P operator is not closed under conjunction, and so the semantics is described not in terms of Kripke structures, but rather in terms of *minimal models* [4]. The principles governing the P operator alone are, following [3], as follows:

(N) $P\top$
(D) $\neg(P\phi \wedge P\neg\phi)$
(RM) From $\phi \supset \psi$ infer $P\phi \supset P\psi$.

Kyburg and Teng [21] deal with a similar notion, where in their approach $\Box_\epsilon \phi$ holds just if the probability of $\neg\phi$ is no more than a fixed (small) parameter ϵ. The resulting logic is that of the classical system of modal logic EMN [4]. This logic can be axiomatised by (N) and (RM), above.

van Eijck and Renne [30] extend this analysis to a logic of knowledge and belief where an agent knows a formula ϕ, expressed $K\phi$, just if ϕ has probability 1; and the agent believes ϕ, $B^c\phi$, just if the probability of ϕ is at least some fixed value $c > .5$. The logic of K is S5 while the logic of B is EMND45, which in turn is essentially Burgess's logic (above) but with positive and negative introspection. As noted by the authors, some of this work extends and elaborates on work originally appearing in [23].

Other work examines approaches not necessarily based on an underlying probabilistic semantics. Halpern and Rabin [13] develop a modal logic of *likelihood*, where $L\phi$ is interpreted as "L is likely" but where the meaning of "likely" is relative to a user. Their rationale is that there are situations where we want to reason about likelihood beyond probability theory or where probability theory may not be applicable. In more detail, $L\phi$ has interpretation "ϕ is reasonably likely to be a consistent hypothesis" which, as the authors note, is much weaker than the statement "ϕ holds with probability at least .5". Iterations of L give weaker notions of likelihood, so if $L\phi$ is taken to mean that ϕ is reasonably likely then $LL\phi$ could be taken to mean that ϕ is somewhat likely. The semantics is given in terms of a Kripke structure in which $L\phi$ is true at a possible world w just if there is an accessible possible world w' where ϕ is true. Consequently, $L\phi \wedge L\neg\phi$ is satisfiable and $L\phi \vee L\neg\phi$ is a theorem. [12] examines this approach with regards to a probabilistic interpretation.

Holliday and Icard [18] (see also [16]) explore semantics for notions of qualitative probability, including models based on a measure semantics (such as those of the previous subsection), as well as those based on orderings over possible worlds and orderings over events. Thus for example, for a semantics based on orderings over possible worlds, the position of a world in an ordering may represent its relative likelihood; this in turn can be lifted to an ordering on sets of worlds, or events [14]. While such (world- or event-) based semantics can capture notions of preferential reasoning, arguably they are not suitable for a foundational account of qualitative probability. That is, while orderings over worlds or sets of worlds capture preferential reasoning, there is then another step to show that such semantics structures are compatible with a probability interpretation. As well, it may not be straightforward to get such event-ordered and world-ordered semantics to agree with a probabilistic interpretation. Last, [24] presents a more general theory of probability than Kolmogorov's, addressing in particular questions of qualitative foundations.

4. A logic of qualitative probability

In this section we present our approach to qualitative probability. The language is that of propositional logic augmented by a binary modal operator $\preceq$ on sequences of formulas. The goal is to axiomatically provide conditions on $\preceq$ such that for a set of formulas over this language, this set is consistent if and only if there is guaranteed to be a *realizing* probability measure $P(\cdot)$ that satisfies these formulas. The approach then is purely qualitative in that the axiomatisation makes no reference to numbers, arithmetic, etc., but refers only to formulas in this extended propositional language; properties of $\preceq$, such as transitivity, are explicitly given in the axiomatisation. While the symbol 1 may appear in formulas, it simply stands for some specific tautology, such as $p \vee \neg p$.

The next subsection covers the language and semantics. The semantic theory is not new, and appears in many approaches to probability: given a set of possible worlds, each world is assigned a probability. The probability of a formula in a model then is just the sum of the probability of those worlds at which the formula is true. Then, in our approach, $\Psi \preceq \Phi$ is satisfied at a world just if the summed probabilities of the formulas in Ψ is not greater than those of Φ. This is followed in the second subsection by the axiomatisation, which specifies the (qualitative) principles governing $\preceq$. An extensive exploration of derivable formulas is given and we finish with soundness and completeness results. The third subsection shows how qualitative notions can nonetheless be encoded in the approach.

4.1. Language and semantics

Definition 4.1 *(Language $\mathcal{L}_{\mathsf{LQP}}$).* Fix a nonempty set $\mathcal{P}$ of propositional atoms. The language $\mathcal{L}_{\mathsf{LQP}}$ consists of the *formulas* ϕ and the *sequences* Φ formed by the following recursion:

$$
\begin{array}{lll}
\phi & ::= & p \mid \neg\phi \mid (\phi \vee \phi) \mid (\Phi \preceq \Phi) \qquad p \in \mathcal{P} \\
\Phi & ::= & \phi \mid \phi \oplus \Phi
\end{array}
$$

Formulas occurring in sequences are called *elements*, and expressions $\Phi \preceq \Psi$ are called *inequalities*. We use the symbols ϕ, ψ, and χ, possibly with subscripts or superscripts, as metavariables for formulas. We use Φ, Ψ, and Δ similarly for sequences. Sequences may be written using indexed prefix notation so that, for example, $\bigoplus_{i=1}^{3} \phi_i$ denotes $\phi_1 \oplus \phi_2 \oplus \phi_3$. We use $|\Phi|$ to denote the number of elements in Φ, or the *length* of Φ. Thus for example $|\phi \oplus \phi \oplus \psi| = 3$. Note that the syntax excludes empty sequences, and so for any sequence Φ, we have $|\Phi| > 0$. We may write $\phi \in \Phi$ to indicate that ϕ occurs as an element of Φ. We will often drop parentheses when no ambiguity of meaning results; hence we might write simply $\phi_1 \vee \phi_2 \vee \phi_3$.

We use the standard definitions for the Boolean connectives $\wedge$ (conjunction), $\supset$ (material implication), and $\equiv$ (material equivalence). We define 1 to be some arbitrary fixed tautology in the underlying propositional language (e.g., $p \vee \neg p$), and we define 0 as $\neg 1$. $\Phi \approx \Psi$ abbreviates $(\Phi \preceq \Psi) \wedge (\Psi \preceq \Phi)$, and $\Phi \prec \Psi$ abbreviates $(\Phi \preceq \Psi) \wedge \neg(\Psi \preceq \Phi)$. We define $\Box\phi$ to be $(1 \preceq \phi)$.

Our goal is to capture axiomatically the intended interpretation that, for formula $\Phi \preceq \Psi$, the sum of the probability of the elements on the left of the inequality is less than or equal to that on the right. For example, a theorem in our approach is $1 \prec 1 \oplus 1$. Our theory will ensure that every tautology has probability 1. Hence the symbolic expression 1, which in reality is simply some fixed tautology, can be interpreted as having numerical probability 1. (Thus it ought not cause confusion that we use the same symbol to denote both the number 1 and the symbolic abbreviation 1 for a fixed tautology in the language.) Since the tautology 1 has probability 1, the expression $\Box\phi$, itself an abbreviation for $1 \preceq \phi$, has the intuitive meaning that the probability of ϕ is 1 or, equivalently, that ϕ is probabilistically certain.

We now turn to the semantic theory underlying our approach, beginning with our definition of model. This definition is essentially the same as that of [29], [10], and others. It can be noted that this definition is quite general, and more general than what we require. In particular, the axiomatic theory we will present is sound with respect to what are called the class of simple models, below. However, to facilitate our comparisons with previous work, it is most convenient to have at hand the full class of models encompassed by Definition 4.2. We say more about this later.

Definition 4.2 *(Model).* A *model* is a structure $M = (W, \Pi, V)$ such that:

1. W is a nonempty set of objects, or *possible worlds*.
2. Π maps each world $w \in W$ to a tuple $\Pi_w = (\Omega_w, F_w, P_w)$ that satisfies the following principles making it a *probability space*:
   (a) $\Omega_w \subseteq W$ is a nonempty set of *outcomes*;
   (b) $F_w \subseteq 2^{\Omega_w}$ is a σ-algebra of subsets of Ω_w. $E \in F_w$ is called an *event* or *measurable set*;
   (c) $P_w : F_w \to [0, 1]$ is a probability measure on F_w.
3. $V : W \to 2^{\mathcal{P}}$ is a propositional valuation assigning to each world $w \in W$ a set $V(w) \subseteq \mathcal{P}$ of propositional atoms taken to be true at w.

A model is:

- *finite* iff W is finite;
- *uniform* iff $\Pi_w = \Pi_v$ for each $w, v \in W$;
- *total* iff $\Omega_w = W$ for each $w \in W$;
- *powerset* iff $F_w = 2^{\Omega_w}$ for each $w \in W$;
- *simple* iff it is finite and powerset; and
- *super-simple* iff it is uniform, total, and simple.

A *pointed model* is a pair (M, w) consisting of a model M and a world w (called the *point*) coming from the set of worlds of M.

Our interests will mainly be with simple models, and also on occasion with super-simple models. For now, think of a model as specifying a number of possibilities as to what might come to be—these are the possible *outcomes* of some experiment or occurrence. A *proposition* may be identified with a set of outcomes: this proposition is true if and only if the outcome that comes to pass is among those that make up the set.

We next discuss the various terms in Definition 4.2; those familiar with these notions can skip to Definition 4.3. A probability space is used to assign probabilities to certain propositions. Propositions that are assigned probabilities are called *measurable sets* or *events*. Due to mathematical difficulties that may arise when we work with *infinite* models (i.e., those having infinitely many outcomes), it is not always possible for us to assign a probability to every proposition while guaranteeing that our probability function satisfies Kolmogorov's axioms. So it is not always possible for every proposition to be measurable. However, in certain special cases, we do not run into this difficulty. In particular, if a model is *finite* (i.e., there are finitely many outcomes), then it is always possible to construct a probability space that assigns a probability to every proposition, and doing so yields a situation in which every proposition is measurable (i.e., is an event).

We may think of a probability space as describing an agent's beliefs as to the likelihood that a given event will obtain. Then, a rough way of thinking of measurable sets is as follows: A measurable set is a proposition for which the agent's belief state permits her to make a judgement as to the likelihood that the outcome will be among those that make up the set. This likelihood is represented by the assignment of a real-number probability between 0 and 1, with 0 representing perceived impossibility and 1 representing perceived certainty. Note that just because we can construct a probability space in which every proposition is measurable, it does not follow that every probability space on the same outcome set has this property. In particular, even within the class of finite models, it is possible to construct a probability space in which some propositions are non-measurable.[7] Intuitively, in such models, the agent does not have a belief about every proposition. However, *powerset* models disallow this possibility: in a powerset model, every proposition is measurable, and therefore the agent indeed has a belief about (i.e., assigns a probability to) each and every event.

Each world is associated with its own probability space, and the probability space used at one world need not be the same as that used at another. If every world uses the same probability space, then we say the model is *uniform*. Such models are those in which the agent is certain of what she believes (in the sense that she cannot hypothesize another world in which she uses a different probability space). Non-uniform models permit the agent to have uncertainty in her own beliefs (as to which probability space she should use). Thus in non-uniform models there are two levels of uncertainty: the first is the agent's uncertainty as represented by the probability she assigns to events, and the second is the agent's uncertainty as to which probability space she should use to assign probabilities to events. For example, an agent may entertain one outcome whose probability function describes a fair coin (probability 0.5 for heads) and another whose probability function describes a coin biased 3:1 in favour of heads (probability 0.75 heads). In such a situation, she is uncertain both about whether the coin will land on heads and about what probability space she ought to use in assessing the likelihood of heads.

Definition 4.2 makes a distinction between *worlds* and *outcomes*. Properly speaking, models are made up of worlds. (Think of these as possible states-of-affairs.) Each world w gives rise to a probability space $\Pi_w = (\Omega_w, F_w, P_w)$. The set F_w contains the measurable sets and the function P_w assigns a probability to each measurable set. Measurable sets are made up of worlds; however, these worlds must come from the set Ω_w of *outcomes*, where Ω_w need not contain every world. Since the probability measure uses the set of outcomes, we might end up in a situation in which a certain world v is not among the possible set of outcomes. This world v therefore plays no direct role in probability considerations. However, v may be accounted for when considering probabilities of probabilities or even deeper nestings of probabilities. We say that a model is *total* to mean that every world is an outcome, and have $\Omega_v = W$ for every world $v \in W$.

A *simple* model is both finite and powerset. In such a model, the situation is indeed "simple": finiteness ensures we do not have difficulties in assigning probabilities to events, and powerset-ness requires that every event is indeed assigned a probability. However different worlds can still use different probability spaces, and we still might have non-outcome worlds. So even in simple models, the agent can be "uncertain" about her beliefs (in the sense described above) and some worlds need not be among the possible outcomes. Our most refined class of models, the *super-simple* models, disallows both of these: a super-simple model is a simple model that is also uniform and total (i.e., it is uniform, total, finite, and powerset). In a super-simple model, the agent's beliefs are not world-specific, every world is a possible outcome, there are finitely many worlds (and hence finitely many outcomes), and the agent assigns a probability to every proposition.

We turn next to notions of truth and satisfaction in LQP.

Definition 4.3 *($\mathcal{L}_{\mathsf{LQP}}$ Satisfaction, Validity).* The semantic function $[\![\cdot]\!] : \mathcal{L}_{\mathsf{LQP}} \to 2^W$ and the satisfaction relation $\models$ between pointed models and $\mathcal{L}_{\mathsf{LQP}}$-formulas are defined as follows.

- $[\![\phi]\!]_M \doteq \{w \in W \mid M, w \models \phi\}$, where $\phi \in \mathcal{L}_{\mathsf{LQP}}$.
- $[\![\phi]\!]^w_M \doteq [\![\phi]\!]_M \cap \Omega_w$.

[7] For example, define $M = (W, \Pi, V)$ by setting $W = \{w_1, w_2\}$, $F = \{\emptyset, W\}$, $P(\emptyset) = 0$ and $P(W) = 1$, and $\Pi_{w_1} = \Pi_{w_2} = (W, F, P)$.

- $M, w \models p$ iff $p \in V(w)$, where $p \in \mathcal{P}$.
- $M, w \models \neg\phi$ iff $M, w \not\models \phi$.
- $M, w \models \phi \vee \psi$ iff $M, w \models \phi$ or $M, w \models \psi$.
- $M, w \models \Phi \preceq \Psi$ iff we have $[\![\chi]\!]_M^w \in F_w$ for each $\chi \in \Phi$ or $\chi \in \Psi$, and that

$$\sum_{\phi \in \Phi} P_w([\![\phi]\!]_M^w) \leq \sum_{\psi \in \Psi} P_w([\![\psi]\!]_M^w) \ .$$

We say that ϕ is *valid in* M, written $M \models \phi$, to mean that $M, w \models \phi$ for each world w in M. For a class $\mathcal{C}$ of models, we write $\mathcal{C} \models \phi$ and say that ϕ is *valid with respect to* $\mathcal{C}$, to mean that $M \models \phi$ for each $M \in \mathcal{C}$. If the class $\mathcal{C}$ is not mentioned, it is assumed to be the full class of models.

Theorem 4.4. *If M is a super-simple model, then for any world v in M, we have*

$$M, w \models \Phi \preceq \Psi \quad \textit{iff} \quad \textstyle\sum_{\phi \in \Phi} P_v([\![\phi]\!]_M) \leq \sum_{\psi \in \Psi} P_v([\![\psi]\!]_M) \ .$$

A note on possibility Observe that models permit two kinds of "possibility" with respect to a world v: the probabilistic notion (i.e., the singleton event $\{v\}$ has nonzero probability) and the Kripke-style possible worlds notion captured by the modal logic S5 (i.e., the world v is among the set W of all possible worlds).[8] These two notions of possibility need not coincide: a world v may be Kripke-possible (i.e., $v \in W$) while being probabilistically impossible (i.e., the probability of $\{v\}$ is 0). The language $\mathcal{L}_{\mathsf{LQP}}$ is designed to address only the probabilistic notion. In order to address the Kripke notion as well, one would need to include an additional modal necessity operator. This is the approach of Segerberg [29], who adds to the language $\mathcal{L}_{\mathsf{LQP}}$ a modal necessity operator we denote here as $\boxdot$.[9] In general the semantics then must be extended by taking our models $M = (W, \Pi, V)$ and adding a binary operator $R \subseteq W \times W$, obtaining *generalized models* $M = (W, R, \Pi, V)$. Defining $R(w) \doteq \{v \in W \mid (w, v) \in R\}$, we would then extend Definition 4.3 so as to interpret formulas $\boxdot\phi$ at a pointed generalized model as follows: $M, w \models \boxdot\phi$ means that $M, v \models \phi$ for all $v \in R(w)$. Intuitively, $v \in R(w)$ is a "(epistemically) possible world," whereas $u \in \Omega(w)$ is a "potential outcome" of a probabilistic process. One may then wish to restrict attention to the models satisfying the following property: we have $\Omega_w \subseteq R(w)$ for each $w \in W$, which says that all potential outcomes are epistemically possible. The modal operator $\boxdot$ then allows us to refer to epistemic possibility whereas the defined operator $\Box$ allows us to refer to probabilistic possibility. It is useful to have separate operators for these concepts since these concepts need not be the same in probabilistic measures on infinite spaces. For example, if we denote the setting of a light dimmer using the real interval $[0, 1]$ such that 0 denotes minimal intensity, 1 denotes maximal intensity, and we take the probability of the switch falling in a given interval to be the length of that interval, then it is probabilistically impossible that the switch will be set to 0.5 (since a point has zero length), even though this particular setting is epistemically possible.

From the perspective of generalized models, our approach amounts to assuming finite spaces (i.e., W is finite), assuming that all potential outcomes are epistemically possible, and taking R to be an equivalence relation (i.e., it is reflexive, transitive, and symmetric). Segerberg's [29] approach assumes that all outcomes are possible but permits less restrictive assumptions on R and does not require spaces to be finite (i.e., W can be infinite). This generality is interesting and a full comparison of our work here with that of Segerberg [29] would require us to work with generalized models over infinite spaces and consider the epistemic possibility operator $\Box$. However, our goal here is more modest, in that we wish to look at a basic theory for reasoning about qualitative probability in finite spaces. As such, we set aside considerations of not-necessarily finite generalized models and of the $\boxdot$ operator for future work. In this way, our work is closer to the more perspicuous approach of Gärdenfors [10] and has connections with the quantitative probabilistic approach of [9]. We comment more on our connections with these and other works later.

4.2. Axiomatic theory

The previous subsection described a probabilistic semantics for formulas of the form $\Psi \preceq \Phi$. This semantics provides a means of determining the adequacy (via soundness and completeness results) of our axiomatic theory of qualitative probability, given next.

[8] Recall that S5 is sound and complete with respect to the class of Kripke models for which the accessibility relation is *total* (i.e., each world is accessible from itself and any two worlds are accessible from each other in both directions) [2].

[9] Segerberg [29] uses instead the symbol $\Box$. We have defined the latter symbol as $\Box\phi \doteq 1 \preceq \phi$ already, and so use instead $\boxdot$ to denote Segerberg's necessity operator.

Definition 4.5 *(LQP).* LQP is defined by the following axiom schemes and rules:

(PC)	All tautologies of classical propositional logic
(Triv)	$0 \prec 1$
(Tran)	$(\Phi \preceq \Psi) \supset ((\Psi \preceq \Delta) \supset (\Phi \preceq \Delta))$
(Tot)	$(\Phi \preceq \Psi) \vee (\Psi \preceq \Phi)$
(Sub)	$\Box(\phi_1 \equiv \phi_2) \wedge \Box(\psi_1 \equiv \psi_2) \supset ((\phi_1 \oplus \Phi \preceq \psi_1 \oplus \Psi) \equiv (\phi_2 \oplus \Phi \preceq \psi_2 \oplus \Psi))$
(Com)	$(\Phi_1 \oplus \Phi_2 \preceq \Psi) \equiv (\Phi_2 \oplus \Phi_1 \preceq \Psi)$
	$(\Phi \preceq \Psi_1 \oplus \Psi_2) \equiv (\Phi \preceq \Psi_2 \oplus \Psi_1)$
(Add)	$((\Phi_1 \preceq \Psi_1) \wedge (\Phi_2 \preceq \Psi_2)) \supset (\Phi_1 \oplus \Phi_2 \preceq \Psi_1 \oplus \Psi_2)$
(Succ)	$(1 \oplus \Phi \preceq 1 \oplus \Psi) \supset (\Phi \preceq \Psi)$
(K1)	$0 \preceq \phi$
(K3)	$\Box\neg(\phi \wedge \psi) \supset (\phi \oplus \psi \approx \phi \vee \psi)$
(MP)	From $\phi \supset \psi$ and ϕ, infer ψ
(Nec)	From ϕ, infer $\Box\phi$

The axiom (Triv) avoids triviality, while (Tran) and (Tot) specify that $\preceq$ is transitive and connected, respectively. (Sub) is substitution of necessary equivalents with respect to initial elements of a sequence. (Com) expresses that sequences are commutative. (Add) allows one to "combine" two inequalities, while (Succ) allows one to "remove" initial 1 elements from both sides of $\preceq$. (K1) and (K3) correspond to the first and third Kolmogorov axioms; the second Kolmogorov axiom, which essentially says that a valid proposition has probability 1, is expressed by (Nec) and the abbreviation $\Box\phi \doteq (1 \preceq \phi)$. (PC) and (MP) are straightforward, giving that the resulting system subsumes classical propositional logic. While this set of axioms may not be absolutely minimal, we have endeavoured to be parsimonious in their specification.

We next explore derivations of this axiomatisation, showing inter alia that we can derive a rich set of results in our theory. These results are presented in four groups, each given as a theorem. They give results concerning properties of sequences, probability, modal logic, and some extended principles, respectively. We conclude the subsection with soundness and completeness results.

Theorem 4.6. *[Sequences]*

1. *(Ref):* $\vdash_{\mathsf{LQP}} \Phi \preceq \Phi$
2. $\approx$ *is an equivalence relation*
3. $\vdash_{\mathsf{LQP}} (\Phi \preceq \Psi) \supset (\Delta \oplus \Phi \preceq \Delta \oplus \Psi)$
4. *Substitution for length-*1 *sequences:*
$$\vdash_{\mathsf{LQP}} \Box(\phi_1 \equiv \phi_2) \wedge \Box(\psi_1 \equiv \psi_2) \supset ((\phi_1 \preceq \psi_1) \equiv (\phi_2 \preceq \psi_2))$$
5. *General substitution of necessary equivalences:*
Let $\phi_1 \in \Phi_1$ *and* $\psi_1 \in \Psi_1$. *Let* Φ_2 *be the same as* Φ_1 *but with some instance of* ϕ_1 *replaced by* ϕ_2, *and similarly for* Ψ_2. *Then:*
$$\vdash_{\mathsf{LQP}} \Box(\phi_1 \equiv \phi_2) \wedge \Box(\psi_1 \equiv \psi_2) \supset ((\Phi_1 \preceq \Psi_1) \equiv (\Phi_2 \preceq \Psi_2))$$
6. *Replacement principles:*
$$\vdash_{\mathsf{LQP}} (\Phi_1 \approx \Phi_2) \supset ((\Phi_1 \oplus \Phi \preceq \Psi) \equiv (\Phi_2 \oplus \Phi \preceq \Psi))$$
$$\vdash_{\mathsf{LQP}} (\Psi_1 \approx \Psi_2) \supset ((\Phi \preceq \Psi_1 \oplus \Psi) \equiv (\Phi \preceq \Psi_2 \oplus \Psi))$$
7. *Cancellation principle:*
$$\vdash_{\mathsf{LQP}} (\Delta \oplus \Phi \preceq \Delta \oplus \Psi) \supset (\Phi \preceq \Psi)$$
8. *Ordering principle:*
$$\vdash_{\mathsf{LQP}} (\Phi_1 \preceq \Psi_1) \supset ((\Psi_1 \oplus \Psi_2 \preceq \Phi_1 \oplus \Phi_2) \supset (\Psi_2 \preceq \Phi_2))$$

The first three parts of the theorem are straightforward. Part 4 is necessary since in (Sub) we have that a sequence Φ is nonempty. Similarly, Part 5 shows, not unreasonably, that substitution of necessary equivalents holds for arbitrary elements of a sequence. Part 6 extends substitution of necessary equivalents to substitution under $\approx$. Part 7 extends (Succ)

to arbitrary sequences. The Ordering Principle, Part 8, which we will subsequently generalise, is a key for many later results.
The next theorem gives various results from classical probability.

Theorem 4.7 *(Probability).*

1. $\vdash_{\mathsf{LQP}} \phi \approx \phi \oplus 0$
2. $\vdash_{\mathsf{LQP}} (\phi \oplus \psi \approx 0) \supset (\phi \approx 0)$
3. $\vdash_{\mathsf{LQP}} \phi \approx (\phi \wedge \psi) \oplus (\phi \wedge \neg\psi)$
4. $\vdash_{\mathsf{LQP}} \phi \oplus \psi \approx (\phi \vee \psi) \oplus (\phi \wedge \psi)$
5. $\vdash_{\mathsf{LQP}} \phi \preceq 1$
6. $\vdash_{\mathsf{LQP}} \phi \oplus \neg\phi \approx 1$
7. $\vdash_{\mathsf{LQP}} \phi \vee \psi \preceq \phi \oplus \psi$
8. $\vdash_{\mathsf{LQP}} (\phi \preceq \psi) \supset (\neg\psi \preceq \neg\phi)$

All of these results are standard from elementary probability theory. Thus Part 3 expresses that

$$P(\phi) = P(\phi \wedge \psi) + P(\phi \wedge \neg\psi),$$

while Part 4 expresses that

$$P(\phi) + P(\psi) = P(\phi \vee \psi) + P(\phi \wedge \psi),$$

which is usually expressed as

$$P(\phi \vee \psi) = P(\phi) + P(\psi) - P(\phi \wedge \psi).$$

Other results may be read off analogously.
The next theorem concerns the necessitation operator. The first and last part relate $\Box$ to $\preceq$, while the second and third items show that the underlying modal logic for $\Box$ is KD.[10] Since this means that $\Box$ is a *normal* modal operator, we henceforth use results regarding normal modal logics freely.

Theorem 4.8 *(Modal Logic).*

1. $\vdash_{\mathsf{LQP}} \Box(\phi \supset \psi) \supset (\phi \preceq \psi)$
2. $\vdash_{\mathsf{LQP}} \Box(\phi \supset \psi) \supset (\Box\phi \supset \Box\psi)$
3. $\vdash_{\mathsf{LQP}} \neg\Box 0$
4. $\vdash_{\mathsf{LQP}} \Box\phi \supset ((\phi \wedge \psi) \approx \psi)$

The following results generalise earlier results from the binary case to the general n-ary case. While of limited interest as independent results, they are used in the completeness proof and for relating the approach to other work. The first part extends Theorem 4.7.3 to an arbitrary number of pairwise-inconsistent formulas. The second part is straightforward but useful. The third and fourth parts generalise Theorem 4.6.8.

Theorem 4.9 *(Extended Principles). For formulas ϕ_i and ψ_i for $1 \le i \le n$, and sequences Φ_i and Ψ_i for $1 \le i \le n$, we have:*

1. $\vdash_{\mathsf{LQP}} \bigwedge_{1 \le i \ne j \le n} \Box\neg(\phi_i \wedge \phi_j) \supset \left(\bigoplus_{i=1}^n \phi_i \approx \bigvee_{i=1}^n \phi_i\right)$
2. $\vdash_{\mathsf{LQP}} \bigwedge_{i=1}^n \Box(\phi_i \equiv \psi_i) \supset \left(\bigoplus_{i=1}^n \phi_i \approx \bigoplus_{i=1}^n \psi_i\right)$
3. $\vdash_{\mathsf{LQP}} \left(\left(\bigoplus_{i=1}^n \Phi_i \approx \bigoplus_{i=1}^n \Psi_i\right) \wedge \bigwedge_{i=1}^n (\Phi_i \preceq \Psi_i)\right) \supset (\Psi_k \preceq \Phi_k)$
4. $\vdash_{\mathsf{LQP}} \left(\left(\bigoplus_{i=1}^n \Phi_i \approx \bigoplus_{i=1}^n \Psi_i\right) \wedge \bigwedge_{i=1}^{n-1} (\Phi_i \preceq \Psi_i)\right) \supset (\Psi_n \preceq \Phi_n)$

The axiomatisation is sound and complete with respect to the class of simple models:

Theorem 4.10 (LQP *Soundness). Let $\mathcal{C}_s$ be the class of simple models. For each $\phi \in \mathcal{L}_{\mathsf{LQP}}$, we have the following:*

$$\textit{if} \quad \vdash_{\mathsf{LQP}} \phi \quad \textit{then} \quad \mathcal{C}_s \models \phi \ .$$

Corollary 4.10.1 *(*LQP *Consistency).* LQP *is consistent:* $\nvdash_{\mathsf{LQP}} 0$.

[10] More accurately, these items show that the underlying logic is *at least* KD. Since we share the same definition of *model*, the theorem in [10][p. 183], showing that the logic is exactly KD, applies here.

Theorem 4.11 *(LQP Completeness). Let $\mathcal{C}_s$ be the class of simple models. For each $\phi \in \mathcal{L}_{\mathsf{LQP}}$, we have the following:*

$$\text{if} \quad \mathcal{C}_s \models \phi \quad \text{then} \quad \vdash_{\mathsf{LQP}} \phi \ .$$

Proof outline. We use the standard completeness proof method of showing that any consistent formula is satisfiable. Specifically, given a consistent formula $\neg\theta$, we use a modal filtration and Theorem 1.2[11] of [28] to obtain a simple model (W, P, V) that satisfies $\neg\theta$.

The filtration is used to define the finite set of worlds W (as the set of maximal consistent subsets of a finite set A generated from θ) and the valuation function V (mapping each world $w \in W$ to the set of proposition letters in w).

For each world $w \in W$, a probability space $(\Omega_w, 2^{\Omega_w}, P_w)$ is constructed as follows. First, the set $\Omega_w \subseteq W$ is the collection of worlds that each contain the formulas in A that can be proved from w to be probabilistically certain. We then define a probability function P_w from a linear functional f_w that we obtain by applying Scott's [28] Theorem as follows.

The basis S for the vector space $L(S)$ is the set W of worlds. We identify vectors in $L(W)$ with functions from W to $\mathbb{R}$. For X and N, we have local sets X_w and N_w defined as follows.

The set N_w is a collection of functions from W to $\mathbb{Z}$, each defined by sums and differences of characteristic functions $\iota([\phi]_w)$ of sets of worlds $[\phi]_w$ that satisfy certain formulas ϕ; in particular, a characteristic function $\iota([\phi]_w)$ is added each time ϕ is in an appropriate sequence Φ that is provable from w to be at least as likely as an appropriate other sequence, and $\iota([\phi]_w)$ is subtracted each time ϕ is in an appropriate sequence that is provable from w to be at most as likely as an appropriate other sequence

The set X_w is defined to be $N_w \cup (-N_w)$, guaranteeing that it is symmetric in $L(W)$ and that it inherits from N_w the properties of being finite and rational.

Then Scott's Theorem guarantees the existence of a linear functional $f_w : L(W) \to \mathbb{R}$ that realizes N_w in X_w, from which we construct the desired probability function P_w on subsets E of Ω_w by normalizing f_w to Ω_w as follows: $P_w(E) \doteq f_w(\iota(E))/f_w(\iota(\Omega_w))$.

The linearity of f_w guarantees that P_w is a probability function, and the fact that f_w realizes N_w in X_w helps to ensure that the constructed model M satisfies the original given formula $\neg\theta$.

See Appendix B for details.

4.3. Quantitative aspects of the theory

The results of the previous subsection are fundamentally qualitative in that, in the proof theory, probabilities, numbers, and arithmetic operators are not mentioned. Although we have used the symbols "1" and "0", these are just abbreviations for some fixed tautology and its negation, respectively. In this subsection we show how we can nonetheless encode seemingly quantitative notions, culminating with rational linear inequalities.

We proceed incrementally, adding expressivity with each new step.

- Define $k \cdot \phi$ (also written $k\phi$) by the recurrence relation:

  $$0 \cdot \phi \doteq 0 \qquad \text{and} \qquad (k+1) \cdot \phi \doteq \phi \oplus (k \cdot \phi).$$

  Example: $2\phi \preceq 3\psi$ abbreviates $\phi \oplus \phi \oplus 0 \preceq \psi \oplus \psi \oplus \psi \oplus 0$, which is equivalent to $\phi \oplus \phi \preceq \psi \oplus \psi \oplus \psi$.
- Negation in sequences is introduced by replacing a sequence element $k\phi$ by 0, and appending $-k\phi$ to the sequence on the other side of $\preceq$.
  Example: From $3\psi \oplus 2\phi \preceq \chi$ we obtain $0 \oplus 2\phi \preceq -3\psi \oplus \chi$, which is equivalent to $2\phi \preceq -3\psi \oplus \chi$.
- We introduce integer ratios p/q (where $q \neq 0$) for coefficients of sequence elements in an inequality: In the binary case, $(p/q) \cdot \phi \preceq (r/s) \cdot \psi$ abbreviates $(p \times s) \cdot \phi \preceq (r \times q) \cdot \psi$. (Note that, in contrast with "$\cdot$", the symbol "$\times$" is not a component of our encoding, but rather is a meta-level symbol denoting the regular multiplication of its two arguments.) In general

  $$\bigoplus_{i \leq n} \left(\frac{p_i}{q_i}\right) \cdot \phi_i \ \preceq\ \bigoplus_{j \leq m} \left(\frac{r_j}{s_j}\right) \cdot \psi_j$$

  abbreviates

  $$\bigoplus_{i \leq n} \left(t \times \frac{p_i}{q_i}\right) \cdot \phi_i \ \preceq\ \bigoplus_{j \leq m} \left(t \times \frac{r_j}{s_j}\right) \cdot \psi_j,$$

  where the product $t \doteq \prod_{i,j} q_i s_j$ and the integer values $t \times p_i/q_i$ and $t \times r_j/s_j$ are computed using the usual rules of arithmetic.
  Example: $0 \preceq \frac{1}{2}\psi \oplus -\frac{1}{3}\phi$ denotes $\frac{1}{3}\phi \oplus 0 \preceq \frac{1}{2}\psi \oplus 0$.

[11] Interestingly, Gärdenfors and Segerberg instead use Theorem 4.1 from [28].

The latter denotes $2 \cdot \phi \oplus 6 \cdot 0 \preceq 3 \cdot \psi \oplus 6 \cdot 0$, which is equivalent to $2\phi \preceq 3\psi$. This in turn is equivalent to $\phi \oplus \phi \preceq \psi \oplus \psi \oplus \psi$.

- Last, we identify $0 \in \mathbb{Q}$ with the formula $0 \in \mathcal{L}_{\mathsf{LQP}}$ and each nonzero rational number $x \in \mathbb{Q}$ with the formula $(p/q) \cdot 1$, where $p/q = x$, $q > 0$, and $\mathsf{gcd}(p, q) = 1$.
Example: Using the previous example, it can be shown that LQP derives $(2/3) \preceq 1$ and $0 \preceq (1/2) \oplus -(1/3)$.

It follows that the language of LQP can express every inequality between finite, rational-coefficient sums of rational numbers and formulas. In particular:

- for all non-negative integers n_a, n_b, n_c, and n_d such that $n_a + n_b > 0$ and $n_c + n_d > 0$;
- for all possibly empty sets $\{a_i\}_{i=1}^{n_a}$, $\{b_i\}_{i=1}^{n_b}$, $\{c_i\}_{i=1}^{n_c}$, and $\{d_i\}_{i=1}^{n_d}$ of rational numbers; and
- for all possibly empty sets $\{\phi_i\}_{i=1}^{n_b}$ and $\{\psi_i\}_{i=1}^{n_d}$ of formulas,

it follows that the following is a well-defined formula of our language:

$$\bigoplus_{i=1}^{n_a} a_i \oplus \bigoplus_{i=1}^{n_b} (b_i \cdot \phi_i) \preceq \bigoplus_{i=1}^{n_c} c_i \oplus \bigoplus_{i=1}^{n_d} (d_i \cdot \psi_i). \tag{2}$$

We refer to expressions of this form as *rational linear inequalities*. Based on the definition of our language, at least one of the sides of the inequality (2) must be nonempty, although we could take either side to be something as simple as 0.

Theorem 4.12 *(Quantitative Probability Principle). For $a_1, \dots, a_n, b_1, \dots, b_m \in \mathbb{Q}$:*

$$\sum_{i=1}^{n} a_i \;\le\; \sum_{i=1}^{m} b_i \quad \Leftrightarrow \quad \vdash_{\mathsf{LQP}} \bigoplus_{i=1}^{n} a_i \preceq \bigoplus_{i=1}^{m} b_i.$$

5. Related work

In this section we compare our approach with that of other related work. In the first subsection we discuss Gärdenfors's logic QP of qualitative probability, after which we briefly consider Segerberg's more general logic PK. In the next subsection, we look at the Fagin et al. logic of quantitative probability $\mathsf{AX}_{\mathsf{meas}}$. The final subsection briefly discusses other work dealing with notions related to qualitative probability. To focus the discussion, some details are deferred to Appendix A.

5.1. The Gärdenfors and Segerberg logics of qualitative probability

The language of Gärdenfors's QP is that of propositional logic augmented with a binary operator of comparative probability $\preceq$. As with our approach, the modal operator of probabilistic certainty $\Box$ is introduced by definition, $\Box\phi \doteq (1 \preceq \phi)$, and nesting of $\preceq$ is allowed. Recall that Gärdenfors's approach, following [29], hinges on the abbreviation schema

$$(\phi_i)_{i=1}^{m} \mathbb{E} (\psi_i)_{i=1}^{m} \quad \doteq \quad \Box \bigvee_{i=0}^{m} C_i$$

where C_i is the disjunction of all conjunctions

$$e_1\phi_1 \wedge \cdots \wedge e_m\phi_m \wedge f_1\psi_1 \wedge \cdots \wedge f_m\psi_m \tag{3}$$

satisfying the property that exactly i of the e_k's are the negation symbol $\neg$, exactly i of the f_k's are $\neg$, and the rest of the e_k's and f_k's are the empty string ϵ.

Gärdenfors gives the following axiomatisation, and shows that it is sound and complete with respect to the class of simple models.

Definition 5.1 *(QP).* QP is defined by the following:

(PC)	All tautologies of classical propositional logic
(A0)	$\Box(\phi_1 \equiv \phi_2) \wedge \Box(\psi_1 \equiv \psi_2) \supset ((\phi_1 \preceq \psi_1) \equiv (\phi_2 \preceq \psi_2))$
(A1)	$0 \preceq \phi$
(A2)	$(\phi \preceq \psi) \vee (\psi \preceq \phi)$
(A3)	$0 \prec 1$
(A4)	$((\phi_i)_{i=1}^{m} \mathbb{E} (\psi_i)_{i=1}^{m} \wedge \bigwedge_{k=1}^{m-1} (\phi_k \preceq \psi_k)) \supset (\psi_m \preceq \phi_m)$ for all $m \ge 1$
(MP)	From $\phi \supset \psi$ and ϕ, infer ψ
(Nec)	From ϕ, infer $\Box\phi$

With the exception of (A4), the above axioms and rules of inference are clearly contained in LQP. The next result relates the Segerberg and Gärdenfors $\mathbb{E}$ notation to our sequences. The first part gives the relation between the $\mathbb{E}$ schema and sequences, while the second part shows that the key Segerberg/Gärdenfors axiom (A4) is derivable in our approach.

Theorem 5.2 *($\mathbb{E}$-Schema and Sequences). For formulas ϕ_i and ψ_i, $1 \le i \le m$, we have:*

1. $\vdash_{\mathsf{LQP}} (\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m \supset \left(\bigoplus_{i=1}^m \phi_i \approx \bigoplus_{i=1}^m \psi_i\right)$
2. $\vdash_{\mathsf{LQP}} ((\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m \wedge \bigwedge_{k=1}^{m-1} (\phi_k \preceq \psi_k)) \supset (\psi_m \preceq \phi_m)$

Since the language of QP is that of LQP but where $\preceq$ is a binary relation on formulas only, it is not surprising that QP is less expressive than LQP:

Theorem 5.3 *(QP Expressivity). LQP is strictly more expressive than QP.*

For example $p \oplus q \preceq r$ denotes a satisfiable $\mathcal{L}_{\mathsf{LQP}}$ formula that is inexpressible in QP. Moreover, noting that $\mathcal{L}_{\mathsf{QP}} \subset \mathcal{L}_{\mathsf{LQP}}$ and making use of the fact that QP and LQP are both sound and complete with respect to the class of simple models, we obtain:

Theorem 5.4. *For any $\phi \in \mathcal{L}_{\mathsf{QP}}$, if $\vdash_{\mathsf{QP}} \phi$ then $\vdash_{\mathsf{LQP}} \phi$.*

Corollary 5.4.1. *LQP is a conservative extension of QP.*

We next consider the earlier logic PK of Segerberg [29]. In contrast to QP and LQP, PK does not allow nested occurrences of the $\preceq$ operator. On the other hand, also in contrast to QP and LQP, PK has a distinct operator of modal necessity and allows formulas with non-measurable truth sets. For those readers that may be interested, the Segerberg axiomatisation is given in Appendix A. We have the following relations between PK, QP, and LQP.

Theorem 5.5 *($\mathcal{L}_{\mathsf{PK}}$ Expressivity). Let $\mathcal{L}^-_{\mathsf{QP}}$ be the fragment of $\mathcal{L}_{\mathsf{QP}}$ obtained by deleting all formulas that contain nesting of $\preceq$'s.*

1. *$\mathcal{L}_{\mathsf{PK}}$ is strictly more expressive than $\mathcal{L}^-_{\mathsf{QP}}$ over the class of super-simple PK models.*
2. *$\mathcal{L}_{\mathsf{PK}}$ and $\mathcal{L}_{\mathsf{QP}}$ are expressively incomparable over the class of super-simple PK models.*
3. *$\mathcal{L}_{\mathsf{PK}}$ and $\mathcal{L}_{\mathsf{LQP}}$ are expressively incomparable over the class of super-simple PK models.*

In the restriction of PK in which the modal operators of probabilistic and epistemic necessity coincide and where all formulas have measurable truth sets, PK is the same logic as QP over the language with unnested occurrences of $\preceq$. Let PK′ be this restriction of PK (again, see Appendix A for details). We obtain:

Theorem 5.6. *For any $\phi \in \mathcal{L}^-_{\mathsf{QP}}$ we have: $\vdash_{\mathsf{QP}} \phi$ iff $\vdash_{\mathsf{PK}'} \phi$*

Corollary 5.6.1. *LQP is strictly more expressive than PK′.*

Hence LQP is a conservative extension of PK′.

5.2. The Fagin et al. logic of quantitative probability

A different approach to reasoning about probability is the theory $\mathsf{AX}_{\mathrm{meas}}$ of *quantitative* probability of [9]. Their language permits Boolean combinations of linear inequalities[12] of the form $c \preceq a_1 w(\phi_1) + \cdots + a_n w(\phi_n)$, where c and the a_i's are integers and the ϕ_i's do not contain $\preceq$'s or $w(\chi)$'s. The expression $w(\phi)$ is mapped in the semantics to a real number called the "weight" of ϕ, where the weight ends up being the probability of the event ϕ.

The axiomatisation of $\mathsf{AX}_{\mathrm{meas}}$ is given as follows:

Definition 5.7 *(AX_{meas}).* $\mathsf{AX}_{\mathrm{meas}}$ is defined by the following:

(PC)	All tautologies of classical propositional logic
(Ineq)	All instances of valid formulas about linear inequalities
(W1)	$0 \preceq w(\phi)$
(W2)	$1 \approx w(1)$
(W3)	$w(\phi \wedge \psi) + w(\phi \wedge \neg\psi) \approx w(\phi)$
(W4)	$w(\phi) \approx w(\psi)$, where $\phi \equiv \psi$ is a tautology
(MP)	From $\phi \supset \psi$ and ϕ, infer ψ

[12] Fagin et al. use $\geq$; for uniformity we remain with $\preceq$.

Fagin et al. also show that (Ineq) can be replaced by a set of explicit schemas that derive the same theorems. In common with QP and PK but in contrast with LQP, this axiomatisation is not schematically finite. $\mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}}$ does not allow nesting of $\preceq$'s in formulas; for example, $1 \preceq (1 \preceq w(p))$ is not a formula in $\mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}}$. To compare the expressivity of $\mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}}$ with $\mathcal{L}_{\mathsf{LQP}}$, we consider the fragment $\mathcal{L}^{-}_{\mathsf{LQP}}$ of $\mathcal{L}_{\mathsf{LQP}}$ that excludes nesting of $\preceq$'s. We obtain:

Theorem 5.8.

1. *$\mathcal{L}_{\mathsf{AX}_{meas}}$ and $\mathcal{L}^{-}_{\mathsf{LQP}}$ are equally expressive over the class of simple models.*[13]
2. *$\mathcal{L}_{\mathsf{LQP}}$ is strictly more expressive than $\mathcal{L}_{\mathsf{AX}_{meas}}$ over the class of super-simple models.*

Thus, although $\mathsf{AX}_{\mathrm{meas}}$ is a logic of quantitative probability, we can express it in LQP. However, it should be noted that $\mathsf{AX}_{\mathrm{meas}}$ is significantly more succinct that LQP. Specifically, a number n can be represented in $\mathsf{AX}_{\mathrm{meas}}$ with $\ln(n)$ bits whereas it would take on the order of n bits in LQP.

An extension of $\mathsf{AX}_{\mathrm{meas}}$ is studied by [8]. The language of this extension is obtained by adding a set of S5 modalities, to give a multi-agent epistemic logic, and allowing nesting of inequalities. While the single-agent, non-S5 fragment of their extended theory would presumably correspond with LQP, we leave an exploration of this correspondence to future work.

5.3. Likelihood

As discussed in Section 3, there has been considerable work addressing a notion of *probably* or *likely*. The former is usually interpreted as having probability of at least (or, strictly greater than) .5. The latter notion, of likelihood, may or may not be given a probabilistic interpretation. Our view is that likelihood is best regarded as a probabilistic notion. It may be difficult to assign a probability to a given event (say, that there is an accident involving two busses in Vancouver tomorrow in which no-one is hurt) but that doesn't mean that, given a sufficiently nuanced theory, that such a probability cannot be determined in principle. As a result we can interpret "ϕ is likely" as $\neg\phi \preceq \phi$ (or perhaps $\neg\phi \prec \phi$, depending on one's taste), that is, that the probability is at least (resp. greater than) .5. Then "ϕ is very likely" can be expressed by $\neg\phi \preceq c \cdot \phi$ for some fixed integer $c > 0$, using the convention of Section 4.3.

5.4. Discussion

To recap the results of this section: Gärdenfors's and Segerberg's logics include an infinite number of complex and (in our opinion) unintuitive counting schemes, as expressed in their $\mathbb{E}$ notation. Fagin et al.'s logic seems simple at first glance—just classical logic, Kolmogorov's axioms, and an axiom (Ineq) concerning valid linear inequalities. However, (Ineq) hides a substantial amount of logic about integers, inequalities and arithmetic (see Theorem A.7 in Appendix A). Our setting has the advantage that there are finitely many schemes, and each scheme is arguably intuitive and expresses a simple logical principle.

$\mathcal{L}_{\mathsf{LQP}}$ is expressively incomparable with Segerberg's logic PK. Like Gärdenfors's $\mathcal{L}_{\mathsf{QP}}$, Segerberg's $\mathcal{L}_{\mathsf{PK}}$ cannot express the $\mathcal{L}_{\mathsf{LQP}}$-formula $p \oplus q \preceq r$. On the other hand, $\mathcal{L}_{\mathsf{PK}}$ includes a Kripke-style necessity operator, which we expressed as $\boxdot$. The meaning of $M, w \models \boxdot p$ is that $M, v \models p$ at all worlds v contained in some possibly strict superset $R(w) \supseteq \Omega_w$ of the set Ω_w of outcomes at w. As a result, $\boxdot p$ is not expressible in $\mathcal{L}_{\mathsf{LQP}}$, even if we restrict to the class of super-simple models. Therefore we must extend our language $\mathcal{L}_{\mathsf{LQP}}$ to a language containing the $\boxdot$ operator in order to regain expressive comparability. On doing this, we would find that the resulting language is strictly more expressive than both $\mathcal{L}_{\mathsf{PK}}$ and $\mathcal{L}_{\mathsf{LQP}}$ over the class of super-simple models (to which of course we add an accessibility function $R : W \to 2^W$ satisfying $\Omega_w \subseteq R(w)$ for each $w \in W$).

With regards to $\mathsf{AX}_{\mathrm{meas}}$ and LQP, we have shown that there is a translation between these two approaches. Therefore, except for the fact that we allow nesting of the operator $\preceq$, the two theories are in a sense "equivalent" from the point of view of derivability. However, the theories differ in a key respect. Notably, $\mathsf{AX}_{\mathrm{meas}}$ is a *quantitative* approach, in that integers, and the operations of addition and multiplication, are assumed to be given a priori. In contrast, we have presented a foundational, *qualitative* approach, beginning from first principles and adopting a minimal set of underlying assumptions. An analogy might be made with the notion of computation: On the one hand, Turing machines provide a formal model for studying fundamental notions involving computation, what can and cannot be computed, and the like. On the other hand, a Turing machine is an impractical model to actually get anything done; rather one would use some higher-level (Turing-equivalent) programming language.

6. Conclusion

In this paper we have addressed the foundations of qualitative probability. While work in this area goes back many years, we have argued that no approach has provided a wholly satisfactory characterisation of qualitative probability. In

[13] That is, there are functions $G : \mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}} \to \mathcal{L}^{-}_{\mathsf{LQP}}$ and $H : \mathcal{L}^{-}_{\mathsf{LQP}} \to \mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}}$ that are satisfaction-preserving over the class of simple models, and in which $G \circ H$ and $H \circ G$ are both the identity function.

earlier work, a binary operator $\preceq$ on formulas (or "events") is given, describing the relation "is no more probable than". The central intuition and innovation of our approach is that $\preceq$ should be regarded as an operator on finite sequences of formulas, with intended interpretation that the combined probabilities of the formulas on the left hand side of $\preceq$ does not exceed that of the formulas on the right hand side.

The resulting logic, LQP, is a general, intuitive formalism for reasoning about the full gamut of qualitative and axiomatic probability. The axiomatisation is finite and intuitive, where each axiom captures a simple logical principle. Consequently, the approach provides a foundational theory of qualitative probability, in that the proof theory adopts minimal assumptions, specifying properties of sequences and the operator $\preceq$. Nonetheless, the approach is expressive enough to encode rational linear inequalities. By identifying 0 with an inconsistent formula and 1 with a tautology, we can build other more complex concepts, ending with rational linear inequalities. Hence our approach can be viewed as specifying a foundation for qualitative probability that nonetheless provides a "bridge" to quantitative approaches.

We give a completeness result for the probability interpretation. Our construction is an adaptation of one due to Lenzen [23]. Overall, our work is part of a recent renewed interest in modal logics for qualitative probability that includes [16,18,30], which in turn can be seen as contributing to ongoing work on the combination of logic with probability [1,27].

We have compared this approach to previous work, and we have shown that ours subsumes those theories for finite outcome sets. Our system captures qualitative probabilistic reasoning, axiomatic probability, the major qualitative system due to Gärdenfors, and the major quantitative system due to [9]. A question as to the extension of LQP to the class of general models is left for future work, as is extension to include a modal operator of epistemic belief.

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgements

James Delgrande gratefully acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada grant number RGPIN-2018-05983. Bryan Renne was funded by Veni grant 275-20-030 from the Netherlands Organisation for Scientific Research (NWO) hosted January 2012 through December 2014 by the Institute for Logic, Language, Information and Computation (ILLC) at the University of Amsterdam. We thank the reviewers for their helpful comments.

Appendix A. Related work

A.1. The Segerberg logic of qualitative probability

We review here the details of Segerberg's logic PK. The language of Segerberg [29] is that of propositional logic augmented with the binary operator of comparative probability $\preceq$ and a unary operator of necessitation. The language does not allow nesting of $\preceq$. We write the necessity operator as $\boxdot$, where $\boxdot\phi$ asserts that ϕ is true in every accessible world. This notation is to distinguish this operator and the modal operator $\Box$ which is used to describe probabilistic certainty (i.e., $\Box\phi \doteq 1 \preceq \phi$). We define the abbreviation $(\phi_i)_{i=1}^m \dot{\mathbb{E}} (\psi_i)_{i=1}^m$ (notice the dot over the $\mathbb{E}$) just as we defined $(\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m$, except that $\boxdot$ is used in place of $\Box$.

Definition A.1 *(PK; [29]).* PK is defined by the following axiom schemes and rules:

(PC)	All tautologies of classical propositional logic
(#0)	$\boxdot(\phi \supset \psi) \supset (\boxdot\phi \supset \boxdot\psi)$
(#1)	$\boxdot(\phi \equiv \phi') \wedge \boxdot(\psi \equiv \psi') \supset ((\phi \preceq \psi) \supset (\phi' \preceq \psi'))$
(#2)	$0 \preceq 0$
(#3)	$(\phi \preceq \phi) \wedge (\psi \preceq \psi) \supset ((\phi \supset \psi) \preceq (\phi \supset \psi))$
(#4)	$(\phi \preceq \psi) \supset (\phi \preceq \phi)$
(#5)	$(\phi \preceq \psi) \supset (\psi \preceq \psi)$
(#6)	$(\phi \preceq \phi) \wedge (\psi \preceq \psi) \supset ((\phi \preceq \psi) \vee (\psi \preceq \phi))$
(#7)	$(\phi_i)_{i=1}^m \dot{\mathbb{E}} (\psi_i)_{i=1}^m \supset ((\bigwedge_{i=1}^m \phi_i \preceq \psi_i) \supset (\bigwedge_{i=1}^m \psi_i \preceq \phi_i))$ for all $m \geq 1$
(#8)	$0 \prec 1$
(#9)	$(\phi \preceq \phi) \supset (0 \preceq \phi)$
(MP)	From $\phi \supset \psi$ and ϕ, infer ψ
(Nec)	From ϕ infer $\boxdot\phi$

Definition A.2 *(PK model).* A PK *model*[14] is a structure $M = (W, \Pi, R, V)$ such that

- (W, Π, V) is a model (Definition 4.2); and
- $R : W \to 2^W$ is an *accessibility function* satisfying $\Omega_w \subseteq R(w)$ for each $w \in W$.

The function $[\![\cdot]\!] : \mathcal{L}_{\mathsf{PK}} \to 2^W$ and the relation $\models$ between pointed PK models and $\mathcal{L}_{\mathsf{PK}}$-formulas is as expected; Definition 4.3 gives all cases except for epistemic necessity:

$$M, w \models \boxdot\phi \text{ iff } M, v \models \phi \quad \text{for each} \quad v \in R(w).$$

Segerberg shows that PK is sound and complete with respect to the class $\mathcal{C}_{\mathsf{PK}}$ of PK-models. He also suggests that the logic of the modal operator $\boxdot$ is K; however as the next results show, the logic in fact is KD.

Theorem A.3 *(PK Seriality). For any PK model $M = (W, \Pi, R, V)$ the relation R is serial.*

Theorem A.4 *(D).* $\vdash_{\mathsf{PK}} \boxdot\phi \supset \neg\boxdot\neg\phi$

The restriction PK′ of PK (used in Theorem 5.6) is defined as follows:

Definition A.5. Define PK′ to be the logic consisting of the axiom schemes and rules of PK together with the axioms:

(PMN)	$(1 \preceq \phi) \supset \boxdot\phi$
(Symm)	$\phi \preceq \phi$

A.2. The Fagin et al. logic of quantitative probability

The language of [9] is defined as follows:

Definition A.6. Let $\mathcal{P}$ be a nonempty set of propositional atoms. The language $\mathcal{L}_{\mathsf{AX}_{\text{meas}}}$ consists of *formulas* ϕ and *weight terms* W given by the following recursion:

$$\begin{array}{lll} \phi & ::= & p \mid \neg\phi \mid (\phi \vee \phi) \mid (a \preceq W) \quad p \in \mathcal{P},\ a \in \mathbb{Z} \\ W & ::= & w(\psi) \mid a \cdot w(\psi) \mid W + W \quad a \in \mathbb{Z} \\ \psi & ::= & p \mid \neg\phi \mid (\phi \vee \phi) \qquad\qquad p \in \mathcal{P} \end{array}$$

The explicit specification of the axiom (Ineq) is given as follows:

Theorem A.7 *([9]). Replacing (Ineq) in the theory AX_{meas} (from Definition 5.7) with the following schemes yields a theory that derives the same theorems:*

(I1)	$w(\phi) \preceq w(\phi)$
(I2)	$(c \preceq a_1 w(\phi_1) + \cdots + a_k w(\phi_k)) \equiv (c \preceq a_1 w(\phi_1) + \cdots + a_k w(\phi_k) + 0 w(\phi_{k+1}))$
(I3)	$(c \preceq a_1 w(\phi_1) + \cdots + a_k w(\phi_k)) \equiv (c \preceq a_{j_1} w(\phi_{j_1}) + \cdots + a_{j_k}(\phi_{j_k}))$ *with* $j_1, \ldots, j_k$ *a permutation of* $1, \ldots, k$
(I4)	$(c \preceq a_1 w(\phi_1) + \cdots + a_k w(\phi_k)) \wedge (c' \preceq a'_1 w(\phi_1) + \cdots + a'_k w(\phi_k)) \supset ((c + c') \preceq (a_1 + a'_1) w(\phi_1) + \cdots + (a_k + a'_k) w(\phi_k))$
(I5)	$(c \preceq a_1 w(\phi_1) + \cdots + a_k w(\phi_k)) \supset (dc \preceq da_1 w(\phi_1) + \cdots + da_k w(\phi_k))$ *with* $d > 0$
(I6)	$(c \preceq a_1 w(\phi_1) + \cdots + a_k w(\phi_k)) \vee (a_1 w(\phi_1) + \cdots + a_k w(\phi_k) \preceq c)$
(I7)	$(c \preceq a_1 w(\phi_1) + \cdots + a_k w(\phi_k)) \supset (d \prec a_1 w(\phi_1) + \cdots + a_k w(\phi_k))$ *with* $d < c$

Appendix B. Proofs

B.1. Proofs of Section 4

Proof of Theorem 4.4. Since M is total, we have $\Omega_v = W$. Hence we have for every formula χ that $[\![\chi]\!]_M^v = [\![\chi]\!]_M \cap W = [\![\chi]\!]_M$. Further, since M is powerset, we have $[\![\chi]\!]_M \in F_v$ for every χ. Finally, since M is uniform, we have $P_w = P_v$. Applying Definition 4.3, the result follows.

Proof of Theorem 4.6. For each item, we reason in LQP. We consider each item in turn.

[14] Segerberg's notation has been adapted to conform to our's. The correspondence with [29] is as follows. [29] defines a model to be a structure $\mathcal{U}^S = (U^S, R^S, V^S, B^S, M^S)$ (where the superscript S is added to remove ambiguity). For PK-model $M = (W, \Pi, R, V)$ where $\Pi_w = (\Omega_w, F_w, P_w)$ we have the correspondences: $M \leftrightarrow \mathcal{U}^S$, $W \leftrightarrow U^S$, $V \leftrightarrow V^S$, $R \leftrightarrow R^S$, $\Omega_w \leftrightarrow \{w' \mid wR^S w'\}$, $F_w \leftrightarrow B^S_w$, $P_w \leftrightarrow M^S_w$.

1: An instance of (Tot) is $(\Phi \preceq \Phi) \vee (\Phi \preceq \Phi)$. Via propositional logic we obtain $\Phi \preceq \Phi$.
2: Reflexivity of $\approx$ is an immediate consequence of the reflexivity of $\preceq$ and the definition of $\approx$. For symmetry we have that $\Phi \approx \Psi$ iff $(\Phi \preceq \Psi) \wedge (\Psi \preceq \Phi)$ iff $(\Psi \preceq \Phi) \wedge (\Phi \preceq \Psi)$ iff $\Psi \approx \Phi$. Transitivity is a simple consequence of (Tran).

To ease proofs, we will freely make use of the symmetry of $\approx$ without further comment.

3: Assume that we have $\Phi \preceq \Psi$. An instance of (Ref) is $\Delta \preceq \Delta$, and applying (Add) to these inequalities yields $\Delta \oplus \Phi \preceq \Delta \oplus \Psi$.
4: Assume that $\Box(\phi_1 \equiv \phi_2)$ and $\Box(\psi_1 \equiv \psi_2)$; as well assume that we have $\phi_1 \preceq \psi_1$. An instance of (Ref) is $1 \preceq 1$ and applying (Add) to $\phi_1 \preceq \psi_1$ and $1 \preceq 1$ yields $1 \oplus \phi_1 \preceq 1 \oplus \psi_1$. Now using (Sub) yields $1 \oplus \phi_2 \preceq 1 \oplus \psi_2$ and finally (Succ) gives $\phi_2 \preceq \psi_2$. The reverse direction for the equivalence follows from the symmetry between $(\phi_1 \preceq \psi_1)$ and $(\phi_2 \preceq \psi_2)$ in the statement of the theorem.
5: Assume that $\Box(\phi_1 \equiv \phi_2)$ and $\Box(\psi_1 \equiv \psi_2)$; and assume that we have $(\Phi_1 \preceq \Psi_1)$. Φ_1 can be written as $\Phi_1' \oplus \phi_1 \oplus \Phi_1''$ and Ψ_1 can be written as $\Psi_1' \oplus \psi_1 \oplus \Psi_1''$. (The case where ϕ_1 or ψ_1 are elements at the "end" of a sequence is handled by a trivial modification to the following argument.) Since we have $\Phi_1' \oplus \phi_1 \oplus \Phi_1'' \preceq \Psi_1' \oplus \psi_1 \oplus \Psi_1''$, two applications of (Comm) yield $\phi_1 \oplus \Phi_1'' \oplus \Phi_1' \preceq \psi_1 \oplus \Psi_1'' \oplus \Psi_1'$. From (Sub) we obtain that $\phi_2 \oplus \Phi_1'' \oplus \Phi_1' \preceq \psi_2 \oplus \Psi_1'' \oplus \Psi_1'$ and two further applications of (Comm) give $\Phi_1' \oplus \phi_2 \oplus \Phi_1'' \preceq \Psi_1' \oplus \psi_2 \oplus \Psi_1''$, which was to be shown.
6: For the first statement, assume $\Phi_1 \approx \Phi_2$ and $\Phi_1 \oplus \Phi \preceq \Psi$. By the definition of $\approx$ and classical reasoning, it follows from $\Phi_1 \approx \Phi_2$ that we have $\Phi_2 \preceq \Phi_1$. Further, by (Ref) we have $\Phi \preceq \Phi$. Therefore, from $\Phi_2 \preceq \Phi_1$ and $\Phi \preceq \Phi$, we have by Part 3 of this theorem that $\Phi_2 \oplus \Phi \preceq \Phi_1 \oplus \Phi$. But then it follows from $\Phi_2 \oplus \Phi \preceq \Phi_1 \oplus \Phi$ and our assumption $\Phi_1 \oplus \Phi \preceq \Psi$ by (Tran) and classical reasoning that we have $\Phi_2 \oplus \Phi \preceq \Psi$. That is, we have shown that $\Phi_1 \approx \Phi_2$ and $\Phi_1 \oplus \Phi \preceq \Psi$ together imply $\Phi_2 \oplus \Phi \preceq \Psi$. By a similar argument, it follows that $\Phi_1 \approx \Phi_2$ and $\Phi_2 \oplus \Phi \preceq \Psi$ together imply $\Phi_1 \oplus \Phi \preceq \Psi$. Consequently, we get that $\Phi_1 \approx \Phi_2$ implies that $\Phi_1 \oplus \Phi \preceq \Psi \equiv \Phi_2 \oplus \Phi \preceq \Psi$, which is the first part of the Replacement Principle. The second part is shown by a similar argument.
7: Assume that we have $\Delta \oplus \Phi \preceq \Delta \oplus \Psi$; and assume further that $|\Delta| = 1$. We can then write our initial assumption as $\phi \oplus \Phi \preceq \phi \oplus \Psi$ where $\phi = \Delta$. By (Ref), we have $\neg\phi \preceq \neg\phi$. It follows by (Add) and classical reasoning that

$$\neg\phi \oplus \phi \oplus \Phi \preceq \neg\phi \oplus \phi \oplus \Psi. \tag{4}$$

We have that $\neg\phi \vee \phi$ is a tautology and therefore that $\Box(\neg\phi \vee \phi)$ by (Nec). It follows from $\Box(\neg\phi \vee \phi)$ by (K3) and classical reasoning that $\neg\phi \oplus \phi \approx \neg\phi \vee \phi$. We also have $(\neg\phi \vee \phi) \equiv 1$ by classical reasoning (recalling that 1 is just some fixed tautology) and therefore that $\Box((\neg\phi \vee \phi) \equiv 1)$ by (Nec). But from $\Box((\neg\phi \vee \phi) \equiv 1)$ and $\neg\phi \oplus \phi \approx \neg\phi \vee \phi$ it follows by (Sub) and classical reasoning that $\neg\phi \oplus \phi \approx 1$. It follows from $\neg\phi \oplus \phi \approx 1$ and (4) by classical reasoning and Part 6 of this theorem that $1 \oplus \Phi \preceq 1 \oplus \Psi$. Applying (Succ) and classical reasoning, we obtain $\Phi \preceq \Psi$.
This shows our result for $|\Delta| = 1$; the result for arbitrary Δ then follows by a straightforward inductive argument.
8: Assume that $\Phi_1 \preceq \Psi_1$ and $\Psi_1 \oplus \Psi_2 \preceq \Phi_1 \oplus \Phi_2$. From $\Phi_1 \preceq \Psi_1$ and Part 3 it follows that $\Phi_1 \oplus \Psi_2 \preceq \Psi_1 \oplus \Psi_2$. From this and our assumption $\Psi_1 \oplus \Psi_2 \preceq \Phi_1 \oplus \Phi_2$ it follows by (Tran) and classical reasoning that $\Phi_1 \oplus \Psi_2 \preceq \Phi_1 \oplus \Phi_2$. Applying Part 7 and classical reasoning, we obtain $\Psi_2 \preceq \Phi_2$. $\Box$

Proof of Theorem 4.7.
For each item, we reason in LQP. We consider each item in turn.

1: We obtain by classical reasoning that $\neg(\phi \wedge 0)$ is a tautology (recalling that 0 is the negation of some tautology), from which it follows by (Nec) that $\Box\neg(\phi \wedge 0)$. From this it follows by (K3) and classical reasoning that $\phi \oplus 0 \approx \phi \vee 0$. By classical reasoning, we have $(\phi \vee 0) \equiv \phi$, from which it follows by (Nec) that $\Box((\phi \vee 0) \equiv \phi)$. But from $\Box((\phi \vee 0) \equiv \phi)$ and $\phi \oplus 0 \approx \phi \vee 0$ it follows by (Sub) and classical reasoning that $\phi \oplus 0 \approx \phi$.
2: Assume $\phi \oplus \psi \approx 0$ and therefore that $\psi \oplus \phi \approx 0$ by (Com) and classical reasoning. It follows by Part 1 of this theorem (specifically, that $0 \approx 0 \oplus 0$), (Tran), and classical reasoning that $\psi \oplus \phi \approx 0 \oplus 0$, and consequently that $\psi \oplus \phi \preceq 0 \oplus 0$. We also have $0 \preceq \psi$ by (K1). But from $\psi \oplus \phi \preceq 0 \oplus 0$ and $0 \preceq \psi$ we obtain by Theorem 4.6.8 that $\phi \preceq 0$. Since we also have $0 \preceq \phi$ by (K1), it follows by the meaning of $\approx$ and classical reasoning that $\phi \approx 0$.
3: We obtain by classical reasoning that $\neg((\phi \wedge \psi) \wedge (\phi \wedge \neg\psi)$ is a tautology, from which it follows by (Nec) that $\Box\neg((\phi \wedge \psi) \wedge (\phi \wedge \neg\psi)$. From this it follows by (K3) and classical reasoning that

$$(\phi \wedge \psi) \oplus (\phi \wedge \neg\psi) \approx (\phi \wedge \psi) \vee (\phi \wedge \neg\psi). \tag{5}$$

By classical reasoning, we also obtain that $((\phi \wedge \psi) \vee (\phi \wedge \neg\psi)) \equiv \phi$; call this formula χ. It follows from χ by (Nec) that $\Box\chi$. But from $\Box\chi$ and (5), it follows by (Sub) and classical reasoning that $(\phi \wedge \psi) \oplus (\phi \wedge \neg\psi) \approx \phi$ and so $\phi \approx (\phi \wedge \psi) \oplus (\phi \wedge \neg\psi)$, as desired.
4: A theorem of propositional logic is $\neg(\phi \wedge (\neg\phi \wedge \psi))$, from which via (Nec) we obtain $\Box(\neg(\phi \wedge (\neg\phi \wedge \psi)))$. From the latter formula, (K3), and modus ponens we obtain that $\phi \oplus (\neg\phi \wedge \psi) \approx \phi \vee (\neg\phi \wedge \psi)$. Since $(\phi \vee (\neg\phi \wedge \psi)) \equiv (\phi \vee \psi)$, an application of (Nec) together with substitution of necessary equivalents in the preceding formula gives $\phi \oplus (\neg\phi \wedge \psi) \approx (\phi \vee \psi)$.

An application of Theorem 4.6.3 to $\phi \oplus (\neg\phi \wedge \psi) \approx (\phi \vee \psi)$ and using commutativity gives

$$\phi \oplus (\neg\phi \wedge \psi) \oplus (\phi \wedge \psi) \approx (\phi \vee \psi) \oplus (\phi \wedge \psi). \tag{6}$$

We have from Part 3 that $\psi \approx (\neg\phi \wedge \psi) \oplus (\phi \wedge \psi)$ and so applying Theorem 4.6.5 to (6) gives $\phi \oplus \psi \approx (\phi \vee \psi) \oplus (\phi \wedge \psi)$, which was to be shown.

5: We have $\phi \oplus \neg\phi \approx (\phi \vee \neg\phi) \oplus (\phi \wedge \neg\phi)$ by Part 4. Since we have $\Box((\phi \vee \neg\phi) \equiv 1)$ and $\Box((\phi \wedge \neg\phi) \equiv 0)$ by classical reasoning and (Nec), it follows by (Sub) and classical reasoning that $\phi \oplus \neg\phi \approx 1 \oplus 0$. Since we also have $0 \preceq \neg\phi$ by (K1), it follows by the definition of $\approx$, Theorem 4.6.8, and classical reasoning that $\phi \preceq 1$.

6 We have by classical reasoning and (Nec) that $\Box\neg(\phi \wedge \neg\phi)$, from which it follows by (K3) and classical reasoning that $\phi \oplus \neg\phi \approx \phi \vee \neg\phi$. Further, by classical reasoning and (Nec) we have $\Box((\phi \vee \neg\phi) \equiv 1)$, from which it follows by $\phi \oplus \neg\phi \approx \phi \vee \neg\phi$, (Sub), and classical reasoning that $\phi \oplus \neg\phi \approx 1$.

7: From Part 4 we have that $\phi \oplus \psi \approx (\phi \vee \psi) \oplus (\phi \wedge \psi)$. From Part 1 we have that $\psi \approx \psi \oplus 0$, and via (Ref) and (Add) we obtain $\phi \oplus \psi \approx \phi \oplus \psi \oplus 0$. Applying (Tran) to this and our original formula yields

$$\phi \oplus \psi \oplus 0 \approx (\phi \vee \psi) \oplus (\phi \wedge \psi).$$

We have $0 \preceq (\phi \wedge \psi)$ by (K1); from this and our previous formula, it follows by (Com), the definition of $\approx$, Theorem 4.6.8, and classical reasoning that $\phi \vee \psi \preceq \phi \oplus \psi$.

8: An instance of Theorem 4.6.8 is $(\phi \oplus \neg\phi \preceq \psi \oplus \neg\psi) \supset (\psi \preceq \phi \supset \neg\phi \preceq \neg\psi)$. The antecedent condition $(\phi \oplus \neg\phi \preceq \psi \oplus \neg\psi)$ is easily shown to be a theorem: from (K3) and Theorem 4.6.6 we obtain that the antecedent is equivalent to $(\phi \vee \neg\phi \preceq \psi \vee \neg\psi)$ which again from Theorem 4.6.6 is equivalent to $1 \preceq 1$, which is an instance of (Ref). Consequently, applying modus ponens to the original formula, we obtain that $\psi \preceq \phi \supset \neg\phi \preceq \neg\psi$. $\Box$

Proof of Theorem 4.8.

1: Assume that $\Box(\phi \supset \psi)$ or, applying the definition of $\Box$, that $1 \preceq (\phi \supset \psi)$. We have by classical reasoning and (Nec) that $\Box((\phi \supset \psi) \equiv (\neg\phi \vee \psi))$, from which it follows by $1 \preceq (\phi \supset \psi)$, (Sub), and classical reasoning that $1 \preceq (\neg\phi \vee \psi)$. But we have $(\neg\phi \vee \psi) \preceq \neg\phi \oplus \psi$ by Theorem 4.7.7. As well, Theorem 4.7.6 is $\neg\phi \oplus \phi \approx 1$. We therefore have

$$\neg\phi \oplus \phi \approx 1, \quad 1 \preceq (\neg\phi \vee \psi), \quad \text{and} \quad (\neg\phi \vee \psi) \preceq \neg\phi \oplus \psi.$$

It follows by the definition of $\approx$, (Tran), and classical reasoning that $\neg\phi \oplus \phi \preceq \neg\phi \oplus \psi$. Applying Theorem 4.6.7, we obtain $\phi \preceq \psi$.

2: We assume $\Box(\phi \supset \psi)$ and $\Box\phi$. It follows from $\Box(\phi \supset \psi)$ by Part 1 and modus ponens that $\phi \preceq \psi$. From $\Box\phi$ we obtain by the definition of $\Box$ that $1 \preceq \phi$. Applying (Tran) to $1 \preceq \phi$ and $\phi \preceq \psi$, we obtain that $1 \preceq \psi$; that is, $\Box\psi$.

3: We have $\neg(1 \preceq 0)$ by (Triv), the definition of $\prec$, and classical reasoning. But $\neg(1 \preceq 0)$ is just $\neg\Box 0$.

4: Assume that $\Box\phi$. It follows by modal reasoning (using Parts 2 and 3) that $\Box((\psi \wedge \neg\phi) \equiv 0)$. From Theorem 4.7.3, we have that $\psi \approx (\psi \wedge \phi) \oplus (\psi \wedge \neg\phi)$ is a theorem. Applying (Sub) using the above necessitation to this formula yields $\psi \approx (\psi \wedge \phi) \oplus 0$. An instance of Theorem 4.7.1 is $(\psi \wedge \phi) \approx (\psi \wedge \phi) \oplus 0$. From (Tran) and classical reasoning we obtain that $\psi \approx \psi \wedge \phi$. $\Box$

Proof of Theorem 4.9.

1: The proof is by induction on n. In the base case, with $n = 2$, the formula is $\Box\neg(\phi_1 \wedge \phi_2) \supset (\phi_1 \oplus \phi_2 \approx \phi_1 \vee \phi_2)$ which is just Axiom (K3).
For the induction step, we assume the result holds for $n - 1$ and show that the result holds for n. So assume that $\bigwedge_{1 \le i \ne j \le n} \Box\neg(\phi_i \wedge \phi_j)$. It follows that $\bigwedge_{1 \le i \ne j < n} \Box\neg(\phi_i \wedge \phi_j)$ and so by the induction hypothesis we have $\bigoplus_{i=1}^{n-1} \phi_i \approx \bigvee_{i=1}^{n-1} \phi_i$.
Since we have $\phi_n \preceq \phi_n$ by (Ref), it follows by Theorem 4.6.3 that

$$\bigoplus_{i=1}^{n} \phi_i \approx (\bigvee_{i=1}^{n-1} \phi_i) \oplus \phi_n. \tag{7}$$

We have noted that $\Box$ is a normal modal operator; a consequence of our antecedent conditions is that we have $\Box\neg((\bigvee_{i=1}^{n-1} \phi_i) \wedge \phi_n)$.
From (K3) and modus ponens we obtain that $(\bigvee_{i=1}^{n-1} \phi_i) \oplus \phi_n \approx (\bigvee_{i=1}^{n-1} \phi_i) \vee \phi_n$ or

$$(\bigvee_{i=1}^{n-1} \phi_i) \oplus \phi_n \approx \bigvee_{i=1}^{n} \phi_i \tag{8}$$

We have that $\approx$ is an equivalence relation; consequently from (7) and (8) we obtain that $\bigoplus_{i=1}^{n} \phi_i \approx \bigvee_{i=1}^{n} \phi_i$ as desired. This satisfies the induction step and so our result follows by induction.

2: The proof is by induction on n. For $n = 1$ the result is immediate from Theorem 4.8.1.
For the induction step, assume that the result holds for $n = k$, and assume that we are given $\bigwedge_{i=1}^{k+1} \Box(\phi_i \equiv \psi_i)$. By the induction hypothesis and classical reasoning, this implies that $\Box(\phi_{k+1} \equiv \psi_{k+1}) \wedge \bigoplus_{i=1}^{k} \phi_i \approx \bigoplus_{i=1}^{k} \psi_i$.
An instance of Theorem 4.6.3 is $\bigoplus_{i=1}^{k} \phi_i \approx \bigoplus_{i=1}^{k} \psi_i \supset \bigoplus_{i=1}^{k} \phi_i \oplus \phi_{k+1} \approx \bigoplus_{i=1}^{k} \psi_i \oplus \phi_{k+1}$.
Since we also have $\Box(\phi_{k+1} \equiv \psi_{k+1})$ an application of Theorem 4.6.5 yields $\bigoplus_{i=1}^{k+1} \phi_i \approx \bigoplus_{i=1}^{k+1} \psi_i$, which was what was to be shown. Hence by induction our result obtains.
3: Assume that $\Phi_i \preceq \Psi_i$ for each $i \leq n$. Since it follows that $\Phi_i \preceq \Psi_i$ for all $i \neq k$, it follows by $n-1$ applications of (Add) that

$$\bigoplus_{i \neq k} \Phi_i \preceq \bigoplus_{i \neq k} \Psi_i. \tag{9}$$

If we assume that $\bigoplus_{i=1}^{n} \Phi_i \approx \bigoplus_{i=1}^{n} \Psi_i$, it then follows by the definition of $\approx$ and (Com) that

$$\left(\bigoplus_{i \neq k} \Psi_i\right) \oplus \Psi_k \preceq \left(\bigoplus_{i \neq k} \Phi_i\right) \oplus \Phi_k. \tag{10}$$

But from (9) and (10) it follows by Theorem 4.6.8 and modus ponens that $\Psi_k \preceq \Phi_k$.
4: The proof is analogous to that of the previous part. $\Box$

Proof of Theorem 4.10. By induction on the length of derivation of a LQP-theorem ϕ, we prove we have $M, w \models \phi$ for each simple pointed model (M, w). We omit the cases from propositional logic, which are straightforward. In the base case, we must show that each axiom scheme of LQP is valid with respect to the class of simple models. So let M be an arbitrary simple model and w a possible world in M.

- (Triv) is valid: Since 0 is a contradiction and 1 is a tautology, we have $0 = P_w([\![0]\!]_M^w) < P_w([\![1]\!]_M^w) = 1$. But then we have by the definition of satisfaction and of the abbreviation $0 \prec 1$ that $M, w \models 0 \prec 1$.
- (Tran) and (Tot) are valid: By the definition of satisfaction, and the transitivity and totality of $\leq$ over the real numbers.
- (Sub) is valid: Assume $M, w \models \Box(\phi_1 \equiv \phi_2)$ and $M, w \models \Box(\psi_1 \equiv \psi_2)$. This means we have $P_w([\![\phi_1 \equiv \phi_2]\!]_M^w) = 1$ and $P_w([\![\psi_1 \equiv \psi_2]\!]_M^w) = 1$. It follows that $P_w([\![\phi_1]\!]_M^w) = P_w([\![\phi_2]\!]_M^w)$ and $P_w([\![\psi_1]\!]_M^w) = P_w([\![\psi_2]\!]_M^w)$. By the definition of satisfaction, we obtain the consequent of (Sub).
- (Com) is valid: By the definition of satisfaction and the commutativity of sum in arithmetic.
- (Add) is valid: By the definition of satisfaction and the truth of the analogous property for the natural order $\leq$ over the reals.
- (Succ) is valid: Assume $M, w \models 1 \oplus \Phi \preceq 1 \oplus \Psi$. This means we have

  $$P_w([\![1]\!]_M^w) + \sum_{\phi \in \Phi} P_w([\![\phi]\!]_M^w) \leq P_w([\![1]\!]_M^w) + \sum_{\psi \in \Psi} P_w([\![\psi]\!]_M^w).$$

  Cancelling the leftmost element on each side, we obtain a statement that is equivalent by the definition of satisfaction to $M, w \models \Phi \preceq \Psi$.
- (K1) is valid: We have $0 \leq P_w([\![\phi]\!]_M^w)$. Since $[\![0]\!]_M^w = \emptyset$ and $P_w(\emptyset) = 0$, it follows by the definition of satisfaction that $M, w \models 0 \preceq \phi$.
- (K3) is valid: Suppose $M, w \models \Box\neg(\phi \wedge \psi)$. This means

  $$1 \ \leq\ P_w([\![\neg(\phi \wedge \psi)]\!]_M^w) \ =\ P_w([\![\neg(\phi \wedge \psi)]\!]_M \cap \Omega_w),$$

  from which it follows that

  $$1 = P_w(\Omega_w - ([\![\phi]\!]_M \cap [\![\psi]\!]_M)) = 1 - P_w([\![\phi]\!]_M \cap [\![\psi]\!]_M \cap \Omega_w).$$

  As a result, $P_w([\![\phi]\!]_M \cap [\![\psi]\!]_M \cap \Omega_w) = 0$. Hence

  $$\begin{aligned} P_w([\![\phi \vee \psi]\!]_M \cap \Omega_w) &= P_w(([\![\phi]\!]_M \cup [\![\psi]\!]_M) \cap \Omega_w) \\ &= P_w([\![\phi]\!]_M \cap \Omega_w) + P_w([\![\psi]\!]_M \cap \Omega_w). \end{aligned}$$

  Applying the definition of satisfaction and the meaning of $\approx$, we get $M, w \models \phi \oplus \psi \approx \phi \vee \psi$.

This completes the induction base. For the induction step, we assume that a derivable hypothesis ϕ of (Nec) is valid (this is the induction hypothesis), and we prove that the (Nec)-derivable consequence $\Box\phi$ is as well. It follows by the induction hypothesis that $M', w' \models \phi$ for each pointed model (M', w'). As a result, it follows that $M \models \phi$ and therefore $P_w([\![\phi]\!]_M \cap \Omega_w) = P_w(\Omega_w) = 1$. But then it follows by the definition of satisfaction and the meaning of $\Box$ that $M, w \models \Box\phi$.

Proof of Corollary 4.10.1. Let M be a one-world simple model. By Definition 4.3, $M, w \not\models 0$ and therefore $\mathcal{C}_s \not\models 0$. Applying Theorem 4.10 gives $\nvdash_{\mathsf{LQP}} 0$.

Completeness makes use of Theorem 1.2 from [28]. We start with some preliminary notions: For a finite nonempty set S, let $L(S)$ be the real vector space with coordinates in S; this is just like $\mathbb{R}^n$ but with coordinate set S instead of $\{1, \ldots, n\}$. A *linear functional on* $L(S)$ is a function $f : L(S) \to \mathbb{R}$ that is *linear*, meaning $f(ax + by) = af(x) + bf(y)$ for all reals $a, b \in \mathbb{R}$ and $x, y \in L(S)$. A set $X \subseteq L(S)$ is *rational* iff each $x \in X$ has its range in the set $\mathbb{Q}$ of rational numbers, and X is *symmetric* iff each $x \in X$ implies $-x \in X$.

Theorem B.1 *(Scott [28, Theorem 1.2]). Let S be a finite nonempty set and let X be a finite, rational, symmetric subset of $L(S)$. For each $N \subseteq X$, there exists a linear functional f on $L(S)$ that* realizes N in X, *meaning $N = \{x \in X \mid f(x) \geq 0\}$, if and only if the following conditions are satisfied:*

1. *for each $x \in X$, we have $x \in N$ or $-x \in N$; and*
2. *for each $n \geq 1$ and $x_1, \ldots, x_n \in N$ we have: $\sum_{i=1}^{n} x_i = 0$ implies $-x_1 \in N$.*

We use Scott's Theorem B.1 to prove completeness of LQP with respect to the class of simple models. Scott's result will be used to generate a probability function on a simple model we will construct that satisfies a given non–LQP-provable formula θ. This requires a number of ideas that we have developed based on an idea due to Lenzen [23]. Though intricate, we attempt to provide some intuition along the way.

Proof of Theorem 4.11. For convenience, we write $\vdash$ to mean $\vdash_{\mathsf{LQP}}$. We make use of Corollary 4.10.1 without mention. To prove completeness, assume $\nvdash_{\mathsf{LQP}} \theta$. It suffices to construct a pointed simple model for $\neg\theta$. Since simple models are finite, we shall base the model on an initial finite set made up of subformulas of θ, their negations, and finitely many other formulas whose role we explain shortly. Worlds of the model will be maximal LQP-consistent subsets of this initial finite set, and the model will be arranged so that a formula ϕ in this set is true at a world w if and only if w contains ϕ. The difficulty will be in defining the probability function on such models, and for this we will turn to Scott's Theorem B.1. To do so, we will define, for each world, sets that play the role of N and X in Scott's Theorem. We explain how this works as the proof progresses.

For $\phi \in \mathcal{L}_{\mathsf{LQP}}$, let $\mathsf{sub}(\phi)$ denote the set of subformulas of ϕ; this is the set consisting of ϕ and all formulas that are constructed according to the grammar of $\mathcal{L}_{\mathsf{LQP}}$ along the way to the construction of ϕ.[15] Lift our subformula function $\mathsf{sub}(-)$ so that it operates on sets of formulas by joining together the subformulas of all formulas in the original set: for $S \subseteq \mathcal{L}_{\mathsf{LQP}}$, let $\mathsf{sub}(S) \doteq \bigcup_{\phi \in S} \mathsf{sub}(\phi)$. We now define some initial machinery we will use to include some of the formulas we require: given $S \subseteq \mathcal{L}_{\mathsf{LQP}}$ and $E \subseteq 2^{\mathcal{L}_{\mathsf{LQP}}}$, define

$$\begin{aligned}
\pm S &\doteq S \cup \{\neg\phi \mid \phi \in S\},\\
\circlearrowleft S &\doteq \pm\mathsf{sub}(S) \cup \pm\mathsf{sub}\{(\Psi \preceq \Phi) \mid (\Phi \preceq \Psi) \in \pm\mathsf{sub}(S)\},\\
E^d &\doteq \textstyle\bigvee_{S' \in E} \bigwedge S'.
\end{aligned}$$

Note: $\emptyset^d \doteq 1$. The set $\pm S$ extends a set by adding in all negations of formulas in the set. The set $\circlearrowleft S$ takes S, adds in the reversal of every inequality in S, and then closes the resulting set under subformulas and negations. This makes sure we have both directions of each inequality in the resultant set (along with one-step closure under subformulas and negations). Given a set E of sets of formulas, E^d is the disjunction of the conjunction of every set in E. Later we will think of worlds as certain sets of formulas and E as a set of worlds; E^d will then be a formula that defines the set E of worlds (hence the "d" in "E^d") in the sense that the formula E^d will be true at a world w if and only if $w \in E$ if and only if $E^d \in w$. For now, however, E^d is merely a disjunction of conjunctions.

Given $S \subseteq T \subseteq \mathcal{L}_{\mathsf{LQP}}$, to say that S is *maxcons* in T means S is *consistent* (i.e., for no finite $S' \subseteq S$ do we have $\vdash (\bigwedge S') \supset 0$) and adding to S any $\phi \in T$ not already present would produce a set that is *inconsistent* (i.e., not consistent). Define

$$\begin{aligned}
A &\doteq \circlearrowleft\{\theta\};\\
W &\doteq \{w \subseteq A \mid w \text{ is maxcons in } A\};\\
w^{\Box} &\doteq \{\phi \in A \mid \vdash (\textstyle\bigwedge w) \supset \Box\phi\};\\
\Omega_w &\doteq \{v \in W \mid w^{\Box} \subseteq v\} \text{ for } w \in W;\\
V(w) &\doteq w \cap \mathcal{P} \text{ for } w \in W.
\end{aligned}$$

[15] To be clear: $\mathsf{sub}(p) = \{p\}$ for each $p \in \mathcal{P}$, $\mathsf{sub}(\neg\phi) = \{\neg\phi\} \cup \mathsf{sub}(\phi)$, $\mathsf{sub}(\phi \vee \psi) = \{\phi \vee \psi\} \cup \mathsf{sub}(\phi) \cup \mathsf{sub}(\psi)$, $\mathsf{sub}(\Phi \preceq \Psi) = \{\Phi \preceq \Psi\} \cup \mathsf{sub}(\Phi) \cup \mathsf{sub}(\Psi)$, and $\mathsf{sub}(\phi \oplus \Psi) = \mathsf{sub}(\phi) \cup \mathsf{sub}(\Psi)$.

It is easy to see that A and W are finite and W is nonempty. By modal reasoning, we have for each $w \in W$ that $w^{\Box}$ is consistent and so may be extended to $v \in W$ satisfying $w^{\Box} \subseteq v$. Hence $\Omega_w \neq \emptyset$. The set A is the basic set we need to construct our simple model for θ. It is the closure under the operator $\circlearrowright$ of the singleton consisting of our initial formula θ.

We then defined our set W of worlds: W is the collection of all subsets w made up of formulas from A such that w is maxcons in A. That is, W contains all subsets of A that are as large as possible but still consistent with LQP. In this way, any extraneous (and incorrect) formulas coming from A will be ignored. Note that A is finite and therefore each world is also finite. The conjunction $\bigwedge w$ consisting of the formulas making up a world (i.e., set of formulas) $w \in W$ is the formula $\{w\}^d$ that defines that world.

Given a world $w \in W$, we defined the set $w^{\Box}$ to be the set of formulas $\phi \in A$ such that LQP derives $(\bigwedge w) \supset \Box\phi$. (Recall that $\Box\phi$ abbreviates $(1 \preceq \phi)$.) Intuitively, $w^{\Box}$ is the set of formulas in A that must have probability 1 according to world w. We then define the set Ω_w of outcomes at w to be the set of all worlds $v \in W$ such that v makes true every formula that must have probability 1 according to w. This ensures that the outcome space Ω_w contains just those outcomes w says should be present. Finally, we defined the propositional valuation V in the usual way: a letter p is true at a world w if and only if p is a member of w. This will be the base case of a forthcoming *Truth Lemma*, which will shows a formula in A is true at a world w if and only if that formula is a member of w. We now provide some more definitions and observations to help apply Scott's Theorem and define a probability function P_w on Ω_w for each $w \in W$.

For $\chi \in \mathcal{L}_{\mathsf{LQP}}$, define $[\chi] \doteq \{v \in W \mid \vdash (\bigwedge v) \supset \chi\}$ and $[\chi]_w \doteq [\chi] \cap \Omega_w$. Thus $[\chi]$ is the set of worlds that derive χ; intuitively, this is the set of worlds at which χ may be thought of as "true." (Later the Truth Lemma will make it so for those in χ in A, but for now this is merely intuition.) $[\chi]_w$ is the restriction of $[\chi]$ to the set of worlds that are also outcomes of w.

We now make some observations about what can be derived from a world w. These will be used later in this completeness proof, some of which will be used in multiple parts of the proof. Our first observation is that from w, one has certainty of Ω_w^d:

$$\vdash (\bigwedge w) \supset \Box\Omega_w^d. \tag{11}$$

To prove (11), first note that by the definition of Ω_w and the fact that each $v \in W$ is maxcons in A that $\vdash (\bigwedge w^{\Box}) \equiv \Omega_w^d$. Applying modal reasoning (i.e., using Theorem 4.8 and the rules of LQP), we obtain $\vdash (\bigwedge_{\phi \in w^{\Box}} \Box\phi) \equiv \Box\Omega_w^d$. But for each $\phi \in w^{\Box}$, we have by definition that $\vdash (\bigwedge w) \supset \Box\phi$. Therefore, we obtain $\vdash (\bigwedge w) \supset \Box\Omega_w^d$.

Applying Theorem 4.8.4 to (11), we obtain for any $\chi \in \mathcal{L}_{\mathsf{LQP}}$:

$$\vdash (\bigwedge w) \supset (\chi \wedge \Omega_w^d \approx \chi)\ . \tag{12}$$

We finally observe that from w one can prove at least one $v \in \Omega_w$ has non-zero probability:

$$\vdash (\bigwedge w) \supset (0 \prec \bigwedge v) \qquad \text{for some } v \in \Omega_w. \tag{13}$$

Otherwise, $\vdash (\bigwedge w) \wedge ((\bigwedge v) \preceq 0)$ for all $v \in \Omega_w$; hence $\vdash (\bigwedge v) \preceq 0$ for all $v \in \Omega_w$, and by repeated applications of (Add) and Theorem 4.7.1, we have that $\vdash \bigoplus_{v \in \Omega_w} (\bigwedge v) \preceq 0$. Since worlds in Ω_w are pairwise inconsistent, we can apply (K3) to get $\vdash \Omega_w^d \preceq 0$. But together with (11), which states that $\vdash (\bigwedge w) \supset (1 \preceq \Omega_w^d)$, we obtain $\vdash (\bigwedge w) \supset (1 \preceq 0)$, which together with (Triv) implies w is inconsistent, a contradiction. Thus we conclude that $\vdash (\bigwedge w) \supset (0 \prec \bigwedge v)$ for some $v \in \Omega_v$.

For $E \subseteq W$, let $\iota(E)$ be the *characteristic function* of E: $\iota(E)(v) \doteq 1$ if $v \in E$, and $\iota(E)(v) \doteq 0$ if $v \in W - E$. Let

$$B \doteq A \cup \{0 \preceq 1\} \cup \{0 \preceq \{v\}^d \mid v \in W\}.$$

Remark B.2. We observe that for any world $w \in W$ and $\psi \in B$ either ψ is provable from w or inconsistent with w, that is either $\vdash (\bigwedge w) \supset \psi$ or $\vdash (\bigwedge w) \supset \neg\psi$.

To see this, first observe that for any $\psi \in A$, this property holds, as w is a maximally consistent subset of A. Any formula in B not in A is itself provable; for example $\vdash 0 \preceq \{v\}^d$ by (K1). This observation will be useful in proving (14) ahead.

For each $w \in W$, define:

$$N_w \doteq \{\textstyle\sum_{\psi \in \Psi} \iota([\psi]_w) - \sum_{\phi \in \Phi} \iota([\phi]_w) \mid \Phi \preceq \Psi \in B \text{ and } \vdash (\bigwedge w) \supset (\Phi \preceq \Psi)\},$$
$$X_w \doteq N_w \cup (-N_w).$$

We view each real-valued function f on W (such as $\iota(E)$) as a vector in the vector space $L(W)$ with basis W, by mapping each basis element $v \in W$ to its weight $f(v)$. N_w is our way of expressing in $L(W)$ the inequalities $(\Phi \preceq \Psi) \in B$ consistent with w. Recalling that $L(W)$ is the real vector space with coordinates in our set of worlds W (similar to the vector space $\mathbb{R}^n$ except that the coordinates come from W instead of from $\{1, 2, \ldots, n\}$), we will represent a set $E \subseteq W$ of worlds by its characteristic function $\iota(E)$. For example, supposing for simplicity we were to have $W = \{a, b, c\}$ and $E = \{a, b\}$ (with a, b, and c some symbols we use here for worlds), then we would represent E by the function $\iota(E) : W \to \mathbb{R}$ such that

$\iota(E)(a) = 1$ (since $a \in E$), $\iota(E)(b) = 1$ (since $b \in E$), and $\iota(E)(c) = 0$ (since $c \notin E$). But now instead of writing an inequality as $\Phi \preceq \Psi$, we will use "subtraction" (in the vector space) to move all the terms on the left-hand side of the inequality $\preceq$ to the right-hand side, writing the vector equivalent of $0 \preceq \Psi - \Phi$. However, we will be able to drop the "$0 \preceq$" part of this expression because, following the statement of Scott's Theorem B.1, our set N_w will be the set "N" in the statement of that theorem and the result of the theorem will guarantee that all of the members of this set N_w will be mapped by the linear functional to a non-negative value. Thus simply including the vector version of "$\Psi - \Phi$" in the set N_w will suffice. So long as we can prove the conditions of Scott's Theorem B.1 hold, we will obtain a linear functional that ensures the vector space equivalent of the statement $0 \preceq \Psi - \Phi$. Finally, we define X_w to play the role of the set "X" from the statement of Scott's Theorem B.1. X_w is defined so as to be as small as possible so that it contains N_w and is closed under vector negation (to ensure symmetry holds, as required by the theorem). This way of choosing N_w and X_w is our adaptation of an idea due to Lenzen [23], who used a similar kind of definition to prove a result for a different theory. We have adapted his idea and extended it for our own purposes here. Our task now will be to prove that N_w and X_w so chosen satisfy the conditions of Scott's Theorem B.1. From this we will obtain a linear functional from which we can construct a probability function P_w enabling us to complete the proof.

We begin with an observation that will also be useful when proving the probability function is well defined: for any $\chi \in B$, we have

$$\vdash \bigoplus_{v\in[\chi]_w} \{v\}^d \approx \chi \wedge \Omega_w^d \ . \tag{14}$$

Equation (14) is proved as follows:

Take an arbitrary $\chi \in B$. For $v, v' \in [\chi]_w$ with $v \neq v'$, we have $\vdash \neg(\{v\}^d \wedge \{v'\}^d)$ by classical reasoning and hence $\vdash \Box\neg(\{v\}^d \wedge \{v'\}^d)$ by (Nec). It therefore follows by Theorem 4.9.1 that

$$\vdash \bigoplus_{v\in[\chi]_w} \{v\}^d \approx \bigvee_{v\in[\chi]_w} \{v\}^d \ . \tag{15}$$

It remains to prove

$$\vdash \bigvee_{v\in[\chi]_w} \{v\}^d \equiv \chi \wedge \Omega_w^d \ . \tag{16}$$

The right-hand-side of the equivalence in (16) is equal to $\chi \wedge \bigvee_{v\in\Omega_w} \bigwedge v$ which, by distributivity of $\wedge$ over $\vee$, is provably equivalent to $\bigvee_{v\in\Omega_w}(\chi \wedge \bigwedge v)$. By Remark B.2, for each $\chi \in B$ and $v \in W$, either $\vdash (\bigwedge v) \supset \chi$ or $\vdash (\bigwedge v) \supset \neg\chi$. Thus we observe for each disjunct that

$$\begin{aligned} &\vdash \chi \wedge \bigwedge v \equiv \bigwedge v && \text{if } \vdash (\bigwedge v) \supset \chi \ , \\ &\vdash \chi \wedge \bigwedge v \equiv 0 && \text{if } \vdash (\bigwedge v) \supset \neg\chi \ . \end{aligned}$$

Hence by classical reasoning, $\bigvee_{v\in\Omega_w}(\chi \wedge \bigwedge v)$ is logically equivalent to $\bigvee_{v\in\Omega_w\cap[\chi]} \bigwedge v$, which is equal to the left hand side of (16).

Apply (Nec) to (16) and combine the result with (15) using Theorem 4.6.4 to obtain (14).

Lemma B.3. *For each w, the set X_w is a finite, rational, symmetric subset of $L(W)$, and $N_w \subseteq X_w$, satisfying the conditions of Scott's Theorem:*

1. *for each $x \in X_w$, either $x \in N_w$ or $-x \in N_w$; and*
2. *for each $n \geq 1$ and $x_1, \dots, x_n \in N_w$: $\sum_{i=1}^n x_i = 0$ implies $-x_1 \in N_w$.*

Proof. We have $N_w \subseteq X_w$; also, X_w is a finite, rational, symmetric subset of $L(W)$. It is obvious that N_w satisfies Item 1. We prove N_w also satisfies Item 2. So assume we have $x_1, \dots, x_n \in N_w$ satisfying $\sum_{i=1}^n x_i = 0$. Each vector x_i has the form $\sum_{\psi\in\Psi_i} \iota([\psi]_w) - \sum_{\phi\in\Phi_i} \iota([\phi]_w)$ for some $\Phi_i \preceq \Psi_i \in B$ where $\vdash (\bigwedge w) \supset (\Phi_i \preceq \Psi_i)$.

For any $\chi \in \mathcal{L}_{\mathsf{LQP}}$ (in particular $\chi \in B$), we have by linearity, $\iota([\chi]_w) = \sum_{v\in[\chi]_w} \iota(\{v\})$, so the assumption implies

$$\sum_{i=1}^n \sum_{\phi\in\Phi_i} \sum_{v\in[\phi]_w} \iota(\{v\}) = \sum_{i=1}^n \sum_{\psi\in\Psi_i} \sum_{v\in[\psi]_w} \iota(\{v\}) \ .$$

So writing out the sums in full without combining any summands, the equation above says that a coordinate $v \in W$ has exactly the same number of appearances in a summand on the left as on the right. So by (Ref) (from Theorem 4.6) and (Com), we obtain

$$\vdash \bigoplus_{i=1}^n \bigoplus_{\phi\in\Phi_i} \bigoplus_{v\in[\phi]_w} \{v\}^d \approx \bigoplus_{i=1}^n \bigoplus_{\psi\in\Psi_i} \bigoplus_{v\in[\psi]_w} \{v\}^d \ . \tag{17}$$

It follows from (17), (14), and (12) by Theorem 4.6.6 and the equality $\bigoplus_{\phi\in\Phi} \phi = \Phi$ that

$$\vdash (\bigwedge w) \supset (\bigoplus_{i=1}^n \Phi_i \approx \bigoplus_{i=1}^n \Psi_i) \ . \tag{18}$$

Since $\vdash (\bigwedge w) \supset (\Phi_i \preceq \Psi_i)$ for each $i \in \{1, \dots, n\}$, it follows from (18) by Theorem 4.9.3 that $\vdash (\bigwedge w) \supset (\Psi_1 \preceq \Phi_1)$. That is, $-x_1 \in N_w$. Therefore, Item 2 is also satisfied. $\square$

Applying Scott's Theorem B.1 yields a linear functional f_w on $L(W)$ that realizes N_w in X_w. It is this linear functional f_w that we shall use to construct a probability function. As we will see, essentially all we will need to do is define the probability of a set $E \subseteq \Omega_w$ to be the value f_w assigns to our vector representation of E divided by the value f_w assigns to our vector representation of the outcome set Ω_w. But for the probability to be well-defined, the denominator must not be zero. Proceeding, it will be helpful to recall that the vector representation of a set $E \subseteq \Omega_w$ is given by its characteristic function $\iota(E)$. Thus the denominator is $f_w(\iota(\Omega_w))$.

To show that $f_w(\iota(\Omega_w)) \neq 0$, first observe that $\iota(\Omega_w) = \iota(1) - \iota(0) \in N_w$, since $0 \preceq 1 \in B$. Let $x = -\iota(\Omega_w)$. As $-x \in N_w$, $x \in X_w$. We wish to show that $x \notin N_w$, and hence $f_w(x) < 0$, since f_w realizes N_w.

Suppose that $x = \sum_{\psi\in\Psi} \iota([\psi]_w) - \sum_{\phi\in\Phi} \iota([\phi]_w) \in N_w$ for some $\Phi \preceq \Psi \in B$. We now show that $\nvdash (\bigwedge w) \supset (\Phi \preceq \Psi)$. For each $v \in \Omega_w$, we have $x(v) = -1$, and hence there are more $\phi \in \Phi$ such that $v \in [\phi]_w$ than $\psi \in \Psi$ such that $v \in [\psi]_w$. Note that by (13), we have that $\vdash (\bigwedge w) \supset (0 \prec \bigwedge v)$ for some $v \in \Omega_w$. Then by repeated applications of Theorem 4.6.8 (which is provably equivalent to $(\Phi_1 \preceq \Psi_1) \supset ((\Phi_2 \prec \Psi_2) \supset (\Phi_1 \oplus \Phi_2 \prec \Psi_1 \oplus \Psi_2))$), (Tot), and Theorem 4.7.1, we obtain

$$\vdash (\bigwedge w) \supset \left(\bigoplus_{\phi\in\Phi} \bigoplus_{v\in[\phi]_w} \{v\}^d \succ \bigoplus_{\psi\in\Psi} \bigoplus_{v\in[\psi]_w} \{v\}^d \right) \quad . \tag{19}$$

Then by (14) and Theorem 4.6.6, we obtain $\vdash (\bigwedge w) \supset \left(\bigoplus_{\phi\in\Phi} (\phi \wedge \Omega_w^d) \succ \bigoplus_{\psi\in\Psi} (\psi \wedge \Omega_w^d) \right)$. By (12) and Theorem 4.6.6, we have that $\vdash (\bigwedge w) \supset (\Phi \succ \Psi)$. As w is consistent, $\nvdash (\bigwedge w) \supset (\Phi \preceq \Psi)$, and hence $x \notin N_w$. Hence $f_w(x) < 0$, and $f_w(-x) = f_w(\iota(\Omega_w)) > 0$. So we may define $P_w : 2^{\Omega_w} \to [0, 1]$ by

$$P_w(E) \doteq f_w(\iota(E)) \,/\, f_w(\iota(\Omega_w)) \quad .$$

We prove that P_w is a probability measure on 2^{Ω_w}.

- $P_w(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n P_w(E_i)$ for pairwise disjoint $E_1, \ldots, E_n \subseteq \Omega_w$.
  Characteristic functions are additive. Thus $\iota(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \iota(E_i)$. Hence

$$\begin{aligned} P_w(\textstyle\bigcup_{i=1}^n E_i) &= \frac{f_w(\iota(\bigcup_{i=1}^n E_i))}{f_w(\iota(\Omega_w))} \\ &= \frac{f_w(\sum_{i=1}^n \iota(E_i))}{f_w(\iota(\Omega_w))} && \text{by additivity of } \iota \\ &= \sum_{i=1}^n \frac{f_w(\iota(E_i))}{f_w(\iota(\Omega_w))} && \text{by linearity of } f_w \\ &= \sum_{i=1}^n P_w(E_i) \quad . \end{aligned}$$

- $P_w(E) \geq 0$ for each $E \subseteq \Omega_w$.
  Because of Ω_w is finite and P_w is additive, it suffices to show that $P_w(\{v\}) \geq 0$ for each $v \in E$. So let $v \in E$. We have $(0 \preceq \{v\}^d) \in B$ and $\vdash (0 \preceq \{v\}^d)$ by (K1); hence $\vdash (\bigwedge w) \supset (0 \preceq \{v\}^d)$ by classical reasoning. Thus $(0 \preceq \{v\}^d) \in N_w$. Since f_w realizes N_w in X_w, it follows that

$$f_w(\iota([0]_w)) \leq f_w(\iota([\{v\}^d]_w)) \quad . \tag{20}$$

  We have $[0]_w = \emptyset$ and it follows by linearity that $f_w(\iota(\emptyset)) = 0$. Also, since $v \in \Omega_w$, we have $[\{v\}^d]_w = \{v\}$. Applying these facts, the definition of P_w, and (20), it follows that $P_w(\{v\}) \geq 0$.
- $P_w(\Omega_w) = 1$ by definition.

So P_w is a probability measure on 2^{Ω_w}, and therefore $M \doteq (W, P, V)$ is a simple model. What remains is for us to make good on our promise that a world will consist of the set of formulas in A that are true at that world. This is the following lemma.

Lemma B.4 *(Truth Lemma). For each $w \in W$ and $\chi \in A$, we have $\chi \in w$ iff $M, w \models \chi$.*

Proof. We prove this by induction on $\mathcal{L}_{\mathsf{LQP}}$-formula construction. The induction base and Boolean induction step cases are standard, so we only address the induction step case for formulas $\Phi \preceq \Psi$.

- Left to right: if $w \in W$ and $(\Phi \preceq \Psi) \in w$, then $M, w \models \Phi \preceq \Psi$.
  Assume $w \in W$ and $(\Phi \preceq \Psi) \in w$. Since f_w realizes N_w in X_w, we have

$$\sum_{\phi\in\Phi} f_w(\iota([\phi]_w)) \leq \sum_{\psi\in\Psi} f_w(\iota([\psi]_w)) \quad . \tag{21}$$

Since $(\Phi \preceq \Psi) \in w \subseteq A$ implies we have $\phi, \psi \in A$ for each $\phi \in \Phi$ and each $\psi \in \Psi$, we may apply the induction hypothesis to obtain $[\phi]_w = [\![\phi]\!]_M^w$ for each $\phi \in \Phi$ and $[\psi]_w = [\![\psi]\!]_M^w$ for each $\psi \in \Psi$. Therefore, we obtain from (21) by the definition of P_w that $\sum_{\phi\in\Phi} P_w([\![\phi]\!]_M^w) \le \sum_{\psi\in\Psi} P_w([\![\psi]\!]_M^w)$. Applying Definition 4.3, we have $M, w \models \Phi \preceq \Psi$.

- Right to left: if $(\Phi \preceq \Psi) \in A$ and $M, w \models \Phi \preceq \Psi$, then $(\Phi \preceq \Psi) \in w$.
  Assume $(\Phi \preceq \Psi) \in A$ and $M, w \models \Phi \preceq \Psi$. Applying Definition 4.3, the definition of P_w, and multiplying both sides of the resulting inequality by $f_w(\iota(\Omega_w))$, we obtain (21). Since $(\Phi \preceq \Psi) \in A$ implies $(\Psi \preceq \Phi) \in A$ by the definition of A, it follows by (Tot) that $(\Phi \preceq \Psi) \in w$ or $(\Psi \preceq \Phi) \in w$. If we had $(\Phi \preceq \Psi) \notin w$, then it would follow that $(\Psi \preceq \Phi \in w)$ and hence we would obtain $\sum_{\psi\in\Psi} f_w(\iota([\psi]_w)) \le \sum_{\phi\in\Phi} f_w(\iota([\phi]_w))$ by the fact that f_w realizes N_w in X_w. This in turn would contradict (21). Conclusion: $(\Phi \preceq \Psi) \in w$. □

Then, since $\nvdash_{\mathsf{LQP}} \theta$, there exists $w_\theta \in W$ such that $\neg\theta \in w_\theta$. Applying the Truth Lemma, $M, w_\theta \not\models \theta$. Completeness follows. Note: (Sub), (Add), and (Succ) are used in the proof of Theorem 4.6.4; (Tran) and (K3) are used in the proof of Theorem 4.9.1. □

Proof of Theorem 4.12. Without loss of generality, we may assume that each of the a_i's and b_j's is non-negative. By the preceding definition of rational linear inequalities in our approach, and the soundness and completeness of LQP (Theorems 4.10 and 4.11), we have the right side of the desired equivalence iff $\sum_{i=1}^n a_i \cdot P_w([\![1]\!]_M \cap \Omega_w) \le \sum_{i=1}^m b_i \cdot P_w([\![1]\!]_M \cap \Omega_w)$. But this is itself equivalent to $\sum_{i=1}^n a_i \le \sum_{i=1}^m b_i$. □

B.2. Proofs of Section 5

We introduce the following notation and terminology. Π_n will denote the set of all permutations of the integers $1, \dots, n$. Thus for $\pi \in \Pi_n$ we have that for $1 \le i \le n$, $\pi(i) \in \{1, \dots, n\}$ and that $\pi(i) = \pi(j)$ iff $i = j$. We also extend our sequence notation to be used with the $\mathbb{E}$ abbreviation. Thus we may write $(\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m$ as $\Phi\mathbb{E}\Psi$, understanding Φ to stand for $(\phi_i)_{i=1}^m$ and similarly for Ψ.[16] $\Phi' \subseteq \Phi$ indicates multiset containment; i.e. each element of Φ' can be paired with a distinct element of Φ. Then $D(\Phi, i)$ is defined to be the set of all sequences of size i with elements drawn from Φ, expressed as a disjunction of conjunctions.

Definition B.5. For sequence Φ, and for $1 \le i \le |\Phi|$, define

$$D(\Phi, i) \doteq \bigvee \{\wedge\Phi' \mid \Phi' \subseteq \Phi \text{ and } |\Phi'| = i\}.$$

Segerberg suggests that the $\mathbb{E}$ schema captures a generalisation of necessary equivalence, but he doesn't elaborate on this point. Part 1 shows that $\Phi\mathbb{E}\Psi$ holds just if, necessarily, elements of Φ and Ψ can be "paired off" such that each such pair of formulas are equivalent. Part 2 of the theorem asserts that $\Phi\mathbb{E}\Psi$ holds iff necessarily, for every i, Φ is true of at least i elements iff Ψ is.

Theorem B.6 *($\mathbb{E}$-Schema). Let $\Phi = \bigoplus_{i=1}^m \phi_i$ and let $\Psi = \bigoplus_{i=1}^m \psi_i$. Then:*

1. $\vdash_{\mathsf{LQP}} \quad \Phi\mathbb{E}\Psi \equiv \Box \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$
2. $\vdash_{\mathsf{LQP}} \quad \Phi\mathbb{E}\Psi \equiv \Box \bigwedge_{i=1}^m (D(\Phi, i) \equiv D(\Psi, i))$
3. $\vdash_{\mathsf{LQP}} \quad \bigoplus_{i=1}^m \phi_i \approx \bigoplus_{i=1}^m D(\Phi, i)$

Proof of Theorem B.6. In Theorem 4.8 we showed that the logic of □ is that of the normal modal logic KD; thus we can freely use results from this modal logic. So for both parts of the theorem we argue with respect to a model $M = \langle W, R, P\rangle$ over a set of atomic sentences $\mathcal{P}$, where W is a set (of possible worlds), $R \subseteq W \times W$ is a serial relation, and $P : \mathcal{P} \mapsto 2^W$ specifies the truth assignment of atomic sentences at worlds. Truth of a formula at a world and other notions are defined in the standard way; see any basic text on modal logic (e.g. [4,19] for details).

1: Left to right: We argue via the model theory; hence assume that for model M and possible world w that $M, w \models \Phi\mathbb{E}\Psi$ or $M, w \models (\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m$. That is, $M, w \models \Box \bigvee_{i=0}^m C_i$ where C_i is specified according to Equation (3). Let $w' \in W$ be such that $(w, w') \in R$; such a w' exists since R is serial.
Then $M, w' \models \bigvee_{i=0}^m C_i$, which is to say, for some j $(0 \le j \le m)$, we have $M, w' \models C_j$. Expressing this in terms of (3) we have

$$M, w' \models e_1\phi_1 \wedge \cdots \wedge e_m\phi_m \wedge f_1\psi_1 \wedge \cdots \wedge f_m\psi_m$$

[16] In $\Phi\mathbb{E}\Psi$, Segerberg and Gärdenfors assume that $|\Phi| = |\Psi|$, which we also adopt. This assumption is inessential since if, say, $|\Phi| < |\Psi|$ the sequence Φ can always be "padded" to the length of Ψ by adding instances of 0.

where each e_k and f_k is either the negation symbol $\neg$ or the empty string, and exactly j of the e and f elements are the negation symbol.
Thus there is a permutation $\pi \in \Pi$ such that $M, w' \models e_1\phi_1 \wedge f_{\pi(1)}\psi_{\pi(1)} \wedge \cdots \wedge e_m\phi_m \wedge f_{\pi(m)}\psi_{\pi(m)}$ and where for each k, e_k is the negation symbol iff $f_{\pi(k)}$ is.
Hence $M, w' \models (\phi_1 \equiv \psi_{\pi(1)}) \wedge \cdots \wedge (\phi_m \equiv \psi_{\pi(m)})$ which is to say $M, w' \models \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$. This means that $M, w' \models \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$.
Since w' is an arbitrary world accessible from w we have $M, w \models \Box \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$.
Right to left: For the other direction, it can be noted that the above argument can be equally well reversed. We omit the details,
This shows that for any model M and world w that $M, w \models \Phi\mathbb{E}\Psi \equiv \Box \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m \left(\phi_i \equiv \psi_{\pi(i)}\right)$.
The result then follows from the completeness result for KD with respect to serial models.

2: Left to right: Assume that for model M and possible world w we have $M, w \models \Phi\mathbb{E}\Psi$ or $M, w \models (\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m$. From the previous part this implies that $M, w \models \Box \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$, or for every $w' \in W$ where $(w, w') \in R$ that $M, w' \models \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$. So for some $\pi \in \Pi$ we have that $M, w' \models \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$.
Let $\Phi' = \{\phi \in \Phi \mid M, w' \models \phi\}$ and let $\Psi' = \{\psi \in \Psi \mid M, w' \models \psi\}$. Clearly $|\Phi'| = |\Psi'|$, and so for $k = |\Phi'|$ we have $M, w' \models D(\Phi, k)$ and $M, w' \models D(\Psi, k)$. Moreover, for $j > k$ we have $M, w' \not\models D(\Phi, j)$ and $M, w' \not\models D(\Psi, j)$, and for $j < k$ we have $M, w' \models D(\Phi, j)$ and $M, w' \models D(\Psi, j)$.
Putting this together yields $M, w' \models \bigwedge_{i=1}^m D(\Phi, i) \equiv D(\Psi, i)$. Since this holds for any w' where $(w, w') \in R$, we obtain $M, w \models \Box \bigwedge_{i=1}^m D(\Phi, i) \equiv D(\Psi, i)$.
Right to left: Assume that for model M and possible world w we have $M, w \models \Box \bigwedge_{i=1}^m (D(\Phi, i) \equiv D(\Psi, i))$. So for any w' where $(w, w') \in R$ we have $M, w' \models \bigwedge_{i=1}^m (D(\Phi, i) \equiv D(\Psi, i))$.
We show that $M, w' \models \bigwedge_{i=1}^m (\phi_i \equiv \psi_{\pi(i)})$ for some $\pi \in \Pi_m$.
Let k be the greatest index such that $M, w \models D(\Phi, k)$. (If there is no such k then we have for every $\phi \in \Phi$ that $M, w' \models \neg\phi$, and for every $\psi \in \Psi$ that $M, w' \models \neg\psi$, whence $M, w' \models \bigwedge_{i=1}^m (\phi_i \equiv \psi_i)$.) Consequently $M, w \models D(\Psi, k)$ and moreover for every $j > k$ we have $M, w \not\models D(\Phi, j)$ and $M, w \not\models D(\Psi, j)$.
Consider the fact that $M, w \models D(\Phi, k)$. This means that for some $\Phi' \subseteq \Phi$ where $|\Phi'| = k$ that $M, w' \models \wedge\Phi'$ and for any $\phi \notin \Phi'$ that $M, w' \not\models \phi$. Hence this defines a partition of Φ into Φ' and $\Phi \setminus \Phi'$ where elements of the former are true at w' and elements of the latter false. In the same fashion we can determine a partition of Ψ into Ψ' (of size k) and $\Psi \setminus \Psi'$, where elements of the former are true at w' and elements of the latter false.
So take any bijection between elements of Φ' and Ψ' and between elements of $\Phi \setminus \Phi'$ and $\Psi \setminus \Psi'$; this bijection defines a permutation $\pi \in \Pi_M$ of elements of Ψ with respect to elements of Φ. This shows that $M, w' \models \bigwedge_{i=1}^m \left(\phi_i \equiv \psi_{\pi(i)}\right)$, which we set out to show at the outset.
From $M, w' \models \bigwedge_{i=1}^m \left(\phi_i \equiv \psi_{\pi(i)}\right)$, we obtain that $M, w' \models \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m \left(\phi_i \equiv \psi_{\pi(i)}\right)$. Since this holds for any w' where $(w, w') \in R$, this means that $M, w \models \Box \bigvee_{\pi\in\Pi} \bigwedge_{i=1}^m \left(\phi_i \equiv \psi_{\pi(i)}\right)$; from the previous part of the theorem we obtain that $M, w \models \Phi\mathbb{E}\Psi$.
This shows for any model M and possible world w that $M, w \models \Phi\mathbb{E}\Psi \equiv \Box \bigwedge_{i=1}^m (D(\Phi, i) \equiv D(\Psi, i))$. From the completeness result for KD with respect to serial models we obtain that $\vdash_{\mathsf{LQP}} \Phi\mathbb{E}\Psi \equiv \Box \bigwedge_{i=1}^m (D(\Phi, i) \equiv D(\Psi, i))$

3: The proof makes use of two lemmas.

Lemma B.7. $D(\phi \oplus \Phi, i+1) \approx (\phi \wedge D(\Phi, i)) \vee D(\Phi, i+1)$

Proof. By Definition B.5, $D(\phi \oplus \Phi, i+1)$ is $\bigvee\{\wedge\Phi' \mid \Phi' \subseteq \phi \oplus \Phi \text{ and } |\Phi'| = i+1\}$.
This is equivalent to

$$\bigvee\{\phi \wedge (\wedge\Phi') \mid \Phi' \subseteq \Phi \text{ and } |\Phi'| = i\} \vee \bigvee\{\wedge\Phi' \mid \Phi' \subseteq \Phi \text{ and } |\Phi'| = i+1\}$$

which is equivalent to

$$(\phi \wedge \bigvee\{\wedge\Phi' \mid \Phi' \subseteq \Phi \text{ and } |\Phi'| = i\}) \vee \bigvee\{\wedge\Phi' \mid \Phi' \subseteq \Phi \text{ and } |\Phi'| = i+1\}.$$

However this is just $(\phi \wedge D(\Phi, i)) \vee D(\Phi, i+1)$. $\square$

Lemma B.8. $(\phi \wedge D(\Phi, i)) \oplus D(\Phi, i+1) \approx (D(\phi \oplus \Phi, i+1)) \oplus (\phi \wedge D(\Phi, i+1))$.

Proof. An instance of Theorem 4.7.4 is

$$(\phi \wedge D(\Phi, i)) \oplus D(\Phi, i+1) \approx ((\phi \wedge D(\Phi, i) \vee D(\Phi, i+1))) \oplus (\phi \wedge D(\Phi, i) \wedge D(\Phi, i+1)). \tag{22}$$

Consider the two terms on the right side of $\approx$.

1. From Lemma B.7, we have that $(\phi \wedge D(\Phi, i)) \vee D(\Phi, i+1) \approx D(\phi \oplus \Phi, i+1)$. Hence we can use Theorem 4.6.6 to simplify this term to $D(\phi \oplus \Phi, i+1)$.

2. The second term is $\phi \wedge D(\Phi, i) \wedge D(\Phi, i+1)$. An immediate consequence of Definition B.5 is that $D(\Psi, i+1) \supset D(\Psi, i)$ is a theorem of propositional logic. Hence this term is equivalent to $\phi \wedge D(\Phi, i+1)$.

Substituting these two parts into (22) gives what was to be shown:

$$(\phi \wedge D(\Phi, i) \oplus D(\Phi, i+1)) \approx D(\phi \oplus \Phi, i) \oplus \phi \wedge D(\Phi, i+1). \quad \square$$

The proof of the theorem is by induction on the length of the sequence Φ.
For $m = 1$, we have that $\phi \approx D(\phi, 1)$ is $\phi \approx \phi$, which is a consequence of (Ref).
For $m = 2$, we are to show $\phi_1 \oplus \phi_2 \approx D(\Phi, 1) \oplus D(\Phi, 2)$. This is the same as $\phi_1 \oplus \phi_2 \approx (\phi_1 \vee \phi_2) \oplus (\phi_1 \wedge \phi_2)$ which is Theorem 4.7.4.
For the induction hypothesis, assume that our result holds for $m = k$, that is for $\Phi = \bigoplus_{i=1}^{k} \phi_i$, we have $\bigoplus_{i=1}^{k} \phi_i \approx \bigoplus_{i=1}^{k} D(\Phi, i)$. We show, for formula ϕ_0, that $\bigoplus_{i=0}^{k} \phi_i \approx \bigoplus_{i=1}^{k+1} D(\phi_0 \oplus \Phi, i)$.[17]
Since $\bigoplus_{i=0}^{k} \phi_i$ is $\phi_0 \oplus \bigoplus_{i=1}^{k} \phi_i$, we can apply the induction hypothesis and use the Principle of Replacement to obtain

$$\bigoplus_{i=0}^{k} \phi_i \approx \phi_0 \oplus \bigoplus_{i=1}^{k} D(\Phi, i). \tag{23}$$

Consider $\phi_0 \oplus D(\Phi, 1)$ from the right hand side of $\approx$ in (23): An instance of Theorem 4.7.4 is $\phi_0 \oplus D(\Phi, 1) \approx (\phi_0 \vee D(\Phi, 1)) \oplus (\phi_0 \wedge D(\Phi, 1))$. The term $\phi_0 \vee D(\Phi, 1)$ is just $D(\phi_0 \oplus \Phi, 1)$ and hence $\phi_0 \oplus D(\Phi, 1) \approx D(\phi \oplus \Phi, 1) \oplus (\phi_0 \wedge D(\Phi, 1))$. Then, via the Replacement Principle, (23) is equivalent to

$$\bigoplus_{i=0}^{k} \phi_i \approx D(\phi \oplus \Phi, 1) \oplus (\phi_0 \wedge D(\Phi, 1)) \oplus \bigoplus_{i=2}^{k} D(\Phi, i). \tag{24}$$

Let the sequence on the right hand side of $\approx$ be Φ_1. We next iteratively obtain a sequence of sequences $\Phi_1, \dots, \Phi_{k-1}$ where $\Phi_i \approx \Phi_{i+1}$ for $1 \le i \le k-1$ and using Lemma B.8 as follows.

- For step i, we have that Φ_i is of the form

$$\bigoplus_{j=1}^{i} D(\phi \oplus \Phi, j) \oplus (\phi_0 \wedge D(\Phi, i)) \oplus \bigoplus_{j=i+1}^{k} D(\Phi, i).$$

- Using Lemma B.8 applied to $(\phi_0 \wedge D(\Phi, i)) \oplus D(\Phi, i+1)$, along with the Replacement Principle yields Φ_{i+1}:

$$\bigoplus_{j=1}^{i+1} D(\phi \oplus \Phi, j) \oplus (\phi_0 \wedge D(\Phi, i+1)) \oplus \bigoplus_{j=i+2}^{k} D(\Phi, i+1).$$

  where clearly $\Phi_i \approx \Phi_{i+1}$.

Continuing in this fashion we obtain Φ_k: $\bigoplus_{j=1}^{k} D(\phi \oplus \Phi, j) \oplus (\phi_0 \wedge D(\Phi, k))$
Finally, it can be noted that since $|\Phi| = k$, $(\phi_0 \wedge D(\Phi, k)) \equiv D(\phi_0 \oplus \Phi, k+1)$.
Thus Φ_k is equivalent to: $\bigoplus_{j=1}^{k+1} D(\phi \oplus \Phi, j)$.
Summing up, Equation (24) can be written $\bigoplus_{i=0}^{k} \phi_i \approx \Phi_1$; we have $\Phi_i \approx \Phi_{i+1}$ for $1 \le i \le k-1$; and Φ_k is equivalent to $\bigoplus_{j=1}^{k+1} D(\phi \oplus \Phi, j)$. Thus by transitivity of $\approx$, we obtain

$$\bigoplus_{i=0}^{k} \phi_i \approx \bigoplus_{i=1}^{k+1} D(\phi \oplus \Phi, i)$$

which was to be shown. $\square$

Proof of Theorem 5.2.

1: This result can be proven entirely within the proof theory of our logic. However, it is also an easy consequence of previous results, some of which were shown by appealing to the model theory of modal logic KD, For brevity, we take the latter alternative.
Let Φ be $\bigoplus_{i=1}^{m} \phi_i$ and let Ψ be $\bigoplus_{i=1}^{m} \psi_i$. From Theorem B.6.2 we have that $\Phi \mathbb{E} \Psi$ is logically equivalent to $\Box \bigwedge_{i=1}^{m} (D(\Phi, i) \equiv D(\Psi, i))$. Using Theorem 4.9.2 (along with repeated application of the equivalence in modal logic K that $\Box(\gamma_1 \wedge \gamma_2) \equiv (\Box\gamma_1 \wedge \Box\gamma_2)$) we have that $\Box \bigwedge_{i=1}^{m} (D(\Phi, i) \equiv D(\Psi, i))$ implies $\bigoplus_{i=1}^{m} D(\Phi, i) \approx \bigoplus_{i=1}^{m} D(\Psi, i)$. But by Theorem B.6.3 and the Replacement Principle, Theorem 4.6.6, we have that this last expression is equivalent to $\bigoplus_{i=1}^{m} \phi_i \approx \bigoplus_{i=1}^{m} \psi_i$, from which we obtain.

[17] Note that on the left side of $\approx$ we run our index from 0 to k, rather than 1 to $k+1$ as would be more usual. Beginning from 0 makes the argument (slightly) more straightforward.

2: This follows by Part 1, Theorem 4.9.3, and classical reasoning. □

Proof of Theorem 5.3. Since $\mathcal{L}_{\mathsf{QP}} \subseteq \mathcal{L}_{\mathsf{LQP}}$, it follows that $\mathcal{L}_{\mathsf{LQP}}$ is at least as expressive as is $\mathcal{L}_{\mathsf{QP}}$ over our class of super simple models. We now show that this expressivity is strict. Define $W = \{w_1, w_2\}$ and $V : W \to 2^{\mathcal{P}}$ by $V(w_1) = \{p\}$ and $V(w_2) = \emptyset$. Let P^1 be the probability measure on 2^W given by $P^1(\{w_1\}) = \frac{1}{3}$ and P^2 be the probability measure on 2^W given by $P^2(\{w_1\}) = \frac{1}{4}$. Define $\Pi^i_x = (W, 2^W, P^i)$ for each $i \in \{1,2\}$ and $x \in W$. Define the super-simple models $M^1 = (W, \Pi^1, V)$ and $M^2 = (W, \Pi^2, V)$. It is easy to verify that $M^1, w_1 \models p \approx \frac{1}{3}$ but that $M^2, w_1 \not\models p \approx \frac{1}{3}$. The language $\mathcal{L}_{\mathsf{LQP}}$ therefore distinguishes (M^1, w_1) and (M^2, w_1).

By induction on the construction of formulas in $\mathcal{L}_{\mathsf{QP}}$, we show that for each $\chi \in \mathcal{L}_{\mathsf{QP}}$, we have $[\![\chi]\!]_{M^1} = [\![\chi]\!]_{M^2}$. The base and propositional Boolean cases are straightforward, so we only consider the induction case for the formula $\phi \preceq \psi$. Proceeding, we observe that for each $P \in \{P^1, P^2\}$, we have $P(\emptyset) < P(\{w_2\}) < P(\{w_1\}) < P(\{w_1, w_2\})$. But then we have by the induction hypothesis that $[\![\chi]\!]_{M^1} = [\![\chi]\!]_{M^2}$ for each $\chi \in \{\phi, \psi\}$, from which it follows that $[\![\phi \preceq \psi]\!]_{M^1} = [\![\phi \preceq \psi]\!]_{M^2}$. □

Proof of Theorem 5.4. By Theorems 4.10, 4.11, and the completeness of QP. Also, we note that: the QP principles (PC), (MP), and (Nec) are present in LQP; QP's (A0) is an instance of LQP's (Sub); QP's (A1) is the same as LQP's (K1); QP's (A2) is an instance of LQP's (Tot); QP's (A3) is LQP's (Triv); and LQP's (A4) is LQP-derivable per Theorem 5.2.2. □

Proof of Theorem 5.5.

1. We have $\mathcal{L}^-_{\mathsf{QP}} \subseteq \mathcal{L}_{\mathsf{PK}}$, so $\mathcal{L}_{\mathsf{PK}}$ is at least as expressive as $\mathcal{L}^-_{\mathsf{QP}}$ over the class of super-simple PK models. To show this expressive relationship is strict, define $W = \{w_1, w_2\}$ and $V : W \to 2^{\mathcal{P}}$ by $V(w_1) = \{p\}$ and $V(w_2) = \emptyset$. Let P be the probability measure on 2^W given by $P(\{w_1\}) = 1$. Define $\Pi_x = (W, 2^W, P)$ for each $x \in W$. Define $R^1 : W \to 2^W$ by $R^1(x) = \{x\}$ for each $x \in W$ and $R^2 : W \to 2^W$ by $R^2(x) = W$ for each $x \in W$. Define the super-simple PK models $M^i = (W, \Pi, R^i, V)$ for each $i \in \{1,2\}$. We have $M^1, w_1 \models \boxdot p$ and $M^2, w_1 \not\models \boxdot p$. However, by an easy induction on the construction of $\mathcal{L}^-_{\mathsf{QP}}$-formulas, it follows that $[\![\phi]\!]_{M^1} = [\![\phi]\!]_{M^2}$ for each $\phi \in \mathcal{L}^-_{\mathsf{QP}}$. So $\mathcal{L}_{\mathsf{PK}}$ is strictly more expressive than $\mathcal{L}^-_{\mathsf{QP}}$ over the class of super-simple PK models.
2. Our argument in Part 1 can be extended to show that no $\mathcal{L}_{\mathsf{QP}}$-formula distinguishes the super-simple PK models (M^1, w_1) and (M^2, w_1) defined in that part. So $\mathcal{L}_{\mathsf{QP}}$ is not more expressive than $\mathcal{L}_{\mathsf{PK}}$ over the class of super-simple PK models.
   To see that the languages are not comparable in the other direction over this class, define $W = \{w_1, w_2\}$ and $V : W \to 2^{\mathcal{P}}$ by $V(w_1) = \{p\}$ and $V(w_2) = \emptyset$. Let P^1 and P^2 be the probability measures on the finite σ-algebra 2^W satisfying the following:
   $$P^1(\{w_1\}) = P^1(\{w_2\}) = \tfrac{1}{2}, \quad P^2(\{w_1\}) = \tfrac{1}{4} \text{ and } P^2(\{w_2\}) = \tfrac{3}{4} \;.$$
   For each $i \in \{1,2\}$, define $\Pi^i_w = (W, 2^W, P^i)$ for each and $w \in W$ and the super-simple PK model $L^i = (W, \Pi^i, R^i V)$, where R^i is immaterial. Last, fix the $\mathcal{L}_{\mathsf{QP}}$-formula $\chi = p \prec (\neg p \preceq p)$. We note that $[\![\neg p \preceq p]\!]_{L^1} = W$ and $[\![\neg p \preceq p]\!]_{L^2} = \emptyset$. Hence
   $$P^1\left([\![\neg p \preceq p]\!]_{L^1}\right) = 1 \quad\text{and}\quad P^2\left([\![\neg p \preceq p]\!]_{L^2}\right) = 0 \;.$$
   Since $P^1([\![p]\!]_{L^1}) = \frac{1}{2} < 1$ and $P^2([\![p]\!]_{L^2}) = \frac{1}{4} \not< 0$, we obtain $L^1, w_1 \models \chi$ and $L^2, w_1 \not\models \chi$. By an easy induction on the construction of $\mathcal{L}_{\mathsf{PK}}$-formulas, it follows that $[\![\phi]\!]_{L^1} = [\![\phi]\!]_{L^2}$ for each $\phi \in \mathcal{L}_{\mathsf{PK}}$. So $\mathcal{L}_{\mathsf{PK}}$ is not more expressive than $\mathcal{L}_{\mathsf{QP}}$ over the class of super-simple PK models. Conclusion: these two languages are incomparable over this class of models.
3. We can extend the induction from Part 1 to show that we have $[\![\phi]\!]_{M^1} = [\![\phi]\!]_{M^2}$ for each $\phi \in \mathcal{L}_{\mathsf{LQP}}$. So $\mathcal{L}_{\mathsf{LQP}}$ is not more expressive than $\mathcal{L}_{\mathsf{PK}}$ over the class of super-simple PK models.
   To see the converse, now let the super-simple models M^1 and M^2 be defined as in the proof of Theorem 5.3. Extend these to super-simple PK models by adding $R : W \to 2^W$ satisfying $R(x) = W$ for each $x \in W$. By an easy induction on the construction of $\mathcal{L}_{\mathsf{PK}}$-formulas, it follows that $[\![\phi]\!]_{M^1} = [\![\phi]\!]_{M^2}$ for each $\phi \in \mathcal{L}_{\mathsf{PK}}$. So $\mathcal{L}_{\mathsf{PK}}$ is not more expressive than $\mathcal{L}_{\mathsf{LQP}}$ over the class of super-simple PK models. But then it follows that $\mathcal{L}_{\mathsf{PK}}$ and $\mathcal{L}_{\mathsf{LQP}}$ are expressively incomparable over the class of super-simple PK models. □

Proof of Theorem A.3. For any PK model $M = (W, \Pi, R, V)$, $\Pi_w = (\Omega_w, F_w, P_w)$, and possible world $w \in W$ we have $M, w \models 0 \prec 1$ by Axiom (#8) and the PK soundness and completeness result. this implies $M, w \models \neg(1 \preceq 0)$ or $M, w \not\models 1 \preceq 0$. From the definition of $\models$ this means that it is not the case that: $P_w([\![1]\!]^w_M) \leq P_w([\![0]\!]^w_M)$, which is to say, it is the case that

$$P_w([\![0]\!]^w_M) < P_w([\![1]\!]^w_M) \;. \tag{25}$$

The definition of $[\![\cdot]\!]$ is $[\![\phi]\!]^w_M \doteq [\![\phi]\!]_M \cap \Omega_w$ or, in the case of PK models $[\![\phi]\!]^w_M \doteq [\![\phi]\!]_M \cap \{w' \mid wRw'\}$. Hence (25) is the same as: $P_w(\{w' \mid wRw' \text{ and } M, w' \models 0\}) < P_w(\{w' \mid wRw' \text{ and } M, w' \models 1\})$ or $0 < P_w(\{w' \mid wRw' \text{ and } M, w' \models 1\})$. Assume that there is no $w' \in W$ such that wRw'. Then in this case $P_w(\{w' \mid wRw' \text{ and } M, w' \models 1\}) = P_w(\{\}) = 0$, which gives $0 < 0$, contradiction.

Since w is an arbitrary possible world, this implies that for any $w \in W$ there is $w' \in W$ such that wRw'. $\square$

Proof of Theorem A.4. An instance of Axiom (#7) is $\phi\dot{\mathbb{E}}\psi \supset ((\phi \preceq \psi) \supset (\psi \preceq \phi))$. The contrapositive of this instance is $((\phi \preceq \psi) \wedge \neg(\psi \preceq \phi)) \supset \neg(\phi\dot{\mathbb{E}}\psi)$ or $((\phi \preceq \psi) \wedge (\phi \prec \psi)) \supset \neg(\phi\dot{\mathbb{E}}\psi)$ which is logically equivalent to $(\phi \prec \psi) \supset \neg(\phi\dot{\mathbb{E}}\psi)$. Taking ϕ as 0 and ψ as 1 yields $(0 \prec 1) \supset \neg(0\dot{\mathbb{E}}1)$. Now $0 \prec 1$ is just Axiom (#8) and so via modus ponens we obtain that $\neg(0\dot{\mathbb{E}}1)$ is a theorem. $\neg(0\dot{\mathbb{E}}1)$ abbreviates $\neg\boxdot((1 \wedge 0) \vee (0 \wedge 1))$ which is equivalent to $\neg\boxdot 0$. Finally, $\neg\boxdot 0 \equiv (\boxdot\phi \supset \neg\boxdot\neg\phi)$ is a theorem of the modal logic K, and so via modus ponens we obtain that $\boxdot\phi \supset \neg\boxdot\neg\phi$ is a theorem. $\square$

Proof of Theorem 5.6. In the presence of (Symm), the axioms of PK simplify considerably: Axioms (#2) – (#5) are redundant, as is the antecedent condition of (#6) and (#9). As a result PK′ is defined by (Symm) and (PMN) together with the following:

(PC) All tautologies of classical propositional logic
(#0) $\boxdot(\phi \supset \psi) \supset (\boxdot\phi \supset \boxdot\psi)$
(#1) $\boxdot(\phi \equiv \phi') \wedge \boxdot(\psi \equiv \psi') \supset ((\phi \preceq \psi) \supset (\phi' \preceq \psi'))$
(#6’) $(\phi \preceq \psi) \vee (\psi \preceq \phi)$
(#7) $(\phi_i)_{i=1}^m \dot{\mathbb{E}} (\psi_i)_{i=1}^m \supset ((\bigwedge_{i=1}^m \phi_i \preceq \psi_i) \supset (\bigwedge_{i=1}^m \psi_i \preceq \phi_i))$ for all $m \geq 1$
(#8) $0 \prec 1$
(#9’) $0 \preceq \phi$
(MP) From $\phi \supset \psi$ and ϕ, infer ψ
(Nec) From ϕ infer $\boxdot\phi$

We have the following results:

1. In the presence of schema (#6’) (resp. (A2)), the schemas (#7) and (A4) are equivalent under propositional reasoning. To show that Gärdenfors's schema (A4) implies Segerberg's (#6’), observe that (A4) (for assumed m) can be considered as a set of i instances $((\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m \wedge \bigwedge_{k=1,k\neq j}^m (\phi_k \preceq \psi_k)) \supset (\psi_j \preceq \phi_j)$ for $1 \leq j \leq m$. Strengthening of the antecedent gives $((\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m \wedge \bigwedge_{k=1}^m (\phi_k \preceq \psi_k)) \supset (\psi_j \preceq \phi_j)$ for $1 \leq j \leq m$; and these together imply $((\phi_i)_{i=1}^m \mathbb{E} (\psi_i)_{i=1}^m \wedge \bigwedge_{k=1}^m (\phi_k \preceq \psi_k)) \supset \bigwedge_{k=1}^m (\psi_k \preceq \phi_k))$ which is (#7). For the other direction, (#7) can be expressed as $((\phi_i)_{i=1}^m \dot{\mathbb{E}} (\psi_i)_{i=1}^m \wedge (\bigwedge_{i=1}^m \phi_i \preceq \psi_i)) \supset (\bigwedge_{i=1}^m \psi_i \preceq \phi_i))$.
   This entails $((\phi_i)_{i=1}^m \dot{\mathbb{E}} (\psi_i)_{i=1}^m \wedge (\bigwedge_{i=1}^m \phi_i \preceq \psi_i)) \supset \psi_m \preceq \phi_m$, which by propositional reasoning is equivalent to $((\phi_i)_{i=1}^m \dot{\mathbb{E}} (\psi_i)_{i=1}^m \wedge (\bigwedge_{i=1}^{m-1} \phi_i \preceq \psi_i)) \supset (\neg(\phi_m \preceq \psi_m) \vee \psi_m \preceq \phi_m)$. Axiom (#6’) is equivalent to $\neg(\phi_m \preceq \psi_m) \supset (\psi_m \preceq \phi_m)$ and so the preceding expression implies $((\phi_i)_{i=1}^m \dot{\mathbb{E}} (\psi_i)_{i=1}^m \wedge (\bigwedge_{i=1}^{m-1} \phi_i \preceq \psi_i)) \supset \psi_m \preceq \phi_m$ which is (A4).
2. $\vdash_{\mathsf{PK}'} (1 \preceq \phi) \equiv \boxdot\phi$
   One direction is given by (PMN). For the other direction, note that $\boxdot\phi$ is equivalent to $\phi\dot{\mathbb{E}}1$ $(= \boxdot((\neg\phi \wedge 0) \vee (\phi \wedge 1)))$. From the previous result, we can use (A4) in place of (#7). An instance of (A4) is $\phi\dot{\mathbb{E}}1 \supset 1 \preceq \phi)$ which is just $\boxdot\phi \supset 1 \preceq \phi)$, which was to be shown.
3. Schema (#0) is a consequence of the axiomatisation of PK′, excluding (#0). See [10], where inter alia it is shown that (#0) is a theorem of QP.

The proof of the theorem is now almost immediate: QP and PK′ share the same rules of inference, (MP) and (Nec). As well, any axiom of QP is an axiom of PK′ (or derivable, in the case of (A4)/(#7). Consequently for $\phi \in \mathcal{L}_{\mathsf{QP}}^-$ a proof of ϕ in QP is a proof in PK′. For the reverse direction, the axioms of PK′ that aren't axioms of QP are (PMN), (Symm) and (#0). However, (PMN) is trivial in QP, since $\Box$ is a defined operator; (Symm) is now a consequence of (#6’); and (#0) is noted above to be redundant. It follows then that for $\phi \in \mathcal{L}_{\mathsf{QP}}^-$ a proof of ϕ in PK′ is transformable to a proof of ϕ in QP. $\square$

Proof of Theorem 5.8.

1 Define $G : \mathcal{L}_{\mathsf{AX}_{\text{meas}}} \to \mathcal{L}_{\mathsf{LQP}}^-$ as follows:

$$G(p) = p\,, \text{for } p \in \mathcal{P}$$
$$G(\neg\phi) = \neg G(\phi)$$
$$G(\phi \vee \psi) = G(\phi) \vee G(\psi)$$
$$G(a \preceq \textstyle\sum_{i=1}^n a_i \cdot w(\psi_i)) = a \preceq \bigoplus_{i=1}^n a_i \cdot G(\psi_i)$$

Notice that the last line defining G makes use of our definition of rational linear inequalities. It is straightforward to verify that for each $\psi \in \mathcal{L}_{\mathsf{AX}_{\text{meas}}}$, we have $\models \phi \equiv G(\phi)$. So G is satisfaction-preserving. It is also straightforward to see that $H := G^{-1}$ is indeed a function with domain $\mathcal{L}_{\mathsf{LQP}}^-$ (by the fact that $\mathcal{L}_{\mathsf{LQP}}^-$ is the non-nested fragment of $\mathcal{L}_{\mathsf{LQP}}$) and that H is also satisfaction-preserving. Hence $\mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}}$ and $\mathcal{L}^{-}_{\mathsf{LQP}}$ are equally expressive over the class of simple models. Further, it is easy to see that each of $G \circ H : \mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}} \to \mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}}$ and $H \circ G : \mathcal{L}^{-}_{\mathsf{LQP}} \to \mathcal{L}^{-}_{\mathsf{LQP}}$ is an identity function.

2 Let L^i for $i \in \{1, 2\}$ be the super-simple model obtained by deleting the relation R^i from the PK model L^i in the proof of Theorem 5.5(2). Take $\chi := p \prec (\neg p \preceq p)$. This is an LQP-formula and, as we saw in the proof of Theorem 5.5(2), we have $L^1, w_1 \models \chi$ but $L^2, w_1 \not\models \chi$. We prove by induction on the construction of formulas $\phi \in \mathcal{L}_{\mathsf{AX}_{\mathrm{meas}}}$ that we have $[\![\phi]\!]_{L^1} = [\![\phi]\!]_{L^2}$. All cases except for one induction step case are straightforward, so let us consider this case. We wish to show that we have

$$[\![d \leq \textstyle\sum_{i \leq n} d_i \cdot w(\psi_i)]\!]_{L^1} = [\![d \leq \textstyle\sum_{i \leq n} d_i \cdot w(\psi_i)]\!]_{L^2}. \tag{26}$$

But by the induction hypothesis, it follows that we have $[\![\psi_i]\!]_{L^1} = [\![\psi_i]\!]_{L^2}$ for each $i \leq n$. So since $P^1_{w_1} = P^2_{w_1}$, it follows that we have $P^1_{w_1}([\![\chi]\!]_{L^1} \cap W) = P^1_{w_1}([\![\chi]\!]_{L^2} \cap W)$ for each $i \leq n$. Applying the definition of satisfaction, (26) follows. □

References

[1] Vaishak Belle, Logic meets probability: towards explainable AI systems for uncertain worlds, in: Carles Sierra (Ed.), Proceedings of the International Joint Conference on Artificial Intelligence, Melbourne, Australia, 2017, pp. 5116–5120.
[2] Patrick Blackburn, Maarten de Rijke, Yde Venema, Modal Logic, Cambridge Tracts in Theoretical Computer Science, vol. 53, Cambridge University Press, Cambridge, UK, 2001.
[3] J.P. Burgess, Probability logic, J. Symb. Log. 34 (2) (1969) 264–274.
[4] B.F. Chellas, Modal Logic, Cambridge University Press, Cambridge, 1980.
[5] Bruno de Finetti, La prévision: ses lois logiques, ses sources subjectives, Ann. Inst. Henri Poincaré 7 (1937) 1–68, Translated into English in H.E. Kyburg, Jr., H.E. Smokler (Eds.), Studies in Subjective Probability, 1964, pp. 93–158.
[6] Bruno de Finetti, La "logica del plausibile" secondo la concezione di polya, in: Atti della XLII Riunione, Societa Italiana per il Progresso delle Scienze, 1951, pp. 227–236.
[7] James P. Delgrande, Bryan Renne, The logic of qualitative probability, in: Q. Yang, M. Wooldridge (Eds.), Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI), AAAI Press, 2015, pp. 2904–2910.
[8] Ronald Fagin, Joseph Y. Halpern, Reasoning about knowledge and probability, J. Assoc. Comput. Mach. 41 (2) (1994) 340–367.
[9] Ronald Fagin, Joseph Y. Halpern, Nimrod Megiddo, A logic for reasoning about probabilities, Inf. Comput. 87 (1990) 78–128.
[10] P. Gärdenfors, Qualitative probability as an intensional logic, J. Philos. Log. 4 (2) (1975) 171–185.
[11] Peter Gärdenfors, Some basic theorems of qualitative probability, Stud. Log. 34 (3) (1975) 257–264.
[12] J.Y. Halpern, D.A. McAllester, Likelihood, probability, and knowledge, Comput. Intell. 5 (2) (1989) 151–160.
[13] Joseph Y. Halpern, Michael O. Rabin, A logic to reason about likelihood, Artif. Intell. 32 (3) (1987) 379–405.
[14] Joseph Y. Halpern, Defining relative likelihood in partially-ordered structures, J. Artif. Intell. Res. 7 (1997) 1–24.
[15] Joseph Y. Halpern, Reasoning About Uncertainty, MIT Press, 2003.
[16] Matthew Harrison-Trainor, Wesley H. Holliday, Thomas F. Icard III, Preferential structures for comparative probabilistic reasoning, in: Satinder P. Singh, Shaul Markovitch (Eds.), Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI Press, 2017, pp. 1135–1141.
[17] Andreas Herzig, Modal probability, belief, and actions, Fundam. Inform. 57 (2–4) (2003) 323–344.
[18] Wesley H. Holliday, Thomas F. Icard III, Measure semantics and qualitative semantics for epistemic modals, in: Proceedings of SALT, vol. 23, 2013, pp. 514–534.
[19] G.E. Hughes, M.J. Cresswell, A New Introduction to Modal Logic, Routledge, London and New York, 1996.
[20] Charles H. Kraft, John W. Pratt, A. Seidenberg, Intuitive probability on finite sets, Ann. Math. Stat. 30 (2) (1959) 408–419.
[21] Henry E. Kyburg Jr., Choh Man Teng, The logic of risky knowledge, reprised, Int. J. Approx. Reason. 53 (3) (2012) 274–285, Precisely Imprecise: A Collection of Papers Dedicated to Henry E. Kyburg, Jr.
[22] H.E. Kyburg Jr., Believing on the basis of evidence, Comput. Intell. 10 (1) (1994) 3–20.
[23] Wolfgang Lenzen, Glauben, Wissen und Wahrscheinlichkeit — Systeme der epistemischen Logik, Springer Verlag, Wien & New York, 1980.
[24] Louis Narens, Theories of Probability: An Examination of Logical and Qualitative Foundations, World Scientific, 2007.
[25] David Poole, Alan Mackworth, Artificial Intelligence: Foundations of Computational Agents, Cambridge University Press, New York, NY, USA, 2010.
[26] Stuart J. Russell, Peter Norvig, Artificial Intelligence: A Modern Approach, third edition, Pearson Education, 2010.
[27] Stuart Russell, Unifying logic and probability, Commun. ACM 58 (7) (2015) 88–97.
[28] Dana Scott, Measurement structures and linear inequalities, J. Math. Psychol. 1 (2) (1964) 233–247.
[29] Krister Segerberg, Qualitative probability in a modal setting, in: J.E. Fenstad (Ed.), Proceedings of the Second Scandinavian Logic Symposium, in: Studies in Logic and the Foundations of Mathematics, vol. 63, Elsevier, 1971, pp. 341–352.
[30] Jan van Eijck, Bryan Renne, Belief as willingness to bet, E-print, arXiv.org, arXiv:1412.5090 [cs.LO], December 2014.