

Bridging Ranking Theory and the Stability Theory of Belief

Eric Raidl¹ · Niels Skovgaard-Olsen^{1,2}

Received: 31 July 2015 / Accepted: 3 September 2016
© Springer Science+Business Media Dordrecht 2016

Abstract In this paper we compare Leitgeb’s stability theory of belief (Annals of Pure and Applied Logic, 164:1338-1389, 2013; The Philosophical Review, 123:131-171, 2014) and Spohn’s ranking-theoretic account of belief (Spohn, 1988, 2012). We discuss the two theories as solutions to the lottery paradox. To compare the two theories, we introduce a novel translation between ranking (mass) functions and probability (mass) functions. We draw some crucial consequences from this translation, in particular a new probabilistic belief notion. Based on this, we explore the logical relation between the two belief theories, showing that models of Leitgeb’s theory correspond to certain models of Spohn’s theory. The reverse is not true (or holds only under special constraints on the parameter of the translation). Finally, we discuss how these results raise new questions in belief theory. In particular, we raise the question whether stability (a key ingredient of Leitgeb’s theory) is rightly thought of as a property pertaining to belief (rather than to knowledge).

Keywords Belief · Probability · Lottery paradox · Stability theory · Ranking theory · Knowledge · Lockean thesis · Odds-threshold

Eric Raidl’s work was supported by the Deutsche Forschungsgemeinschaft (DFG) Research Unit FOR 1614. Niels Skovgaard-Olsen’s work was supported by a grant to Wolfgang Spohn from the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program *New Frameworks of Rationality* (SPP 1516).

✉ Eric Raidl
eric.3.raidl@uni-konstanz.de
Niels Skovgaard-Olsen
niels.skovgaard.olsen@psychologie.uni-freiburg.de

¹ Department of Philosophy, University of Konstanz, Postfach D9, 78457 Konstanz, Germany

² Department of Psychology, University Freiburg, Freiburg, Germany

1 Introduction

The lottery paradox is an old problem which is thought to show that probability theory is incapable to account for full, rational belief as distinct from certainty. In particular, the lottery paradox shows that it is impossible for the following to be jointly satisfied: (1) Belief is closed under conjunction, (2) there is a threshold such that subjective probability above the threshold defines belief. One radical answer to this problem is to abandon probability theory as representing degrees of belief and represent degrees of belief by a function with other properties. Precisely this route was taken by Spohn [14] which led him to introduce ranking functions as representing degrees of belief. These avoid the lottery paradox and allow representing full, rational belief without confusing it with certainty. Recently, the above interpretation of the lottery paradox has been challenged by Leitgeb [3–5]. The lottery paradox arises under the assumption that there is one threshold that defines belief for all probabilities (of an agent). Leitgeb weakens this thesis by reversing quantifiers – each probability may have its own threshold – and shows that this weakened thesis is compatible with belief being closed under conjunction.

The two solutions differ. Spohn's solution is ranking-theoretic, Leitgeb's is probabilistic. Spohn draws the radical conclusion that the lottery paradox is a reason for abandoning probability theory when it comes to modelling epistemic concepts. In contrast, Leitgeb attempts to save probability theory for this purpose. In this article, we argue that the two solutions, although different in nature, are structurally comparable. For this structural comparison, we introduce a translation scheme between (finite) ranking functions and (finite) probabilities. On this background, we show that (modulo translation) the probability functions satisfying Leitgeb's new probabilistic threshold thesis (with belief being closed under conjunction) correspond to ranking functions satisfying Spohn's liberalised ranking-theoretic definition of belief. The inverse is only true under an additional assumption, either on the ranking functions or on the translation. We interpret this as meaning either that Leitgeb's solution is more restrictive (it imposes further constraints) or as meaning that the parameter of the translation is subject to constraints (if the mentioned probability-to-ranking embedding is required to be an isomorphism for belief).

The plan of the paper is as follows. We start by introducing Spohn's ranking-theoretic account of belief and Leitgeb's probabilistic account of belief and discuss them as solutions to the lottery paradox (Section 2). To compare both theories, we introduce a novel translation scheme between finite ranking (mass) functions and probability (mass) functions (Section 3). We state some crucial consequences of this translation scheme which motivate it independently of the purpose for which it is introduced here: (1) a new probabilistic belief notion and (2) the parallelism of ranking-theoretic and probabilistic updating. Using (1), we explore the relation between Leitgeb's probabilistic theory of belief and Spohn's ranking-theoretic account of belief (Section 4). Finally, we indicate how these results raise new questions in belief theory (Section 5). Proofs are provided in the Appendix A.

2 Two Solutions to the Lottery Paradox

In this section we state the lottery paradox and present Spohn’s (Section 2.1) and Leitgeb’s solutions (Section 2.2) to it. In all what follows, we assume algebras to be finite. Finite algebras can always be considered to be power-set algebras over a finite base. The sets in the algebra are interpreted as propositions, sets of possible worlds or, if preferred, sets of state descriptions. A proposition is therefore represented as a set of possible worlds in which the sentence expressing the proposition is true. The term “probability” refers to subjective probability representing the credences or degrees of belief of a rational agent. That the probability of A is p , $P(A) = p$, according to an agent, means that the agent assigns credence p to the proposition represented by A .

Assume the following belief characterisation: the proposition A is *believed* by an agent iff¹ her subjective probability P for A is greater than some fixed threshold t . Formally,

Thesis 1 (probabilistic strong threshold, $ST(P)$) There is a unique $t \in (0, 1)$ such that for all possible probability functions P of an agent over a given algebra \mathcal{A}

$$\forall A \in \mathcal{A}, \quad \mathbf{B}(A) \text{ iff } P(A) > t. \tag{2.1}$$

The lottery paradox shows that it is impossible for belief \mathbf{B} to satisfy both $ST(P)$ and the following proposition, quantified over the domain \mathcal{A} :²

(B4) If $\mathbf{B}(A)$ and $\mathbf{B}(B)$ then $\mathbf{B}(A \cap B)$. (Closure under conjunction)

Fix t and consider a lottery with $n > (1 - t)^{-1}$ tickets that has only one winner. The uniform probability $1/n < 1 - t$ is a reasonable probability assignment for the propositions that ticket i will win ($1 \leq i \leq n$). For every ticket, the probability that it won’t win is $(n - 1)/n > t$. The agent will therefore believe for each ticket that it won’t win. If belief is closed under conjunction then the agent will also believe that no ticket will win. Yet, this is inconsistent with the initial belief that one of the tickets will win. Therefore (B4) and $ST(P)$ are inconsistent (modulo the axioms of probability). The uniform probability provides the counterexample.

One can draw two different conclusions from this paradox, if (B4) and a “threshold belief notion” of the form of Eq. 2.1 are to be maintained. Either (1) the probability P in $ST(P)$ has to be replaced by something else. Or (2) $ST(P)$ has to be weakened.³ Spohn [14, 15] develops the first option, Leitgeb [3, 4] the second, as we will now show. In all that follows, Ω is a finite base, i.e. a set of possible worlds (or state descriptions), and $\mathcal{A} = \wp(\Omega)$ is the power-set algebra over Ω , representing all available propositions given the base. We start with Spohn’s solution.

¹ ‘if and only if’.

²We keep the number “4”, in “(B4)”, in correspondence to the enumeration of the defining properties of a proper filter (cf. Definition 6), i.e. (1) non-empty, (2) proper, (3) upwards closed and (4) closed under finite intersections. This is condition (b) in [15, def. 4.5].

³A third solution is to choose a particular non-uniform probability assignment.

2.1 Spohn’s Solution

In what follows, let $\mathbb{R}_\infty^+ := \mathbb{R}^+ \cup \{+\infty\}$ the non-negative real numbers to which we add infinity.⁴

Definition 1 κ is a (real-valued) **ranking mass** (r.m.) over (the finite base) Ω if $k : \Omega \rightarrow \mathbb{R}_\infty^+$ is a function with at least one 0. The associated (real-valued) **negative ranking function** (n.r.f.) κ over $\mathcal{A} = \wp(\Omega)$ is induced by

$$\kappa(A) = \min_{w \in A} k(w) \tag{2.2}$$

with $\kappa(\emptyset) = +\infty$.

A negative real ranking function can also be defined directly:

Definition 2 Let $\mathcal{A} = \wp(\Omega)$ be a finite algebra. κ is a **(real-valued) negative ranking function** if it is a function $\kappa : \mathcal{A} \rightarrow \mathbb{R}_\infty^+$ that satisfies:

1. $\kappa(\Omega) = 0$, (Minimum)
2. $\kappa(\emptyset) = +\infty$, (Maximum)
3. for all $A, B \in \mathcal{A}$, $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$. (Finite minimativity)

The function κ should be thought of as a *grading of doubt, disbelief* (taking to be false), which is maximal for the empty set \emptyset (representing the contradiction) and minimal for the full set Ω (representing the tautology). Definitions 1 and 2 are equivalent in the finite case. If κ is induced by a ranking mass – in the sense of Definition 1 – then it is a real-valued ranking function – in the sense of Definition 2. Conversely, if κ is a real-valued ranking function on a finite algebra – in the sense of Definition 2 – then it is induced by a unique ranking mass – in the sense of Definition 1.

A slightly different definition starts with a positive ranking function:

Definition 3 Let $\mathcal{A} = \wp(\Omega)$ be a finite algebra. β is a **(real-valued) positive ranking function** (p.r.f.) if it is a function $\beta : \mathcal{A} \rightarrow \mathbb{R}_\infty^+$ that satisfies:

1. $\beta(\emptyset) = 0$, (Minimum)
2. $\beta(\Omega) = +\infty$, (Maximum)
3. for all $A, B \in \mathcal{A}$ $\beta(A \cap B) = \min\{\beta(A), \beta(B)\}$. (Finite minimativity)

The function β should be thought of as a *grading of belief*, which is minimal for the empty set and maximal for the full set. β and κ are interdefinable by $\beta(A) = \kappa(\overline{A})$. κ is a n.r.f. if and only if β is a p.r.f. [15, 5.11]. This results from the fact that the three laws (minimum, maximum, minimativity) in Definitions 2 and 3 are duals of each other, given interdefinability.⁵

Spohn defines belief by

⁴For the definitions in this section see [15, ch. 5].

⁵This is clear for minimum and maximum. For minimativity it follows from the De Morgan law $A \cup B = \overline{\overline{A} \cap \overline{B}}$.

Definition 4 Let κ be a n.r.f. over (the finite algebra) $\mathcal{A} = \wp(\Omega)$ and $z \geq 0$. The **belief predicate** \mathbf{B}_κ^z is defined by

$$\forall A \in \mathcal{A}, \quad \mathbf{B}_\kappa^z(A) \text{ iff } \kappa(\bar{A}) > z. \tag{2.3}$$

This defines full belief explicitly in terms of degrees of doubt (negative ranks). Full belief is relative to a particular ranking function and possibly, but not necessarily, to a particular threshold. Spohn’s initial definition fixes the threshold to $z = 0$. However, allowing $z > 0$ does not change the structure of ranking-based belief, nor the ranking-updating dynamics, as noticed independently by Hans Rott [10, p. 327 and 336] and Matthias Hild, as noted by Spohn [15, p. 76].

The connection between belief as defined by \mathbf{B}_κ^z and ST becomes clear when we switch to the dual function of κ , the positive ranking function:

Definition 5 Let β be a p.r.f. over (the finite) $\mathcal{A} = \wp(\Omega)$ and $z \geq 0$. The **belief predicate** \mathbf{B}_β^z is defined by

$$\forall A \in \mathcal{A}, \quad \mathbf{B}_\beta^z(A) \text{ iff } \beta(A) > z. \tag{2.4}$$

\mathbf{B}_β^z and \mathbf{B}_κ^z are extensionally the same belief predicate for $\beta(A) = \kappa(\bar{A})$. To arrive at ST for positive ranking functions, the following remarks are necessary. If we accept that the threshold z can possibly be above zero, two dimensions of variation arise. The threshold of a single agent could vary over time (*diachronic variability*) and thresholds could vary from agent to agent (*intersubjective variability*). Let us exclude intersubjective variability for the moment. Diachronic variability means that an agent could have different standards over time to set the threshold which determines her ranking-theoretic belief set. It would nonetheless be the case that for a particular agent, a unique threshold is fixed for every specific moment of time (*static invariability*). Such a threshold would therefore determine belief with respect to any of the possible ranking functions the agent could choose among at that specific time. If one threshold z is fixed for all possible β from which the agent can actually choose (i.e. given static invariability), Definition 5 turns into an ST(β), where the probability function P is replaced by a positive ranking function β :

Thesis 2 (ranking strong threshold (ST(β))) There is a unique $z \geq 0$ such that for all positive ranking functions β over a given algebra \mathcal{A} ,

$$\forall A \in \mathcal{A}, \quad \mathbf{B}_\beta^z(A) \text{ iff } \beta(A) > z. \tag{2.5}$$

In the above thesis, the threshold and the possible ranking functions are, as before, relative to an agent. The thesis is to be read counterfactually: Although the agent might have chosen another ranking function, the threshold which determines ranking-theoretic belief would have remained the same. Spohn’s original belief definition with $z = 0$ is such a strong threshold thesis. In fact, it is even stronger, since it excludes not only static variability, but also diachronic and intersubjective variability.

However, the generalised belief definition adopted here, with $z \geq 0$, does not require z to be fixed. This allows not only intersubjective and diachronic variability,

but also static variability. Therefore a weak threshold thesis (WT(β)) is not prohibited by Spohn’s account and even required in the liberalised account, if scale invariance is taken seriously. In effect, standard ranking theory is scale-invariant – any rescaling of a ranking function represents the same ranking function, and their dynamic evolution is parallel or covariant. Therefore the generalised theory with the liberalised belief notion should remain scale-invariant as well. However, this is only possible if the threshold itself is rescaled in the same manner. The threshold $z = 0$ is the only scale-invariant (and therefore intersubjectively constant) threshold; all other thresholds are only covariant. In this sense, the liberalised non-zero threshold could be taken as the scale. Hence different agents can choose different thresholds. In fact, they should, since the threshold 1 has another meaning for κ than for 2κ . Similarly, the same agent should choose different (and covariant) thresholds for rescalings of her initial ranking function. Under scale invariance, κ and 2κ represent the same ranking function, therefore an agent changing her scale from 1 to 2 should also change her threshold from z to $2z$.

We will now outline how Spohn’s proposal solves the lottery paradox. For this, we need to consider two additional properties a belief predicate \mathbf{B} over an algebra \mathcal{A} can have. We call it a “belief operator” to distinguish it from a mere belief predicate. A belief predicate just expresses belief. The following definition puts additional conditions on such a predicate.

Definition 6 A belief predicate \mathbf{B} over the algebra \mathcal{A} is a **belief operator** over \mathcal{A} if it satisfies, for all $A, B \in \mathcal{A}$:

- (B1) $\mathbf{B}(\Omega)$, (Tautologies are believed)
- (B2) $\neg\mathbf{B}(\emptyset)$, (Contradictions are not believed)
- (B3) if $A \subseteq B$ and $\mathbf{B}(A)$ then $\mathbf{B}(B)$, (Closure under implication)
- (B4) if $\mathbf{B}(A)$ and $\mathbf{B}(B)$ then $\mathbf{B}(A \cap B)$. (Closure under conjunction)⁶

In algebraic terms, \mathbf{B} is a belief operator if and only if $\mathcal{B} = \{A \in \mathcal{A} : \mathbf{B}(A)\}$ is a proper filter over \mathcal{A} .

Definition 7 A belief predicate \mathbf{B} over an algebra \mathcal{A} is a **core (belief) notion** over \mathcal{A} if there exists a (unique) non-empty $C \in \mathcal{A}$, called the *belief core* of \mathbf{B} , such that

$$\forall A \in \mathcal{A}, \quad \mathbf{B}(A) \text{ iff } C \subseteq A. \tag{2.6}$$

In algebraic terms, \mathbf{B} is a core notion with core C if and only if the corresponding \mathcal{B} is a principal filter generated by $C \neq \emptyset$. If \mathcal{A} is finite then \mathbf{B} is a core belief notion over \mathcal{A} iff it is a belief operator over \mathcal{A} . This is due to the fact that all filters \mathcal{B} over a finite algebra are principal filters, generated by the core $\bigcap \mathcal{B}$.

Lemma 1 Let $\mathcal{A} = \wp(\Omega)$ be a finite algebra, κ a n.r.f. and $z \geq 0$. \mathbf{B}_κ^z from Definition 4 is a core notion with core $[\leq z]_\kappa := \{w \in \Omega : \kappa(w) \leq z\}$.

⁶The lottery paradox may of course be circumvented by dropping (B4).

Proof We have $\mathbf{B}_\kappa^z(A)$ iff $\kappa(\bar{A}) > z$ iff $\forall w \in \Omega (w \in \bar{A} \rightarrow \kappa(w) > z)$ iff $\forall w \in \Omega (\kappa(w) \leq z \rightarrow w \in A)$ iff $[\leq z]_\kappa \subseteq A$. Therefore $[\leq z]_\kappa$ is the core. \square

Putting these results together, we can conclude that the ranking-theoretic belief notion is closed under finite conjunctions and cannot lead to a lottery paradox. Spohn’s explicit solution (for $z = 0$) is the uniform ranking mass $\mathbf{0}$. In effect, this implies that the agent is neutral about each ticket losing while believing that one ticket will win.

2.2 Leitgeb’s Solution

Whereas Spohn replaces probability by a ranking function in ST, Leitgeb solves the lottery paradox by sticking to probability and weakening ST:

Definition 8 Let \mathbf{B} be a belief predicate and P a probability function, both over the algebra \mathcal{A} . (\mathbf{B}, P) satisfy the **Lockean thesis** if there is a $t \in (0.5, 1]$ such that

$$\forall A \in \mathcal{A}, \quad \mathbf{B}(A) \text{ iff } P(A) \geq t. \tag{2.7}$$

The Lockean thesis (as well as the strong threshold thesis for \geq instead of $>$) can be satisfied trivially for the threshold $t = 1$. However, this is not the intended meaning. Leitgeb has weakened ST. Instead of requiring that *there is* a threshold such that *for all* probabilities Eq. 2.7 holds, he only requires $\text{WT}(P)$, the Lockean thesis quantified over all probabilities over a given algebra: *for each* probability *there is* a threshold such that Eq. 2.7 holds.⁷

Thesis 3 (probabilistic weak threshold, $\text{WT}(P)$). For all possible probability functions P of an agent over a given algebra \mathcal{A} , there is a threshold $t \in (0, 1)$ such that

$$\forall A \in \mathcal{A}, \quad \mathbf{B}(A) \text{ iff } P(A) > t. \tag{2.8}$$

Reversing the order of quantification from $\exists \forall$ to $\forall \exists$ also avoids the lottery paradox. The paradox invokes a situation where an agent tells her opponent the threshold first, who can then devise a lottery for which the uniform attribution jointly with (B4) implies an inconsistency (modulo the axioms of probability). This is not possible under WT, since the threshold is chosen as a function of the probability (or constrained by the chosen probability).

More precisely, Leitgeb [4] shows that the *Lockean thesis* and (B4) can be jointly satisfied. First he proves that the Lockean thesis is equivalent (modulo (B4) and the probability axioms) to a certain kind of stability of the belief core.⁸ The stability notion involved is preservation of (sufficiently) high probability under conditionalisation.

⁷Leitgeb also weakens “ $>$ ” to “ \geq ”. This is irrelevant in the finite case, when $t = 1$ is excluded.

⁸Note that the Lockean thesis already implies (B1 – 3) by the axioms of probability.

Definition 9 Let P be a probability function over the algebra \mathcal{A} . $A \in \mathcal{A}$ is P -stable,⁹ if

$$\forall B \in \mathcal{A} ((B \cap A \neq \emptyset \wedge P(B) > 0) \rightarrow P(A|B) > 1/2). \tag{2.9}$$

The equivalence Leitgeb [4, Theorem 1] proves is¹⁰

Theorem 1 Let $\mathcal{A} = \wp(\Omega)$ finite, \mathbf{B} a belief operator (with core C) and P a probability function over \mathcal{A} . Then the following are equivalent:

1. (\mathbf{B}, P) satisfy the Lockean thesis for $t = P(C)$.
2. (i) C is P -stable, and (ii) if $P(C) = 1$ then C is the smallest probability-1 set.

Proof 2, Appendix. □

The extreme case $t = 1$ satisfies (B4) and the Lockean thesis trivially. More profoundly, Leitgeb shows that the Lockean thesis can be satisfied non-trivially, i.e. for $t \in (0.5, 1)$. By an illuminating construction he is even able to show that, in the finite case, almost all¹¹ probabilities over a given algebra satisfy the Lockean thesis non-trivially.

We will use the abbreviation $p_A := \min_{w \in A} P(w)$ – the minimum atomic probability in A . This minimum always exists for a finite probability space. To prove the existence of such non-trivial probabilities, the following equivalent formulation of P -stability is useful:

Lemma 2 Let P be a probability function over a finite algebra $\mathcal{A} = \wp(\Omega)$. The following are equivalent if $P(A) \neq 1$:

1. $A \in \mathcal{A}$ is P -stable.
2. $p_A > P(\bar{A})$.

If $P(A) = 1$, the equivalence still holds with \geq instead of $>$.

Proof Lemma 5, in the Appendix, proves a more general equivalence. □

Constructing a probability P that satisfies the Lockean thesis non-trivially now becomes easy. Instead of P -stability, it suffices to apply the equivalent definition (2) to a non-trivial core $\emptyset \neq C \subsetneq \Omega$ and require $P(C) < 1$: Choose $p_C \in (0, 1)$, intended to be the minimum atomic probability in C , and another value $q \in (0, 1)$,

⁹Cf. [5]. In terms of Leitgeb’s [3] more general definition of P -stability, this is P -stability $_{\frac{1}{2}}$.

¹⁰Leitgeb shows a slightly different, but equivalent, result: Let \mathbf{B} be a predicate over the algebra \mathcal{A} and P a function over \mathcal{A} with values in $[0, 1]$. Then (a,b,c) is equivalent to (b, d), with letters referring to the following assumptions: (a) \mathbf{B} is a core belief notion, with core $C_{\mathbf{B}}$, (b) P is a probability, (c) together they satisfy the Lockean thesis for $t = P(C_{\mathbf{B}})$, (d) there is a unique non-empty $C = C_{\mathbf{B}}$ such that (2) of Theorem 1. We chose the presentation in Theorem 1 for the following reasons, it is simpler and outsources two hypothesis: (b) and (a’) that \mathbf{B} is a belief operator. (b) figures in both equivalent statements of Leitgeb and (a’) on a finite algebra is equivalent to (a).

¹¹In the sense of the Lebesgue measure on the probability hypersurface.

intended to be the probability of C , such that $1 - q < p_C \leq q < 1$. Then $q \in (0.5, 1)$. Atomic values $P(w)$ for $w \in \overline{C}$ can now be distributed arbitrarily, as long as they are non-negative and their sum is $1 - q$. Atomic values $P(w)$ for $w \in C$ may be distributed arbitrarily, as long as they are non-negative, their sum is q and the minimal value is p_C .

3 The Atomic Translation Scheme

Both solutions to the lottery paradox allow for a threshold notion of belief, in the sense of Eq. 2.1 (c.f. Eqs. 2.5 and 2.8), and satisfy (B4). This suggests that there is a relation between them. Yet, as they stand, the two theories are incomparable because they are formulated in different languages (ranking-theoretic vs. probabilistic). In order to compare them, we introduce a translation scheme between the intended models of the two theories (ranking functions vs. probability functions). We thereby obtain a probabilistic analogue to Spohn’s theory of belief.

On a finite algebra $\mathcal{A} = \wp(\Omega)$, a probability function P is uniquely characterised by the probability mass vector $\mathbf{p} = (p_w)_{w \in \Omega}$ if we set $P(\{w\}) = p_w$. Similarly, a ranking function κ is uniquely characterised by the ranking mass vector $\mathbf{k} = (k_w)_{w \in \Omega}$ if we set $\kappa(\{w\}) = k_w$. To avoid complications, we also write $P(w) = p_w = P(\{w\})$ and $\kappa(w) = k_w = \kappa(\{w\})$.

Let κ be a n.r.f. and P a probability function over a finite algebra $\mathcal{A} = \wp(\Omega)$ and $a \in (0, 1)$. The *ranking-to-probability translations* are induced by the mass translations

$$\Phi_a : \kappa(w) \mapsto P_\kappa^a(w) := \frac{a^{\kappa(w)}}{Z(a, \kappa)}, \tag{3.1}$$

where $Z(a, \kappa) = \sum_{v \in \Omega} a^{\kappa(v)}$ is a normalising factor. The probability mass is then extended to a probability function at the algebraic level (by applying additivity). The *probability-to-(real-valued)-ranking translations* are induced by the mass translations

$$\Psi_a : P(w) \mapsto \kappa_P^a(w) := \log_a \frac{P(w)}{p}, \tag{3.2}$$

where $p := \max_{v \in \Omega} P(v)$. This negative real-valued ranking mass is then extended to a negative real-valued ranking function (applying minimativity). When κ , or κ and a , are given by the context, we write $Z(a)$ and Z , respectively, instead of $Z(a, \kappa)$. When we intend to refer to a specific translation (Φ_a, Ψ_a) in the family, we use the term *a-translation*. When we intend to refer to the family of translations, we simply speak of *the translation scheme*. We call the translation scheme *atomic* because it takes place at the atomic level and cuts off the algebraic level, so to speak. The complete probability or ranking functions are generated from the mass functions. The functional correspondence only holds at the atomic level; not at the level of the algebra.

However, we have:

Lemma 3 *Let $\mathcal{A} = \wp(\Omega)$ be a finite algebra and $a \in (0, 1)$. Then:*

1. *If κ is a n.r.f. over \mathcal{A} then P_κ^a is a probability function over \mathcal{A} .*

2. If P is a probability function over \mathcal{A} , then κ_p^a is a n.r.f. over \mathcal{A} .
3. $\Phi_a\Psi_a = id$ and $\Psi_a\Phi_a = id$.
4. $\kappa(w) > \kappa(v)$ iff $P(w) < P(v)$ (for $P = P_\kappa^a$ or $\kappa = \kappa_p^a$).

Proof 3, Appendix. □

By (1) and (2) the translations are well behaved. By (3) the translations are inverses to each other. We can therefore write $\Phi^{-1} := \Psi$. By (4), the translations invert the strict order relation. Therefore the 0 atomic κ -rank is mapped to the maximum atomic P_κ^a -probability. Conversely, the maximum atomic P -probability is mapped to an atomic κ_p^a -rank of value zero. We exclude the two extreme translations, $a = 0, 1$, for the following reason: $a = 1$ maps any ranking function to the same probability, viz. the uniform probability on the support of κ , i.e. $\Omega(\kappa) = \{w \in \Omega : \kappa(w) < +\infty\}$. And $a = 0$ maps all ranking function sharing the same zeros to the uniform probability with support on these zeros. None of these extreme translations mirrors (or folds over) the strict order, this is, they violate condition (4).

The above translation scheme has several consequences:

- Analogues:** Every ranking-theoretic notion which is definable in terms of ranking masses has a *probabilistic analogue* (and vice versa).
- Isomorphism:** The class of ranking masses (over a finite algebra) is ‘isomorphic’, by any particular translation, to the class of probability masses (over the same algebra).¹² The isomorphism extends to the reduced language with mass functions and without sums (provided the ranking-mass language $<, \min, \max, \infty$ is mapped to the probability-mass language $>, \max, \min, 0$). This is, the probability mass P_κ^a and the ranking mass κ are order-dual functions.
- Simulation:** Finite probability theory can be simulated (or defined) within finite ranking theory (in a sufficiently rich language, including a constant a , exponentiation and addition). Conversely, finite real-valued ranking theory can be simulated (or defined) in finite probability theory (in a sufficiently rich language, including a constant a , the logarithm to base a and \max).

Each of these consequences opens a wide field of unsettled questions for exact formal and philosophical analysis. In what follows, we concentrate on some crucial instances of the (finite) ranking–probability parallelism.¹³ We derive (1) the probabilistic analogue to the ranking-theoretic notion of belief and discuss (2) the probabilistic analogues to the dynamics of ranking-updating. Definition (1) leads to a new probabilistic belief notion. Definition (2) leads to a new probabilistic update,

¹²This is due to the fact that Φ_a induces a bijection between $\mathbb{K}^n := \{\mathbf{k} \in (\mathbb{R}_{\infty}^+)^n : \exists i < n, k_i = 0\}$ and $\mathbb{P}^n := \{\mathbf{p} \in [0, 1]^n : \sum p_i = 1\}$.

¹³This parallelism differs from the parallelism proposed by Pearl to explain the parallelism of the operations (($\min, +, -$) v.s. ($+, \times, \div$)). The later is based on equating algebraic rank values with the standard part of a logarithm to an infinitesimal base of a non-standard probability: $\kappa(A) = \text{st log}_i P(A) - \text{see [15, Theorem 10.1]}$.

as well as to a parallelism between ranking- and probabilistic-updating methods. The more complex relation between Leitgeb’s notion of belief and the ranking-theoretic notion of belief (in its probabilistic guise) is analysed in the next section (Section 4), based on (1).

3.1 Belief Analogue

A first surprising consequence of the translation scheme is that it induces a new probabilistic notion of belief.

Ranking-theoretic belief is defined (in Definition 4), on the atomic level, as

$$\forall A \in \mathcal{A}, \quad \mathbf{B}(A) \text{ iff } \forall w \in \bar{A}, \kappa(w) > z, \text{ for some fixed } z \geq 0. \quad (3.3)$$

The probabilistic analogue is

$$\forall A \in \mathcal{A}, \quad \mathbf{B}(A) \text{ iff } \forall w \in \bar{A}, P(w) < y, \text{ for some fixed } y \in (0, 1). \quad (3.4)$$

The two definitions are linked under the translation scheme. The ranking-to-probability translation inverts the order, therefore we have $\kappa(w) > z$ iff $P_{\kappa}^a(w) < a^z/Z(a)$. Taking $y(z, a) = a^z/Z(a)$ yields the expected isomorphism. The induced probabilistic belief notion is of course a core notion. Whereas the core of the ranking notion of belief is $[\leq z]_{\kappa} = \{w \in \Omega: \kappa(w) \leq z\}$, the core of the probabilistic analogue is $[\geq y]_P = \{w \in \Omega: P(w) \geq y\}$. Whereas z functions as a ranking-theoretic *upper* bound for atoms in the core, y functions as a probabilistic *lower* bound for atoms in the core. Note, however, that the so-defined probabilistic belief notion does not necessarily satisfy the probabilistic Lockean thesis. Furthermore we should keep in mind that the *atomic* threshold y should be distinguished from the *algebraic* threshold t .

3.2 Dynamical Analogues

A second instance of probabilistic analogues induced by the translation scheme are the probabilistic analogues to ranking dynamics. This induces a *dynamics parallelism*. The conditional ranking corresponds to the conditional probability, simple result-oriented ranking-conditionalisation corresponds to simple Jeffrey conditionalisation (on a partition into two cells) and general result-oriented ranking conditionalisation corresponds to general Jeffrey conditionalisation. On the other hand, evidence-oriented conditionalisation corresponds to a probability update which resembles, but is not exactly, Field conditionalisation [8, pp. 301].

More precisely, for conditionalisation, this means that the following diagram commutes:

$$\begin{array}{ccc}
 \kappa & \xrightarrow{a} & P \\
 \downarrow A & & \downarrow A \\
 \kappa_A & \xrightarrow{a} & P_{\kappa_A} = P_A
 \end{array} \quad (3.5)$$

Here $\kappa_A = \kappa(\cdot|A)$ is the conditional ranking function given A . This new function is defined whenever $\kappa(A) < +\infty$ by $\kappa_A(w) = \kappa(w) - \kappa(A)$ for $w \in A$ and by $\kappa_A(w) = +\infty$ for $w \notin A$. Furthermore $P = P_\kappa^a$ and $P_A = P(\cdot|A)$ is the conditional probability and $P_{\kappa_A} = P_{\kappa_A}^a$.

The commutativity of the above diagram may be read in several ways, because of the possibility to invert the horizontal arrows. First, it means that conditionalisation is the probability update induced by the conditional ranking function. For this, invert the northern horizontal arrow and take the route $P \xrightarrow{\Phi_a^{-1}} \kappa \xrightarrow{A} \kappa_A \xrightarrow{\Phi_a} P_{\kappa_A}$. Conversely, it also means that the conditional ranking function is the ranking-update induced by probabilistic conditionalisation. For this, invert the southern horizontal arrow and take the route $\kappa \xrightarrow{\Phi_a} P \xrightarrow{A} P_A \xrightarrow{\Phi_a^{-1}} \kappa_{P_A} = \kappa_A$. Finally, it means that probabilistic and ranking conditionalisation evolve *in parallel* under any fixed a -translation. This parallelism is represented by the two downward arrows, the fixed translation being the horizontal arrows both indexed by the same a . Similar commutative diagrams may be stated for result-oriented and evidence-oriented conditionalisation and their probabilistic analogues.

3.3 Finiteness

Our restriction to finite algebras may be criticised for lack of generality, since ranking theory can be developed for complete algebras over infinite bases [15] and probability theory can be developed for σ -algebras over infinite bases. There are several replies to this objection.

First, we should note that the essential reason for the restriction to finite algebras is a technical one. The translation we propose between ranking theory and probability theory is mainly restricted to the finite case. In effect, at least one direction of the translation stated in Section 3 may cause problems as soon as we leave the finite realm. Furthermore, we should also note that some extensions to infinite algebras are possible. In particular, the direction from probability to ranking functions allows translations of probability densities into ranking masses, whenever the inverse order $>$ over the density values is a well-ordering (that is, there is no infinite increasing chain of mass values – which permits only approximations of many interesting densities). For this direction it is therefore not so much the algebra that needs to be finite, but the density values that need to admit an inverse well-ordering. For the other direction, the condition is more complex: the ranking function has to admit a ranking mass with the particular feature of being integrable and normalisable (in the measure-theoretic sense).

Second, the technical restriction stated above might not be that much of a philosophical restriction since the probabilistic theory of belief formulated by Leitgeb is presented on finite algebras, and most examples given by Spohn refer to finite algebras. Third, since ranking theory and its belief notion are invariant under ranking embeddings, we may always consider finite coarse-grainings of ranking functions, if the latter operate on infinite algebras. This secures also the other direction of the translation.

Finally, the finiteness assumption can be pragmatically justified and might even be preferable. In effect, when an agent forms her beliefs for purposes of practical reasoning, we may assume that she is only interested in certain propositions arising as supersets over a finite partition of the base of the fine-grained algebra. The cells of the partition represent the questions which are of pragmatic interest to her at that moment. If the agent’s rational capacities slightly deviate from those of an ideal rational agent, in particular in computational power, then finite coarse-graining may even be a desirable requirement. Therefore the restriction to finite algebras may be welcomed pragmatically. At worst our restriction means that the parallel established here between ranking theory and probability theory, and the comparison of the belief notions, are confined to situations where agents restrict their considerations for pragmatic reasons to finite coarse-grainings of an algebra.

4 Ranking-Theoretic and Probabilistic Belief

In this section we use the probabilistic analogue of ranking-belief to state the relation between Spohn’s ranking-theoretic belief and Leitgeb’s probabilistic belief.

4.1 Framework for Comparing Belief Models

Let us start with setting up the three different belief models to be compared.

Definition 10 Let $\mathcal{A} = \wp(\Omega)$ finite, F a function and \mathbf{B} a predicate over \mathcal{A} .

1. (\mathbf{B}, F) is a **Spohn belief model** if F is a ranking function over \mathcal{A} and \mathbf{B} is defined by Definition 4, i.e. there is a $z \geq 0$ such that for all $A \in \mathcal{A}$

$$\mathbf{B}(A) \text{ iff } F(\bar{A}) > z. \tag{4.1}$$

2. (\mathbf{B}, F) is a **Leitgeb belief model** if F is a probability function over \mathcal{A} , \mathbf{B} is closed under conjunction and (\mathbf{B}, F) satisfies the Lockean thesis, i.e. there is a $t \in (0.5, 1]$ such that for all $A \in \mathcal{A}$

$$\mathbf{B}(A) \text{ iff } F(A) \geq t. \tag{4.2}$$

3. (\mathbf{B}, F) is an **atomic belief model** if there is a $y \geq 0$ such that for all $A \in \mathcal{A}$

$$\mathbf{B}(A) \text{ iff } \forall w \in \Omega (F(w) \geq y \rightarrow w \in A). \tag{4.3}$$

We call y , in the atomic belief model, the **atomic threshold**, to distinguish it from an **algebraic threshold** as in the Lockean thesis. Whenever $F := \kappa$ is a ranking function, we write (\mathbf{B}, κ) and whenever $F := P$ is a probability function, we write (\mathbf{B}, P) . If $F := P$ is a probability, atomic belief is called **atomic probabilistic belief**. From the previous remarks it is clear that for the above belief models, we always have that

1. \mathbf{B} is a belief operator,
2. \mathbf{B} is a core belief notion (provided $\mathcal{A} = \wp(\Omega)$ is finite).

Additionally $(\mathbf{B}_\kappa^z, \kappa)$ satisfies the Lockean thesis, but for the corresponding positive ranking function β , and $(\mathbf{B}_\kappa^z, \kappa)$ is an atomic belief model, but with \geq replaced by \leq .

Atomic probabilistic belief has not been extensively studied in the literature,¹⁴ but it can be motivated as follows. Suppose the rational agent knew which among the possible worlds Ω is the actual world – say w . In that case, closure under implication recommends that she should also believe every proposition that is true of w , i.e. every proposition to which w belongs. The totality of her beliefs is therefore obtained by closure under implication starting with $\{w\}$. However, in a state of uncertainty (or doubt), the ideally rational agent does not know which possible world is the actual world. Yet, the agent may have ways to select some best candidates for the actual world. This is exactly what atomic belief does. The atomic probabilistic belief model recommends that she selects the set of her best candidates for the actual world. The “best” candidates are those that have a probability score above some fixed subjective atomic threshold. The principle of closure under implication generates the totality of the beliefs, i.e. the agent should believe every proposition that is true for all her candidates. Her set of candidates for the actual world will then be a belief core for her beliefs. In logical terms, the agent believes every sentence which is implied by the disjunction of the descriptions of the best candidates for the actual world.

Let us now investigate the relation between the different belief notions.

Theorem 2 *Let $a \in (0, 1)$ and $\mathcal{A} = \wp(\Omega)$ be finite.*

1. *If (\mathbf{B}, κ) is a Spohn belief model, with threshold z , then (\mathbf{B}, P_κ^a) is an atomic probabilistic belief model (with atomic threshold $y = a^z/Z(a)$) and with the same core).*
2. *If (\mathbf{B}, P) is an atomic probabilistic belief model, with threshold y , then (\mathbf{B}, κ_P^a) is a Spohn belief model (with threshold $z = \log_a \frac{y}{p}$ and with the same core).*

Proof 4, Appendix. □

Spohn belief models and atomic probabilistic belief models are one and the same notion, up to a translation. The translation would be perfect and would extend to the algebraic level, if we were to replace the law of additivity by the law corresponding to minimativity – the resulting law would impose ‘maximativity’ (see below, Definition 11).

Theorem 3 *An atomic probabilistic belief model is a Leitgeb belief model if and only if $y > P(\overline{C})$ (with y the atomic threshold and C the belief core).*

¹⁴However, see [7, equation 17], and [6]. We became aware of this work after having developed atomic probabilistic belief as a probabilistic analogue to ranking-theoretic belief. The odds-threshold rule formulated by Lin–Kelly amounts to believing a proposition A iff it is implied by the disjunction of the most plausible propositions of a partition $(C_j)_{j \in J}$. The most plausible C_i ’s are those that score not below a certain fraction of the most probable C_j , i.e. $P(C_i) \geq \max_{j \in J} P(C_j)/x$, where x might be a function of i . For fixed x the odds-threshold rule determines an atomic probabilistic belief model with threshold $y = \max_{j \in J} P(C_j)/x$.

Proof 5, Appendix. □

This can be interpreted as saying that, in the finite case, Leitgeb’s theory of probabilistic belief implies the probabilistic analogue of ranking-theoretic belief. The reverse implication, however, does not hold. Not every atomic probabilistic belief model is a Leitgeb model. Therefore an atomic probabilistic belief model does not in general satisfy the Lockean thesis. For this, the atomic threshold would have to satisfy the additional separation condition $y > P(\overline{C})$.

The above Theorems 2 and 3 have as an immediate corollary:

Corollary 1 *Let $a \in (0, 1)$ and $\mathcal{A} = \wp(\Omega)$ finite.*

1. *If (\mathbf{B}, P) is a Leitgeb belief model with core C , then (\mathbf{B}, κ_p^a) is a Spohn belief model with threshold $z = \log_a \frac{pC}{p}$.*
2. *If (\mathbf{B}, κ) is a Spohn belief model with core C , $z = \max_C \kappa(w)$ and $x = \min_{\overline{C}} \kappa(w)$, then (\mathbf{B}, P_κ^a) is a Leitgeb belief model, provided:*

$$a > |\overline{C}|^{\frac{1}{z-x}}. \tag{4.4}$$

Proof 6, Appendix. □

Let us summarise the two central results (Corollary 1). Every Leitgeb belief model corresponds to a class of equivalent Spohn belief models. Conversely, not every Spohn model corresponds to a class of equivalent Leitgeb belief models. Rather, a Spohn belief model corresponds to such a class only under a suitable constraint (4.4) on the translation. From the point of view of Leitgeb’s belief model, we may say that this imposes a constraint on the admissible translations of ranking functions into probability functions. Yet, if all translations are considered equally admissible, then Leitgeb’s belief is stricter than Spohn’s belief. Rott [11] already provided an argument why Leitgeb’s belief model imposes constraints which are too strict for belief, and in Section 4 further arguments will be added.

4.2 Atomic Probabilist Belief and Leitgeb Belief Compared.

Since we have a probabilistic analogue to Spohn’s ranking-theoretic belief – namely atomic probabilistic belief – we can investigate the relation between this belief model and Leitgeb’s belief model directly.

Lemma 4 *Atomic probabilistic belief satisfies for all $A \in \mathcal{A}$: if $\mathbf{B}(A)$ then $P(A) \geq P(C)$, where C is the belief core.*

Proof By sub-additivity and because C is a core. □

It is clear that Leitgeb belief is an atomic belief. However the reverse is not true. By the above lemma, atomic probabilistic belief satisfies the left-to-right implication of the Lockean thesis when the threshold is taken to be $t = P(C)$

and if $P(C) \in (0.5, 1]$. However, in general the probability value of the belief core of an atomic probabilistic belief model will not necessarily be greater than 0.5, nor will it generally satisfy the right-to-left implication of the Lockean thesis.

Theorem 4 *An atomic probabilistic belief model (\mathbf{B}, P) , with core C and atomic threshold y is a Leitgeb belief model iff one of the following (equivalent) conditions holds:*

1. *it is **separating**, i.e. $y > P(\overline{C})$,*
2. *(i) C is P -stable and (ii) if $P(C) = 1$, then C is the smallest probability-1 set.*
3. *(\mathbf{B}, P) satisfies the Lockean thesis,*
4. *(\mathbf{B}, P) satisfies the right-to-left implication of the Lockean thesis.*

Proof (1) by Theorem 3. (1) implies (2) by lemma 2. (2) implies (3) by Theorem 1. (3) trivially implies (4). And (4) implies (3), since an atomic probabilistic belief already satisfies the other implication of the Lockean thesis. Therefore (4) implies (1), since (3) implies P -stability (Theorem 1) and therefore separation (Lemma 2). \square

Atomic probabilistic belief and Leitgeb's belief fall apart exactly when the Lockean thesis (more exactly, the right-to-left implication) or the separation assumption is not satisfied.

The purpose of this article is neither to defend the atomic probabilistic belief model, nor to reject it. Atomic probabilistic belief was solely introduced to compare, in the same probabilistic framework, Spohn's ranking-theoretic belief and Leitgeb's probabilistic belief. Atomic probabilistic belief, as already said, is the probabilistic analogue to ranking-theoretic belief (as long as the algebra is finite). This being said, let us continue with the comparison.

4.3 Examples

The essential difference between atomic probabilistic belief and Leitgeb belief consists in the fact that atomic belief is not necessarily separating (see previous Theorem 4). This difference can be explained intuitively as follows. Consider all atoms of an algebra and their values. Order them according to their value on a line (from left to right) such that low atoms come before higher atoms. Draw an arbitrary second line, orthogonal to the first line. This line separates low atoms from high atoms, with respect to an arbitrarily selected standard for low/high. The arbitrary line separates the belief core (to the right) from its complement (to the left). According to atomic probabilistic belief, the line can be drawn everywhere, whereas, according to Leitgeb's belief, the separation line can only be drawn such that the minimum of the high atoms (fixed in this way) is higher than the sum of the (corresponding) low atoms, i.e. the line *separates* the minimum of the core atoms from the sum of the core complement atoms. This separation condition introduces a requirement which is not an atomic

requirement, but an algebraic one, since the complement of the core is (in general) not an atom.

Some examples may be useful to see when an atomic probabilistic belief model is not a Leitgeb belief model.

Example 1 Let $\Omega = \{w_1, \dots, w_4\}$ and $\mathcal{A} = \wp(\Omega)$. Assume the following probability mass, with the separation line being marked by “||”, the core appearing to the right of this line:

$$\frac{w_1 \mid w_2 \parallel w_3 \mid w_4}{0.1 \mid 0.2 \parallel 0.3 \mid 0.4}$$

Then $C := \{w_3, w_4\}$ is an admissible core for atomic probabilistic belief if the atomic threshold is some $y \in (0.2, 0.3]$. This is an inadmissible core for Leitgeb belief, since the value of w_3 is not greater than the value of $\bar{C} = \{w_1, w_2\}$. A Leitgeb belief model would only accept the three-element core $\{w_2, w_3, w_4\}$ or the trivial four-element core. Notice further that belief generated by the atomic probabilistic belief core C does not satisfy the Lockean thesis for probability. Since C is the core, the algebraic threshold would have to satisfy $t \leq 0.7$. But then the set $\{w_1, w_2, w_4\}$ also meets the criterion. Yet it is not believed according to the atomic probabilistic belief model, because $\{w_3, w_4\}$ is not a subset of it.

Slightly lowering the value of w_1 or w_2 , and readjusting those of w_3 and w_4 accordingly (e.g. lower w_2 to 0.1999 and raise w_3 to 0.3001), would make C acceptable as a core for a Leitgeb belief model and would also imply that the atomic probabilistic belief model satisfies the Lockean thesis, since the atomic probabilistic belief model now satisfies the criterion for being a Leitgeb belief model. Let us therefore look at a more extreme example:

Example 2 Assume the following probability mass over $\Omega = \{w_1, \dots, w_4\}$:

$$\frac{w_1 \mid w_2 \parallel w_3 \mid w_4}{0.1 \mid 0.2 \parallel 0.200001 \mid 0.499999}$$

Here $C := \{w_3, w_4\}$ is an admissible core for atomic belief, for an atomic threshold $y \in (0.2, 0.20001]$. But it is an inadmissible core for a Leitgeb belief model, since the value of w_3 is not greater than the value of $\bar{C} = \{w_1, w_2\}$. Again the belief set generated by the belief core C does not satisfy the Lockean thesis for probability. Since C is the core, we have to choose the threshold $t \leq 0.7$. But then $\{w_1, w_2, w_4\}$ would again be believed according to the Lockean thesis, contrary to an atomic probabilistic belief model.

In both examples the core has a probability > 0.5 . But atomic probabilistic belief admits of cases where the core has a probability smaller than (or equal to) 0.5!

Example 3 Assume the following probability mass over $\Omega = \{w_1, \dots, w_8\}$:

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2

Here $C := \{w_7, w_8\}$ can count as a core for atomic belief, if $y \in (0.1, 0.2]$. But it cannot count as a core for Leitgeb belief, for similar reasons as above (since $0.6 \not\leq 0.2$). Similarly, the generated belief does not satisfy the Lockean thesis, since $\{w_1, \dots, w_4\} \subsetneq \overline{C}$ already has the same probability as the core.

In this example the core has a lower probability than 0.5! The example may be varied by changing the values of w_1, \dots, w_6 slightly by, say, raising to no more than 0.19, but with a total of 0.6 remaining, and the values of w_7, w_8 are changed accordingly but remain above 0.19.

Since Example 3 violates the Lockean thesis and therefore the conditions for Leitgeb belief, one pressing question is whether there are any intuitive cases that would meet this description. Suppose that a witness to a murder is given the following questions (because they are discriminatory for the potential suspects): (1) Did the killer wear a red coat? (2) Did the killer wear an orange coat? (3) Did the killer wear a yellow coat? (4) Did the killer wear a green coat? (5) Did the killer wear shorts? As it was dark, the witness is unsure about her testimony and reports that the only thing she remembers is that the killer wore a green coat. She shows no preference w.r.t. shorts, but exhibits a clear preference for the models in which the killer wore a green coat – say w_7, w_8 – over the models in which the killer wore some other colour of coat, although she will readily admit that she might be mistaken, since her probability for that proposition is below 0.5.

If this is the only lead for the detective to go on, then the following situation might be seen as paradoxical. On the one hand, a Lockean thesis with a low threshold of $t \in (0.5, 0.6]$ recommends that the detective should believe that the killer's coat was not green, as $P(\{w_1, w_2, w_3, w_4, w_5, w_6\}) \geq t > 0.5 > P(\{w_7, w_8\})$. Leitgeb's account, on the other hand, recommends that the detective should not believe anything other than the disjunction of all propositions. In effect, Ω is the only set which, taken as core, meets the criterion of separation. In contrast, the atomic probabilistic belief model licenses the detective to form a full belief that the killer wore a green coat, and to use it as an assumption for further inquiry, in spite of the substantial uncertainty that is associated with it. If $p(w) = 0.2$ is translated into $k(w) = 0$, and $p(w) = 0.1$ into $k(w) = 1$, then ranking theory will agree with this verdict. Moreover, due to its minimum aggregation rule, it will prevent the proposition. "The killer did not wear a green coat" from being believed since the negation of this proposition has rank 0. Standard ranking belief with threshold $z = 0$ will in fact agree with this verdict for every ranking function obtained from the above P under any translation with $a \in (0, 1)$.

4.4 Criteria for Deciding between Atomic Belief and Leitgeb Belief

Given the previous Theorem 4 one could argue against Leitgeb’s belief (and for atomic belief) as follows:

- I. The separation assumption is counterintuitive.
- II. Belief is not, and should not necessarily be, stable.
- III. The Lockean thesis (for probability) is too strong.

(I) The separation assumption can be challenged by an extreme counterexample.

Example 4 Consider an algebra $\mathcal{A} = \wp(\Omega)$ with $2^7 = 128$ atoms, generated from a language with 7 propositional letters.

- 1. The first 90 each have a probability 0.01 (yielding a total of 0.9),
- 2. the next 14 have a probability of 0.002 each (yielding a total of 0.028),
- 3. the next 24 have a probability of 0.003 each (yielding a total of 0.072).

Atomic belief allows the first 90 atoms to constitute a belief core for an atomic threshold $y \in (0.003, 0.01]$. Even though the probability of C is quite high, Leitgeb belief prohibits this, since the probability of these atoms is much lower than $P(\overline{C}) = 0.1$.

This example can be exaggerated as much as we want by adding sufficiently many 1’s after “0.01”, yielding a total of $0.9 \dots 9$ with as many 9’s as desired, and correspondingly adding the same number of zeros before the 2 and the 3 appearing in “0.002” and “0.003”, respectively. It seems odd to refuse to believe C .

From the perspective of the separation condition C cannot be believed, because the single probabilities of atoms making up C does not exceed the probability of \overline{C} . The separation condition can be reformulated as follows:

- C is an acceptable belief core iff for any $w \in C$, $P(C) > 1 - P(w)$.

In other words, C is an acceptable belief core if and only if any atom making up this core is essential to the core. That is to say, if this atom were “removed” from all the possibilities, then the resulting probability $P(\Omega \setminus \{w\})$ would drop below the probability of the core, and therefore below the highest level the Lockean threshold can be.

From the perspective of the atomic threshold thesis, C can be believed. The only criterion that counts is a singular evaluation of worlds for being candidates for the actual world. The atomic threshold is very low in this case, because the agent has an extreme degree of uncertainty w.r.t. her best candidates for the actual world. In this sense the example is extreme. The agent has so little background knowledge to constrain her choice of the best candidates for the actual world that she ends up selecting a set of 90 possibilities that she is unable to discriminate between. The atomic threshold thesis can also be read in terms of essentiality.

- C is an acceptable belief core iff for every $w \in C$ and $v \in \overline{C}$, $P(w) > P(v)$.

In other words, C is an acceptable belief core if and only if any atom in the core is preferable to any atom outside the core.

(II) The following argument could be given against imposing stability of belief (a more detailed line of argument against stability is given in the next section). In the simplest case, stability of belief under updating on evidence, which is not belief-contravening, means the following:

- If $\mathbf{B}(A)$ and E is “compatible” with \mathbf{B} , then $\mathbf{B}_E(A)$.

Here \mathbf{B}_E is the updated belief after accepting E . For the moment no particular updating method is assumed. If \mathbf{B} is a core belief notion (as is the case on a finite algebra if belief is assumed to be consistent and closed under implication) then “compatibility” means that $E \cap C \neq \emptyset$, where C is the non-empty core of \mathbf{B} . The above updating requirement then becomes

- If $C \subseteq A$ and $C \cap E \neq \emptyset$ then $C_E \subseteq A$.

Here C_E is the updated belief core. Again no particular updating method has been assumed. In particular, we obtain:

- If $C \cap E \neq \emptyset$ then $C_E \subseteq C$.

However, this requirement, known as preservation, is a quite conservative belief notion. It means that the agent considers as candidates for the new core only those possibilities which are already among C . This is a bad requirement if the agent was previously misinformed or had wrong beliefs. If an agent starts with wrong beliefs (the actual world is not in the core), then whatever future true but not belief-contravening evidence she might receive will never put her back on the right track. She will restrict her belief core more and more, but will never arrive at a core that contains the actual world. In fact, preservation has to be abandoned as soon as evidence is belief-contravening. So why not abandon it right away?

(III) The Lockean thesis is problematic for the previous Example 4. If we start with the above belief core C , then, according to the Lockean thesis, there is a threshold t such that $P(C) \geq t$. But then, due to construction, the agent could remove ten atoms out of C and replace them by the negation of C . This is, $(C \setminus \{w_1, \dots, w_{10}\}) \cup \overline{C} = \Omega \setminus \{w_1, \dots, w_{10}\}$ would also be believed. From the perspective of the Lockean thesis the atoms constituting the core are not essential. This is the essential difference between the Lockean thesis and the atomic probability: for the atomic probability belief, only atomic probabilities and atoms matter, for the Lockean thesis atoms do not matter and only algebraic probabilities matter.

A proponent of Leitgeb belief could of course employ the converse argument, arguing for the Lockean thesis, or for stability. One reply to this might be to note that atomic probabilistic belief satisfies the Lockean thesis for a function distinct from probability, namely possibility, and is stable under conditionalising that function.

Definition 11 Let $\mathcal{A} = \wp(\Omega)$ be a finite algebra. ρ is a **possibility function** iff it is a function $\rho: \mathcal{A} \rightarrow [0, 1]$ that satisfies, for all $A, B \in \mathcal{A}$:

1. $\rho(\Omega) = \max \rho,$ (maximum)
2. $\rho(\emptyset) = 0,$ (minimum)
3. $\rho(A \cup B) = \max\{\rho(A), \rho(B)\}.$ (finite maximativity)

The definition of a possibility function slightly differs in the first axiom from that of a *possibility measure*, as given in [1], where (1) is replaced by the stricter condition $\max \rho = 1$. (The reason for this is that possibility theory was initially conceived as a fuzzy extension of logic and not as modelling epistemic concepts such as belief.) A probability mass function p over Ω induces a possibility function over $\wp(\Omega)$ by

$$\rho(A) = \max_{w \in A} p(w) \tag{4.5}$$

with $\rho(\emptyset) = 0$. Given a possibility function, we can define an *impossibility function* by $\alpha(A) = \rho(\overline{A})$. Impossibility functions have not been investigated, since traditional possibility theory rather considers the *necessity measure* $\eta := 1 - \alpha$, for α generated from a possibility measure ρ . Then η is a T-norm and ρ the corresponding T-conorm. We prefer considering α directly, since it is a direct mirror of a positive ranking function:

Definition 12 Let $\mathcal{A} = \wp(\Omega)$ be a finite algebra. α is an **impossibility function** iff it is a function $\alpha: \mathcal{A} \rightarrow [0, 1]$ that satisfies, for all $A, B \in \mathcal{A}$:

1. $\alpha(\emptyset) = \max \rho,$ (maximum)
2. $\alpha(\Omega) = 0,$ (minimum)
3. $\alpha(A \cap B) = \max\{\alpha(A), \alpha(B)\}.$ (finite maximativity)

Possibility functions and impossibility functions are interdefinable in the same manner as are negative and positive ranking functions. In fact, possibility and impossibility functions just *are* the *full algebraic* images of negative ranking and positive ranking functions, respectively, under the atomic translation, now extended to the algebraic level. If κ is a negative ranking function then ρ_κ^a is a possibility function, and if ρ is a possibility function then κ_ρ^a is a negative ranking function. In both cases the translation extends to the algebraic level and therefore to the algebraic threshold (in contrast to probability).

Spohn [15, Section 11.8] comments on the isomorphism between negative ranking functions and possibility measures. Yet he does not investigate the weakening of the assumption $\max \rho = 1$, the analogue to positive ranking functions or the possibilistic analogue to the ranking-theoretic belief definition, nor the possibility of having both possibility and probability as two measures arising from the same point function and therefore coexisting.

By the same move as Spohn defines ranking-theoretic belief, we may define possibilistic belief: There is a $y \geq 0$ such that

$$\forall A \in \mathcal{A} \quad \mathbf{B}(A) \quad \text{iff} \quad \rho(\overline{A}) < y \quad [\quad \text{iff} \quad \alpha(A) < y \quad \text{iff} \quad \eta(A) > 1 - y]. \tag{4.6}$$

Therefore belief defined by a possibility function is closed under conjunction and is in fact a belief operator and also satisfies the Lockean thesis (for the corresponding necessity function η). A possibility function, an impossibility function, a necessity function and a probability function can all coexist over the same probability mass. The atomic probabilistic belief, defined by means of the probability mass, will be identical to the belief defined by means of the possibility, impossibility or necessity function. Atomic possibilistic belief now satisfies the Lockean thesis, is invariant under (possibility function) embeddings and is stable under possibility conditionalisation $\rho(A|B) = \frac{\rho(A \cap B)}{\rho(B)}$. Instead of satisfying these conditions with respect to probability, it satisfies them with respect to the corresponding necessity/possibility function!

5 Comparative Discussion

As we have seen, the probabilistic version of ranking-theoretic belief introduced above is more general than the stability-theoretic belief in that the set of models satisfying the latter is a proper subset of the set of models satisfying the former. Accordingly, one pressing question for the comparative discussion is whether stability ought to be considered a necessary condition of rational belief, as there are models satisfying our probabilistic version of ranking theory that lack this property.

5.1 Stability as a Property of Knowledge Rather Than of Rational Belief

In what follows, we are going to motivate a dissenting response to the above question. To a certain extent, the underlying suspicion is already present in [16], where it is implicitly suggested that while stability may plausibly be viewed as a property of *knowledge*, it may be too strict a requirement for *rational belief*. To be sure, the explanations differ on whether a belief should be stable w.r.t. updating on true information or updating on propositions compatible with the agent's current beliefs [9, 12]. But the philosophical motivation presented below for Leitgeb's stability requirement is very similar to the possible motivation for a stability account of knowledge that Rott [9] presents (without endorsing it, however), as we shall see.

In taking up Spohn's suspicion, we start by looking at the intuitive motivation for imposing the stability requirement in Leitgeb's theory. To this aim, we can profitably turn to [5]. Here a generalized version of the stability theory is presented as the *Humean thesis* (p. 12):

$$\forall A \in \mathcal{A} \quad \mathbf{B}(A) \text{ iff } \forall D \in \mathcal{D} (P(D) > 0 \rightarrow P(A|D) > r). \quad (5.1)$$

This generalises the $\frac{1}{2}$ -stability condition (2.9), in that it introduces a generalised threshold $r \geq \frac{1}{2}$, but also because conditionalisation is required to remain stable with respect to a yet-to-be-specified set of conditions.

Interestingly, Leitgeb [5] suggests that we view the set of propositions \mathcal{D} , w.r.t. which the probability of our belief candidate A should remain stably high under conditionalisation, as a set of *potential defeaters* for A . That is, the members of \mathcal{D}

can be conceived as propositions that potentially lower the probability of A under conditionalisation. Yet, these propositions don't count as *sufficient reasons* against A , insofar as A still retains a probability above the threshold after the conditionalisation has taken place. (Of course, \mathcal{D} also contains propositions that raise the probability of A or leave it unchanged, which are not even reasons against A to begin with. But the stability theory is only interesting for probability-lowering cases, where the issue arises of whether the threshold is still met after conditionalisation).

Viewed from this perspective, a substantial issue confronting generalised stability theory is which properties to ascribe to \mathcal{D} , or in other words, what should count as a potential defeater of A . In [4], it was suggested that its members need merely be consistent with A . In [5], several other alternatives are presented. On the version he defends, \mathcal{D} is to be characterized as the set of those of the agent's *doxastic possibilities*, which is defined by lack of full belief in their negations (i.e. $D \in \mathcal{D}$ iff $\neg \mathbf{B}(\bar{D})$). Consequently, potential defeaters are now being defined relative to the belief set of the agent and no longer relative to A . Given the latter interpretation of the stability theory, a case can be made for stability being a requirement better suited for *knowledge* than for *rational belief per se*.

The first step is to reinterpret the Lockean threshold. As Leitgeb [3] points out, in transitioning from degrees of belief (expressed in numerical values) to full belief (expressed in qualitative terms), information is lost. So why bother? One point of having a set of full beliefs is that it allows us to demarcate the propositions we want to use as unqualified factual assumptions in further practical and theoretical reasoning. When deliberating about how to act, we inevitably make assumptions about how the world is. And when deliberating about what to believe, we inevitably leave some of the planks of Neurath's boat intact while replacing others. That is, for both practical and theoretical reasoning, we need to keep some factual assumptions¹⁵ fixed while deliberating. We can thus view the Lockean threshold as our decision rule for determining these assumptions.

The second step consists in drawing a distinction between *individual* and *interpersonal* reasoning and arguing that the issue of what it is rational to believe arises in the first instance in the context of the agent's own (individual) reasoning.¹⁶ Accordingly, the crucial question at stake when theorising about *full rational belief* is, when ought the agent to unqualifiedly endorse A as a factual assumption in her own practical and theoretical reasoning? In contrast, the topic of *knowledge* requires the introduction of a context of *interpersonal* reasoning. There the question becomes, when ought other agents with diverging doxastic perspectives unqualifiedly endorse A as a factual assumption in *their own* practical and theoretical reasoning?

¹⁵By 'factual assumption' something stronger than 'supposing for the sake of the argument' is meant. Factual assumptions are assumptions about facts which the agent is willing to rely on in her practical and theoretical reasoning – the agent uses them as a basis of her actions and for further inquiries.

¹⁶In choosing to theorise about the individual context of reasoning of the rational agent who *has* the beliefs (as opposed to focusing on the intersubjective context of *ascribing* beliefs), we take Leitgeb [5, p. 25] to be following this order of explanation.

Having drawn this distinction, we can regard the set \mathcal{D} of potential defeaters as a set of potential *objections* against A that may be entertained by other doxastic perspectives (irrespective of whether they are salient to the agent). That is, in deciding whether others, who potentially hold divergent background beliefs, ought to endorse A unqualifiedly as a factual assumption in their own practical and theoretical reasoning, our agent needs to take into account potential objections that they may have, and to ensure that A retains a probability above the threshold after conditionalisation. In [5], this is formalised as conditionalising A not only on the agent's own beliefs but also on propositions that she is neutral about. In [4], this is presented as conditionalising on everything that is consistent with A .

Hence, the stability requirement finds a natural motivation in the context of interpersonal reasoning. Yet, this motivation applies primarily to the standards of justification introduced when dealing with *knowledge*. In contrast, the same motivation cannot be alluded to if we are mainly concerned with rational belief in the context of individual reasoning. For as long as we are merely dealing with the question of which propositions the agent ought to treat as factual assumptions in her own practical and theoretical reasoning, there is no need to go beyond conditionalising on her own background beliefs in determining whether the threshold is met. It is only when determining whether others with divergent background beliefs ought to treat the same propositions similarly that there is a need to conditionalise on the members of \mathcal{D} .

This suggests that Leitgeb's theory is based on attributing something to rational belief which should more properly be viewed as belonging to knowledge. Indeed, the philosophical interpretation we have just given of Leitgeb's stability thesis is very similar to the motivation that Rott expresses for introducing stability as a requirement on knowledge:

Like Plato, Lehrer suggests a dialogical construal of the stability idea. The believing subject is imagined as being engaged in a dialogue with a *critic* (a Socratic dialogue partner) who tries to undermine the subject's beliefs. Only if the subject wins the dialogue in the sense that he successfully defends his belief against all the critic's objections, can that belief be called *knowledge*. Rott [9, p. 471]

So whereas other philosophers have sought to make stability a distinguishing feature of knowledge as opposed to rational belief, Leitgeb's account blurs this distinction.

5.2 Counterexample

To drive the point home it may be useful to consider a counterexample showing that while stability may plausibly be associated with knowledge, it imposes restrictions too severe when used as a constraint on rational belief. What we need is a case where Leitgeb's theory forbids believing a seemingly rationally permitted proposition due to the fact that its probability is not stably high under conditionalising on some far-out possibility that we would ordinarily ignore. We are therefore looking for an outlandish possibility that Leitgeb's theory forces us to take into account – in spite of the fact that we would normally set it aside.

Hypotheses concerning the existence of parallel universes seem to fit the bill nicely. In [2], one finds an in-depth treatment of the many ways in which contemporary physics naturally leads to hypotheses about parallel universes (e.g. through quantum mechanics, the potentially infinite expansion of space, inflationary cosmology and string theory). Of course such hypotheses are in a precarious state. Not only do they lack supporting empirical evidence of their own and concern highly counterintuitive matters, but they deal with states of affairs that are epistemically inaccessible to us (and according to some versions, even necessarily so). Greene thus wisely begins his book with the following cautious statement:

The subject of parallel universes is highly speculative. No experiment or observation has established that any version of the idea is realized in nature. So my point in writing this book is not to convince you that we're part of a multiverse. I'm not convinced – and, speaking generally, no one should be convinced – of anything not supported by hard data. Greene [2, p. 8]

Let us then assume that the rational agent we are modelling is following the epistemic norm of not being convinced (i.e. not having full beliefs), without support by hard data, of speculative hypotheses which stand in a tension with our intuitive picture of the world – no matter how well they can be motivated theoretically. As we have empirical evidence neither for nor against the hypothesis, H , that we are living in a multiverse, a rational agent should accordingly be neutral w.r.t. H . Yet, due to the counterintuitive character of H , there will be a great number of things that we would ordinarily find acceptable to believe, which would be undermined if H were true. To design a counterexample that affects both [5] and [4], we only need to identify a proposition which (i) we would ordinarily find rationally permissible to believe, which (ii) is logically consistent with H and which (iii) H nevertheless counts as a sufficient reason against.

One candidate would be, to restrict H to concern the finite versions of a multiverse and allow the number of parallel universes to be so large that it is very likely that there will be an exact copy of our own universe (due to the fact that combinations of finite elements tend to repeat themselves if enough trials are run). In that case, the probability of propositions about things we consider unique in our universe could be lowered by H below the threshold. So let us just take the proposition that there are no exact copies of ourselves – call it p – as our example.

The point is (1) that although we may not be said to *know* p (on some strict notion of knowledge), given that the multiverse hypothesis can be motivated in nine different ways by some of our best theories in physics [2], it certainly appears *rationally permissible* to believe p given our background beliefs, and (2) that the only evidence we have against p consists in speculative scenarios that are themselves (most likely) beyond our epistemic reach.¹⁷ Then the problem with the stability theory of belief is that on reasonable assumptions it forces us to take such outlandish scenarios into account to evaluate whether the threshold is still satisfied. In contrast, on ranking

¹⁷Yet, as Greene [2, ch. 11] carefully argues, there may be some propitious circumstances under which versions of the multiverse hypothesis would turn out to have testable consequences.

theory the agent only needs to consider her own background beliefs in evaluating whether A deserves to be endorsed in an unqualified sense.

5.3 Relativity of Belief to Partitioning

A further battleground where the differences between ranking-theoretic belief and stability-theoretic belief have to be fought out is their behaviour under embeddings. As Leitgeb [4] is keenly aware, one of the perhaps more problematic features of his theory is that it introduces a relativity of belief, not only to the probability function, but also to the partitioning. As a result, different belief sets will be generated for an agent, depending on which questions are posed to her, and logical inferences across contexts are not licensed unrestrictedly (p. 160). In contrast, a central feature of ranking theory is the following invariance principle:

The propositional attitudes, their contents, and their static and dynamic laws must be so conceived as to be invariant under coarse- and fine-graining of the underlying conceptual and propositional framework. Spohn [15, p. 68]

Accordingly, the definition of a ranking function is independent of the algebra, and the dynamic laws of ranking theory are invariant under embeddings of ranking functions (algebraic embeddings that conserve the ranks of the previous coarse-grained ranking function while fine-graining it). Such ranking-embeddings also conserve belief, if z remains fixed. Similar invariances hold for probabilistic (Jeffrey-) conditionalisation w.r.t. measure-space embeddings (i.e. algebraic embeddings that conserve the values of the previous coarse-grained probability function). By contrast, invariance does not hold for Leitgeb's probabilistic account of belief. In particular, on Leitgeb's account, belief is not invariant under probabilistic embeddings.

In an attempt to motivate this language sensitivity of his account, Leitgeb [4, Section 4] argues that the relativity to partitioning works to his theory's advantage when it comes to sorting out the different intuitions at work in the lottery paradox. On the one hand, there is the intuition that insofar as we have no grounds for discriminating between the different tickets, and a contradiction is produced if we believe of every individual ticket that it will not win, we must treat each ticket the same way and hence ultimately suspend belief about the possibility that it is a losing ticket. When a fine-grained partitioning is chosen, this is exactly the outcome that the stability theory will produce. In this case, the Lockean threshold is 1 (*ibid.*). On the other hand, there is the intuition that it is so much more unlikely that the Lottery ticket w_1 wins than that the winner is one of the remaining tickets that we should be allowed to believe of w_1 that it will lose. When a coarse-grained partitioning is chosen, where all the other possibilities are lumped together into one possibility $\{\{w_1\}, \{w_2, \dots, w_n\}\}$, this is exactly the result delivered by the stability theory, since $\{w_2, \dots, w_n\}$ now makes up the least-believed proposition, which determines the threshold (modulo subjective probability).

In contrast, the ranking-theoretic response to the lottery paradox (for $z = 0$) is that consistency is to be maintained through the uniformly-0 ranking function, where all propositions about a particular ticket losing and their complements are assigned a rank of zero, and the only non-zero rank is assigned to \emptyset . Such an agent will remain

neutral about each ticket's losing and will believe that the lottery has at least one winner.

It would thus seem that here the stability theory has a comparative advantage over ranking theory. However, setting aside the special case of the lottery paradox, Leitgeb has not yet shown that partitioning-dependency is generally an attractive feature of the theory when applied to more mundane examples. So, here there is a large unmet burden of justification for proponents of the stability theory.¹⁸

To enhance the intuitive appeal of the ranking-theoretic response to the lottery paradox, we remark that this theory is not restricted to treating the domain of neutrality as a single point value. Indeed, any threshold $z > 0$ can be chosen such that belief is defined by $\mathbf{B}(A)$ iff $\kappa(\bar{A}) > z$. The belief set remains consistent and deductively closed, as noted earlier. This means that the rational agent can continue to differentiate the comparative credibility of various propositions that she neither believes nor disbelieves (e.g. because adopting belief attitudes towards A or \bar{A} would generate an inconsistency, as in the lottery paradox, or because A and its negation merely concern a subject matter that the agent lacks firmly held views about). These differentiations in comparative credibility of propositions that are within the neutrality zone can be thought of as differentiations between how close or remote the propositions are to belief/disbelief [10]. Our rational agent can therefore assign ranks expressing that she is closer to disbelieve that a particular ticket will win, depending on how many tickets there are in the lottery, while keeping inconsistency at bay by never actually disbelieving that this ticket will win. The agent thus modelled will still be very cautious in what she disbelieves, but at least she can display a sensitivity to whether the lottery contains 2 or a million tickets by being ever closer to the verge of disbelieving that ticket i will win in the latter case and closer to believing that it will win in the former case.

All accounts of full belief investigated here involve making some compromise between satisfying the following desiderata: (1) logical closure of rational belief, (2) avoiding the lottery paradox, (3) introducing an algebraic threshold on full beliefs, and (4) invariance of full belief with respect to partitioning. By contrast, both ranking theory and its mirror-image, the generalised possibility theory, are able to satisfy all of these constraints simultaneously. Hence, whether probability theory is a suitable framework for representing full, rational belief depends on whether all of these constraints need to be satisfied.¹⁹

¹⁸This problem carries over to attempts of using Leitgeb's stability thesis as part of an account of knowledge (unless, of course, the latter can be explicated in terms of ranking theory in a way that satisfies the principle of invariance).

¹⁹One particular concern is the problem of logical omniscience. Yalcin [17, 18] has argued that introducing partition-sensitivity of belief makes our models of rational belief less idealised. In the present dialectic this would count as an argument in favour of Leitgeb's stability theory against ranking theory. In [13] a different solution to the problem of logical omniscience was presented, which would require reinterpreting the regulatory norms of ranking theory as providing norms of public, argumentative discourse instead of norms of individual belief.

Due to their aggregation rules at the algebraic level, ranking theory and generalised possibility theory have a structural advantage when it comes to combining an algebraic threshold and a core notion of belief which ensures closure under conjunction. The reason is that the minimum and maximum rules used to form sets of atoms target exactly those properties of individual atoms that make them members of the core. Leitgeb attempts to introduce a constraint on core membership that ensures a parallel relationship on the basis of the sum aggregation rule in probability theory, but it comes at the cost of making belief partition-sensitive.

6 Conclusion

In summary, we have shown that Spohn's ranking-theoretic notion of belief can be translated into the probability calculus by the novel atomic translation scheme introduced in this paper. As we have seen, this translation introduces a new probabilistic notion of belief – atomic probabilistic belief. We have proved that every model satisfying Leitgeb's probabilistic stability theory of belief is a particular instance of atomic probabilistic belief, and its ranking-theoretic translation satisfies Spohn's theory of belief, without the reverse being true. That is, atomic probabilistic belief can be considered as more general than Leitgeb's stability theory. The same holds on the level of theories: Spohn's theory is more general than Leitgeb's stability theory. In our comparative discussion, finally, we have motivated the thesis that stability may more properly be viewed as a property of knowledge than of rational belief. We have also pointed to some of the crucial junctures where the stability theory would have to be tested against the steel of ranking theory.

Acknowledgments We would like to thank Hannes Leitgeb for his helpful comments on an earlier manuscript, Hans Rott for his suggestions and critical remarks and an anonymous referee for pressing us to focus more on the novelty of the translation and its consequences. They all helped to improve the quality of the paper.

Appendix A: Proofs

A.1 Prerequisites

Lemma 5 (cf. lemma 2) *Let P be a probability over a finite algebra $\mathcal{A} = \wp(\Omega)$. If $P(A) \neq 1$, then the following are equivalent:*

1. $A \in \mathcal{A}$ is P -stable,
2. $\forall w \in A, P(w) > P(\bar{A})$,
3. $\min_A P(w) > P(\bar{A})$,
4. For all non-empty $D \subseteq A$ and all $E \subseteq \bar{A}$, $P(D) > P(E)$.

If $P(A) = 1$ these equivalences hold, with $>$ replaced by \geq .

Proof First assume $P(A) \neq 1$.

(1 \Rightarrow 2) Let $A \in \mathcal{A}$ be P -stable. If $A = \emptyset$ then (2) is vacuously satisfied. Suppose $A \neq \emptyset$. Let $w \in A$ and $B = \{w\} \cup \bar{A}$, so that $B \cap A = \{w\} \neq \emptyset$. $P(B) = 0$ is impossible, since then $P(w) = 0$ and $P(\bar{A}) = 0$, contradicting $P(A) \neq 1$. Therefore $P(B) > 0$. Thus the stability condition applies and $P(A|B) > \frac{1}{2}$. Therefore the following are equivalent:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(w)}{P(w) + P(\bar{A})} > \frac{1}{2} \tag{A.1}$$

$$P(w) > \frac{1}{2}(P(w) + P(\bar{A})) \tag{A.2}$$

$$\frac{1}{2}P(w) > \frac{1}{2}P(\bar{A}) \tag{A.3}$$

$$P(w) > P(\bar{A}) \tag{A.4}$$

(2 \Rightarrow 3) If Eq. A.4 holds for all $w \in A$ then it holds for $w \in A$ s.t. $P(w)$ is minimal.

(3 \Rightarrow 4) Let $p_A = \min_A P(w) > P(\bar{A})$. Consider $\emptyset \neq D \subseteq A$ and $E \subseteq \bar{A}$. By sub-additivity and minimality:

$$P(D) \geq p_A > P(\bar{A}) \geq P(E) \tag{A.5}$$

(4 \Rightarrow 1) If $A = \emptyset$, then A is vacuously P -stable. Let us therefore consider $A \neq \emptyset$. Let $B \in \mathcal{A}$, such that $B \cap A \neq \emptyset$ and $P(B) > 0$. Define $D = B \cap A$ and $E = B \setminus D$. We have $\emptyset \neq D \subseteq A$ and $E \subseteq \bar{A}$. So that (4) applies and we have the following equivalent expressions

$$P(D) > P(E) \tag{A.6}$$

$$\frac{1}{2}P(D) > \frac{1}{2}P(E) \tag{A.7}$$

$$P(D) - \frac{1}{2}P(D) > \frac{1}{2}P(E) \tag{A.8}$$

$$P(D) > \frac{1}{2}(P(E) + P(D)) \tag{A.9}$$

$$\frac{P(D)}{P(D) + P(E)} > \frac{1}{2} \tag{A.10}$$

$$P(A|B) > \frac{1}{2} \tag{A.11}$$

If $P(A) = 1$, then in (1 \Rightarrow 2) $P(B) = 0$ is possible, yielding $P(w) \geq P(\bar{A}) = 0$ for $w \in A$. This weak inequality transfers to (3) and (4). (4 \Rightarrow 1) since $P(D) \geq P(E) = 0$ (4). Thus $P(A|B) = \frac{P(D)}{P(D)+P(E)} = 1 > \frac{1}{2}$. \square

Proof 2 of Theorem 1 Assume $\mathcal{A} = \wp(\Omega)$ (finite), \mathbf{B} a belief operator (with core C) and P a probability function over \mathcal{A} .

(1) \Rightarrow (2) Suppose \mathbf{B}, P satisfy the Lockean thesis, for $t = P(C)$.

- i To derive a contradiction, suppose that C is not P -stable. Therefore there is $v \in C$ such that $P(v) \leq P(\overline{C}) = 1 - P(C)$ (Lemma 5.2 or even $P(v) < P(\overline{C}) = 0$ if $P(C) = 1$ creating an immediate contradiction). Consider $B = \overline{C} \cup (C \setminus \{v\})$ which is not a superset of C . Therefore it is not believed. Yet, $P(B) = 1 - P(v) \geq P(C) = t$, contradicting the Lockean thesis.
- ii To derive a contradiction, suppose $P(C) = 1$, and that there is $D \subsetneq C$ with $P(D) = 1$. Then the Lockean thesis implies $\mathbf{B}(D)$, therefore C is not the core.

(2) \Rightarrow (1). Suppose (i) C is P -stable and (ii) if $P(C) = 1$ then C is the smallest probability-1 set. Then $t := P(C) = P(C|\Omega) > \frac{1}{2}$ (since $C \neq \emptyset$). Let us show that (i,ii) imply the Lockean thesis. $C \subseteq A$ implies $P(A) \geq t$ (sub-additivity). Therefore $\mathbf{B}(A)$ implies $P(A) \geq t$. Let us show the reverse.

Suppose there is B such that $\neg\mathbf{B}(B)$ and $P(B) \geq t$ (contradicting the Lockean thesis). Then $C \not\subseteq B$. There are three cases: (1) $B \subsetneq C$, (2) $B \cap C = \emptyset$ or (3) $\emptyset \neq B \cap C \notin \{C, B\}$. All lead to a contradiction. If (1) $B \subsetneq C$ then, by sub-additivity $P(B) \leq P(C) = t$. $P(B) < P(C)$ would contradict $P(B) \geq t$. $P(B) = P(C)$ would contradict $\neg\mathbf{B}(B)$. (2) is excluded because if $P(B) \geq t > \frac{1}{2}$, $P(C) = t$ and $B \cap C = \emptyset$, then $P(B \cup C) \geq 2t > 1$. Consider (3). Let $D = C \cap B \subseteq C$ and $E = B \setminus D \subseteq \overline{C}$. By construction and hypothesis $P(B) = P(D) + P(E) \geq t = P(C)$. Define $B' = D \cup \overline{C}$. $B' \supseteq B$, therefore $P(B') \geq P(B)$. Take a superset D' of D , such that $D' = C \setminus \{w\}$ with $w \in C$ (which exists because $D = B \cap C \neq C$ by (3)). Define $B'' = D' \cup \overline{C}$, therefore $B' \subseteq B''$. Then $P(C) = t \leq P(B) \leq P(B') \leq P(B'') = 1 - P(w)$. Which implies $P(w) \leq P(\overline{C})$. This contradicts (i) P -stability of C (Lemma 5.2) if $P(C) \neq 1$. If $P(C) = 1$, then $P(w) = P(\overline{C}) = 0$ could be possible, but then C is not the smallest probability 1 set, contradicting (ii). \square

Proof 3 of Lemma 3 Let $a \in (0, 1)$.

1. Let $\kappa(w)$ be a ranking mass. Then $P_\kappa^a(w)$ is a probability mass: Its sum is 1 by the normalisation constant, and all $P_\kappa^a(w) \geq 0$. Since $0 \leq a^{\kappa(w)}$ for $a \in (0, 1)$.
2. Let $P(w)$ be a probability mass. Then $\kappa_p^a(w)$ is a ranking mass: Let $p = \max_{v \in \Omega} P(v)$ (which exists, since Ω is finite). Set $r_w = P(w)/p$ for $w \in \Omega$. Then $r_w \in [0, 1]$. Thus $\ln r_w \in [-\infty, 0]$. Yet, for $a \in (0, 1)$, we have $\ln a \in (-\infty, 0)$. Therefore $\kappa_p^a(w) = \log_a r_w = \frac{\ln r_w}{\ln a} \geq 0$. κ_p^a has a zero for $w \in \Omega$ such that $P(w) = p$. Since then $r_w = 1$ and $\ln r_w = \ln 1 = 0$.
3. Because $\Psi_a \Phi_a(\kappa) = \kappa$ and $\Phi_a \Psi_a(P) = P$.

$$\kappa_{P_\kappa^a}^a(w) = \log_a \frac{P_\kappa^a(w)}{\max_{\Omega} P_\kappa^a(v)} = \log_a \frac{a^{\kappa(w)}}{\max_{\Omega} a^{\kappa(v)}} = \log_a \frac{a^{\kappa(w)}}{a^0} = \kappa(w) \tag{A.12}$$

$$P_{\kappa_p^a}^a(w) = \frac{a^{\kappa_p^a(w)}}{Z} = \frac{a^{\log_a \frac{P(w)}{\max_{\Omega} P(v)}}}{Z} = \frac{P(w)}{\max_{\Omega} P(v) \sum_{u \in \Omega} \frac{P(u)}{\max_{\Omega} P(v)}} = P(w) \tag{A.13}$$

4. For $a \in (0, 1)$, a^x and $\log_a y$ are strictly decreasing. \square

A.2 Belief Models

Proof 4 of Theorem 2 Let $a \in (0, 1)$.

1. Let (\mathbf{B}, κ) be a Spohn belief. Since κ is a ranking function, $P = P_K^a$ is a probability function for any $a \in (0, 1)$ (Lemma 3.1). It therefore suffices to show that $[\geq y]_P$ is the core of \mathbf{B} for $y := a^z/Z(a)$. Note that $\kappa(w) \leq z$ iff $P(w) \geq y$ (Lemma 3.4). Therefore for all $A \in \mathcal{A}$:

$$A \in \mathcal{B} \text{ iff } \kappa(\bar{A}) > z$$

$$\text{iff } [\leq z]_\kappa \subseteq A$$

$$\text{iff } [\geq y]_P \subseteq A$$

Therefore $\mathcal{B} := \{A \in \mathcal{A} : \mathbf{B}(A)\}$ is generated by $[\geq y]_P$ (non-empty). Thus \mathcal{B} is a principal filter. Yet, a principal filter is generated by a unique element (its core). Therefore $[\leq y]_P$ is the core.

2. By the fact that $\Psi_a = \Phi_a^{-1}$ (Lemma 3.3). □

Proof 5 of Theorem 3 Assume (\mathbf{B}, P) is an atomic probabilistic belief model over a finite algebra $\mathcal{A} = \wp(\Omega)$. Therefore P is a probability and \mathbf{B} is a belief operator over \mathcal{A} . Additionally \mathbf{B} has a unique core, namely $C = [\geq y]_P$ for some $y \in [0, 1]$.

- (\Rightarrow) Suppose (\mathbf{B}, P) is a Leitgeb model. In particular, (\mathbf{B}, P) satisfy the Lockean thesis, for $t = P(C)$ with C the core. Therefore, by Theorem 1 ($1 \Rightarrow 2$), (i) C is P -stable and (ii) if $P(C) = 1$ then C is the smallest probability-1 set. Set $y = \min_C P(w)$. Then $y > P(\bar{C})$ by Lemma 5.2 and P -stability of C , if $P(C) \neq 1$. If $P(C) = 1$, then C is the smallest probability 1 set and therefore $y > 0 = P(\bar{C})$.
- (\Leftarrow) Suppose that the core C of the atomic belief model (\mathbf{B}, P) satisfies $y > P(\bar{C})$, for y the atomic threshold. It suffices to show that the Lockean thesis is satisfied for $t_C = P(C)$. Let $t_C = P(C)$. Then it suffices to show that $q < t_C$ for $q = \max\{P(A) : \neg\mathbf{B}(A)\}$. Because we may then always chose t with $q < t < t_C$ such that the Lockean thesis is satisfied. By finiteness q exists. A non-believed element with maximal probability will have the form $A = \bar{C} \cup C \setminus \{w\}$, where $w \in C$ has minimal probability in C . Yet, by the separation assumption $P(w) \geq y > P(\bar{C})$. This implies $0 > P(\bar{C}) - P(w)$ which implies $P(C) > P(\bar{C}) + P(C) - P(w)$, which implies $t_C = P(C) > P(A) = q$. □

Proof 6 of Corollary 1 Let $a \in (0, 1)$.

1. In a Leitgeb belief model, \mathbf{B} is a belief operator. \mathbf{B} is closed under conjunction by assumption and \mathbf{B} satisfies (B1-3) by the Lockean thesis. Additionally, the core C is P -stable and if $P(C) = 1$, then it is the smallest probability 1 set. Therefore $\min_C P(w) > P(\bar{C})$. Therefore $y = \min_C P(w)$ can be chosen as

atomic threshold. Therefore a Leitgeb belief model is always an atomic belief model. Thus the claim follows from Theorems 2.2 and 3.

2. Let (\mathbf{B}, κ) be a Spohn belief model with core C and $z = \max_C \kappa(w)$ and $x = \min_{\overline{C}} \kappa(w)$. Let $a \in (0, 1)$. Then $(\mathbf{B}, P_{\kappa}^a)$ is an atomic belief model, with

$$y(a, z) = \frac{a^z}{Z(a)} \quad (\text{A.14})$$

by Theorem 2.1. This model is a Leitgeb model iff $y(a, z) > P(\overline{C})$ (Theorem 3). This holds iff $a^z > Z(\overline{C}) := \sum_{\overline{C}} a^{\kappa(w)}$. It therefore suffices to show this later inequality. Assume

$$a > |\overline{C}|^{\frac{1}{z-x}} = \left(e^{\ln |\overline{C}|} \right)^{1/(z-x)} = e^{\frac{\ln |\overline{C}|}{z-x}} \quad (\text{A.15})$$

The following conditions are equivalent:

$$\ln a > \frac{\ln |\overline{C}|}{(z-x)} \quad (\text{A.16})$$

$$(z-x) \ln a > \ln |\overline{C}| \quad (\text{A.17})$$

$$a^{z-x} > |\overline{C}| \quad (\text{A.18})$$

$$a^z > |\overline{C}| a^x = \sum_{\overline{C}} a^x \quad (\text{A.19})$$

Therefore $a^z > \sum_{\overline{C}} a^x \geq Z(\overline{C})$, since $x = \min_{\overline{C}} \kappa(w)$. \square

References

- Dubois, D., & Prade, H. (1988). *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York: Plenum Press.
- Greene, B. (2012). *The Hidden Reality: Parallel Universes and the Deep Laws of the Cosmos*. London: Penguin Books.
- Leitgeb, H. (2013). Reducing Belief Simpliciter to Degrees of Belief. *Annals of Pure and Applied Logic*, 164(12), 1338–1389. doi:10.1016/j.apal.2013.06.015.
- Leitgeb, H. (2014). The Stability Theory of Belief. *The Philosophical Review*, 123(2), 131–171. doi:10.1215/00318108-2400575.
- Leitgeb, H. (2015). The Humean Thesis on Belief. *Aristotelian Society Supplementary Volume*, 89(1), 143–185. doi:10.1111/j.1467-8349.2015.00248.x.
- Levi, I. (1996). *For the Sake of the Argument: Ramsey Test Conditionals, Inductive Inference and Non-monotonic Reasoning*. Cambridge: Cambridge University Press.
- Lin, H., & Kelly, K.T. (2012). Propositional Reasoning that Tracks Probabilistic Reasoning. *Journal of Philosophical Logic*, 41, 957–981. doi:10.1007/s10992-012-9237-3.
- Raidl, E. (2014). *Probabilité, Invariance et Objectivité*, PhD thesis at the University of Paris 1 Panthéon-Sorbonne, IHPST. <http://www.theses.fr/s123429>.
- Rott, H. (2004). Stability, strength and sensitivity: Converting belief into knowledge. *Erkenntnis*, 61(2-3), 469–93. doi:10.1007/s10670-004-9287-1.
- Rott, H. (2009). Degrees all the way down: Beliefs, non-beliefs and disbeliefs, In Huber, F., & Schmidt-Petri, C. (Eds.) *Degrees of Belief* (pp. 301–339). Dordrecht: Springer.
- Rott, H. (2015a). Stability and scepticism in the generation of plain beliefs from probabilities. Manuscript version of May 26, 2015.
- Rott, H. (2015b). Unstable knowledge, unstable belief. Manuscript version of July 28, 2015.

13. Skovgaard-Olsen, N. (2015). The problem of logical Omniscience, the preface paradox, and doxastic commitments. *Synthese*, 1–26. doi:10.1007/s11229-015-0979-7.
14. Spohn, W. (1988). Ordinal Conditional Functions. A Dynamic Theory of Epistemic States, In Harper, W.L., & Skyrms, B. (Eds.) *Causation in Decision, Belief Change, and Statistics*, Vol. 2 (pp. 105–134). Dordrecht: Kluwer.
15. Spohn, W. (2012). *The Laws of Belief: Ranking Theory and its Philosophical Applications*. Oxford: Oxford University Press.
16. Spohn, W. (presentation). The Value of Knowledge. https://www.tilburguniversity.edu/upload/87af9554-bf5f-4514-be46-021183a63bf0_Presentation%20Spohn.pdf. Accessed 3 March 2015.
17. Yalcin, S. (2011). Nonfactualism about Epistemic Modality, In Egan, A., Weatherson, B., & Yalcin, S. (Eds.) (pp. 295–332): Oxford University Press.
18. Yalcin, S. (forthcoming). Belief as Question-Sensitive. *Philosophy and Phenomenological Research*. https://www.academia.edu/26580337/Belief_as_Question-Sensitive. Accessed 19 August 2016.