The Interaction of Logic and Probability in Accuracy Based Epistemic Utility Arguments

Mathijs Mul
Science Major

Bachelor thesis

Amsterdam University College

Supervisor
T. Achourioti

Reader
R. Pettigrew

May 15th, 2013

Abstract

This paper investigates the relation between logic and probability theory by studying its instantiation in epistemic utility theory. We first provide a general analysis of logic and probability, and identify the issues that may arise in systems that combine both fields. Next, we consider such issues in the context of epistemic utility, which offers a mathematical formalisation of epistemic states and thereby seeks to establish norms for rational beliefs. We approach epistemic utility from the accuracy based justification of probabilism by Leitgeb and Pettigrew, and show how the interactions between logical truth and probabilistic belief in this argument are unproblematic because belief is interpreted as expected truth value, and because the geometric framework is taken to be Euclidean, which assumes that the logical semantics are classical. Dropping this assumption, we continue by investigating what happens to the argument if we consider non-classical semantics as well. Using Williams' proposed generalisation of gradational accuracy, we run the Leitgeb and Pettigrew argument for alternative semantics. This establishes probabilism for several semantics, given that we accept Williams' presuppositions. We argue that these assumptions are too strong to guarantee the preservation of vital semantic characteristics, which we illustrate in a case study for the logic of paradox (LP). We claim that this non-classical logic is relevant to epistemic utility, and that probabilism is still an appropriate norm for rational beliefs. Therefore, we suggest three alternative methods of justifying probabilism for LP, and assess the logical-probabilistic dynamics in which they result. We observe how the maintenance of a Euclidean geometry facilitates a generalisation of gradational accuracy, which allows for unproblematic interactions between truth and probability, provided that we still interpret beliefs as expected truth values. We explain how the case study for LP relates to other non-classical logics, and conclude that an unproblematic relation between truth and probability in epistemic utility generally requires (a) preservation of the Euclidean geometry and (b) adherence to the mentioned interpretation of belief. Returning to the general relation between logic and probability and the class of systems that combine both fields, we conclude that the analysis of epistemic utility demonstrates how unproblematic logical-probabilistic interactions necessitate an adequate geometric understanding of logical truth and probability, and a coherent semantic interpretation of both notions. This conclusion can form the starting point for similarly themed studies into other systems that combine logic and probability theory.

Acknowledgements

I would like to express the greatest gratitude to Dora Achourioti, who supervised me while I was writing this paper. Her interest in the project, continuous feedback and active support were a great source of motivation, inspiration and knowledge, without which this thesis would not have been possible. I would also like to thank Richard Pettigrew for being my second reader, as well as for the useful comments that he made on the draft version of the paper.

Contents

1 Introduction

Logic and probability theory have long been separated disciplines. It is not difficult to develop an intuitive understanding of the distinction between both fields, and the points at which they may differ or even conflict. Yet, such disciplinary contrasts are by no means a reason to keep both fields apart - indeed, the divergence between their respective idiom and domain has created a promising condition for their reconciliation, and attempts have been made to combine the particular strengths of both discourses. Think, for example, of projects such as the logical formalisation of probability theory, or the incorporation of probabilistic elements in logical systems.

However, compromises between logic and probability do not come naturally. The differences between both fields urge us to think about the interactions between logic and probability once integrated in a new system, and the issues that such interactions are likely to raise. In this paper, we will therefore set out to investigate the points where logic and probability theory disagree, and identify the problems that such discrepancies may cause if we choose to combine logic and probability in the same system. In order to develop a concrete awareness of such problems, we will study their manifestation in the particular field of epistemic utility theory. In this context, we will not only explore the relation between probability and classical logic, but also consider the influence of applying non-classical semantics. To create a clear sense of what may happen, we will take the logic of paradox (LP) as a case study, approach it from an epistemic utility perspective and assess the dynamics that we encounter between logic and probability.

The first section of the paper will be dedicated to a general analysis of logic and probability theory. We will explain how both fields are understood in this paper, how they relate, where they differ, and what kind of issues we may expect in systems of combined logic and probability.

Subsequently, in the second section we will examine how the identified issues are instantiated in a specific case, namely that of epistemic utility theory. This theory offers a mathematical formalisation of epistemic states, and thereby seeks to establish norms for rational belief. After a general explanation of the project and its main results, we will study how it concretises the abstract relationship between logic and probability. We will explore which particular challenges are presented, and how these could be faced.

The reason for choosing epistemic utility is that it bears great relevance to contemporary logical and philosophical debates, and that establishing probabilism for rational credences has in fact been one of its primary aims. Hence, the project has generated sophisticated arguments to justify the rationality of probabilism. Such arguments simultaneously legitimise the position of probabilistic elements, namely credences, in a logical framework. This creates an interesting and concrete setting to study the dynamics between logic and probability.

Our treatment of epistemic utility will be centred around the accuracy-based argument in favour of probabilism advanced by Leitgeb and Pettigrew in [LP10a] and [LP10b], and the notion of accuracy as such. We will see how accuracy is instrumental to establishing probabilism, and how it is at the same time the locus that most strongly and urgently draws our attention to the interactions between logic and probability in this context. We will investigate the shape of these dynamics, and study the assumptions on which the legitimacy of the applied accuracy concept is based.

As we will see, the main presupposition underlying the relation between logical truth and probabilistic credence in Leitgeb and Pettigrew's argument is that the applied logical semantics are classical. However, we will argue that non-classical logics are relevant to the epistemic utility project as well, in the sense that we would also like to establish probabilism for rational credences under less traditional circumstances. Therefore, we will investigate how the Leitgeb and Pettigrew argument is affected if we drop the assumption of classicality, and turn to alternative logics. We will investigate a possible generalisation of the accuracy based argument in order to justify probabilism for non-classical semantics, and reassess the relation between truth and probability under such conditions. Not only is this of interest to our study into the relation between logic and probability, as hopefully it will also result in a contribution to the epistemic utility discourse.

Investigating the generalisation of Leitgeb and Pettigrew's argument, we will first consider the applicability of a method proposed by Williams to generalise an earlier epistemic utility argument, namely Joyce's. Williams describes a generalised accuracy notion in [Wil12], which we will demonstrate can be integrated in Leitgeb and Pettigrew's argument in order to extend its conclusion to several non-classical logics as well. We will, however, continue by arguing that Williams' method appeals to strong assumptions, which may have undesirable consequences for characteristic aspects of logical semantics.

The third and last section of the paper will comprise a case study, where we elaborate on the specific relevance of LP to epistemic utility, and demonstrate how Williams' assumptions damage the distinguishing features of this logic's semantics. As we would still like to establish probabilism and evaluate the interactions between probabilistic credence and LP semantics, we will suggest a few alternative justifications. Using a method different from Williams' and paying particular attention to accuracy, we will attempt to translate the original Leitgeb and Pettigrew argument to LP and explain what this implies for the relation between logic and probability. Here, we will have to rethink the issues identified in the first section. In the suggested solutions, we will be careful to maintain those characteristics of LP that determine its relevance to epistemic utility. The suggestions will serve as a detailed, technical illustration of what may happen if we let logic and probability interact under non-classical circumstances, and additionally it may prove useful to epistemic utility as such.

Finally, we will conclude our study on the relation between logic and probability by revisiting our initial, general concerns on the basis of our analysis of probabilism arguments in epistemic utility, and particularly those suggested for LP. We will explain how epistemic utility deals with the issues that the dynamics between logic and probability may raise, and describe what happens if we move beyond classicality. By contextualising the results of our case study for LP, we will reflect on the shape that these dynamics may take for other non-classical logics. Moreover, we want to determine how our conclusions on epistemic utility theory could relate to other theories of combined logic and probability, and to which extent we would be justified in extrapolating our findings to such alternative theories. Hereby, we will aim to establish a clearer understanding of what it means for logic and probability to interact in a single system, and what it takes for such interactions to be unproblematic. With this conclusion, we intend to provide a basis for the investigation of other systems that combine logic and probability.

2 Logic and probability

In this section, we will explore the relation between logic and probability theory in general terms. First, we will specify how logic and probability are understood in this paper. Next, we will analyse the exact respects in which both disciplines differ. Finally, we will elaborate on the issues that may arise in projects that combine logic and probability, and critically assess the interactions that such projects are likely to provoke.

2.1 Logic and probability as understood in this paper

Logic will be understood in the widest, classical and non-classical sense. In our discussion of epistemic utility theory, we will first be interested in the classical logic that typically forms its background. Consequently, we will widen our focus and also consider non-classical logics. Chapter 4 will be dedicated to a case study of LP in relation to epistemic utility theory.

Classical probability theory is understood according to the conventional Kolmogorov axiomatisation. That is, for Ω a non-empty set, $\mathbf{F}$ a field on Ω and P a function from $\mathbf{F}$ to $\mathbb{R}$, P is a probability function, and $(\Omega, \mathbf{F}, P)$ a probability space, if and only if the following conditions are satisfied:

(K1) non-negativity: $P(A) \geq 0$, for all $A \in \mathbf{F}$

(K2) normalisation: $P(\Omega) = 1$

(K3) finite additivity: $P(A \cup B) = P(A) + P(B)$ for all $A, B \in \mathbf{F}$ such that $A \cap B = \emptyset$

As we will be investigating interactions between logic and probability, we would like to restate the measure-theoretic Kolmogorov axiomatisation in logical terms. That is, we would like to capture the axioms by means of a logical rather than a mathematical apparatus. The role of the logical consequence relation is essential here, and allows for the following reformulation of Kolmogorov's axioms, as stated by Paris in [Par01] and advocated by Williams in [Wil12].

(P1) $\vdash_k A \Rightarrow b(A) = 1$
$A \vdash_k \Rightarrow b(A) = 0$

(P2) $A \vdash_k B \Rightarrow b(A) \leq b(B)$

(P3) $b(A \wedge B) + b(A \vee B) = b(A) + b(B)$

where b is a function into $[0, 1]$[1] and $\vdash_k$ is interpreted as classical consequence. This axiomatisation is useful not only because it uses a logical idiom, but also because its treatment of the logical consequence relation turns out to be applicable to a range of non-classical logics as well. We are interested in such a generalisation of the logical probability axioms, because upon moving to non-classical situations, we have to reconsider whether the Kolmogorov axiomatisation is still appropriate. Its appeal to sets of possible worlds $A, B \in \mathbf{F}$ may not be compatible with all non-classical logics, so rather than possible worlds, we would like our probability axioms to depend on a notion that allows for a smoother generalisation to alternative logics. Paris has shown in [Par01] that the consequence relation is such a notion, and that (P1)-(P3) are applicable to all semantics satisfying the following conditions:

[1] We now introduce probabilistically constrained functions b in general terms, but later in the paper we will specificy our interest to credence functions, where satisfaction of the probability axioms can be regarded as the coherence requirement of probabilism.

(i) truth values are taken from $\{0, 1\}$

(ii) $A \vdash_k B$ obeys the 'no drop' restriction

(iii) the following Tarskian truth conditions are satisfied:

(T2) $V(A) = 1 \wedge V(B) = 1 \Leftrightarrow V(A \wedge B) = 1$

(T3) $V(A) = 0 \wedge V(B) = 0 \Leftrightarrow V(A \vee B) = 0$

The 'no drop' restriction amounts to the 'guarantee that there is no drop in truth-value over a valid argument' ([Wil12]): the truth value of a proposition should be equal to or lower than the truth value of any of its logical consequences.

2.2 Differences between logic and probability

2.2.1 Compositionality

Compositionality is to say that '[t]he meaning of a complex expression is determined by its structure and the meanings of its constituents.' ([Sza10]) Traditionally, logics are compositional. Think of classical logic, where we can compose truth tables that directly derive the truth values of complex sentences from the truth values of their atomics, which are structured according to certain configurations of connectives. Many non-classical logics, such as LP, are compositional as well.

Probability, however, is not compositional, as the probability of a complex expression is not necessarily determined by the value assigned to its components. E.g., if $0 < P(A) = P(B) < 1$ and A and B are independent, then $P(A \cap B) = P(A)P(B)$, but $P(A \cap A) = P(A) \neq P(A \cap B)$. We see that the probability function P outputs different probabilities for the identically structured arguments $A \cap B$ and $A \cap A$, even though the probabilities of the constituent parts A and B are the same.

2.2.2 Necessity vs. possibility

Mathematically and theoretically, probability functions constitute a particular class of functions. Only philosophical interpretations of the formal constraints on such functions determine their practical use as devices for computations about likelihoods, chances, relative frequencies or degrees of belief. Probability theory is primarily applied in this domain of uncertainties: it does not establish what is necessary, but what is possible.

Logic, on the contrary, is traditionally concerned with statements that can deductively be proven certain, or true. Logical truth does require the analytic necessity that probability typically fails to achieve. At the same time, logic classically does not express the uncertainties of probability theory, as it can only express either truth or falsity. Non-classical systems such as fuzzy logic also tend to emphasise that logical truth may be partial, but that that an interpretation of truth in degrees is different from probabilistic uncertainty.

2.2.3 Discreteness vs. continuity

Classical logic is bivalent: its truth values are limited to the Boolean domain 0, 1. Thus, it can only assign one of these two values to formulas, which makes its semantics discrete. Anything else than 0 or 1 is meaningless. Although bivalence does not apply to all logics, only few (fuzzy) logics locate their truth values in a continuous interval, so semantic discreteness seems typical to logic.

Probabilities, however, can take any possible value in the interval $[0, 1]$, which is a continuous spectrum, as opposed to the discrete set of logical truth values.

2.2.4 Semantic status

Logical semantics revolve around truth, which is usually granted a more or less primitive status - a semantically non-reducible or non-analysable status, that is. (Reductionist approaches to truth are usually interested in reduction to syntactic or set-theoretic terms; not to more primitive semantic terms.)

Probability does not enjoy this semantically primitive status, but has been subjected to a variety of different philosophical interpretations. Some have considered it to express evidential support, others have identified it with subjective degree of belief, and others again with relative frequency. Semantic disagreement appears to be much stronger in probability theory than it is in logic.

2.3 Combining logic and probability

The above-mentioned contrasts have given us an impression of the different domains where logic and probability theory are traditionally applied. We see that a combined logical-probabilistic approach would be particularly useful in situations that deal with necessity as well as with uncertainty. It is for this reason that logic and probability have been implemented together in projects such as the modelling of inductive reasoning and knowledge revision or belief updates. In order to determine which difficulties we can expect in such combined systems, let us look at the interactions that are likely to occur between logic and probability.

Syntactically, the fields seem to be at peace with each other. The syntax of probability theory is of a set-theoretic nature that is easily accommodated in a logical framework. It is in the semantics that we encounter difficulties. Considering probabilities from the classical Kolmogorov perspective, we could model them in a possible-world setting. Probability functions could then be regarded as quantifications of the accessibility relation, which attach an uninterpreted number between 0 and 1 to the relationship between a possible world and the actual world. Thus, probabilities are measures over the set of possible worlds, and they express a certain, unspecified meaning with respect to these worlds. (Be it evidential support, degree of belief, relative frequency or any other interpretation.) The Paris axiomatisation does not assume a possible worlds framework, but still, probability functions assign some form of quantified signification to propositions. Probability can therefore be treated as a semantic notion. At the same time, however, truth continues to be the primary entity in logical semantics. Logic and probability may thus appeal to conflicting forms of semantic authority, so if we combine the two, we should be careful to investigate this duality.

Furthermore, we may have to rethink the status of logical truth in the first place. If we choose to incorporate a probabilistic calculus in a system, we should consider whether it is even possible for logical truth to maintain its traditional status at all. Especially if we integrate analytic statements (with logical semantics) and synthetic ones (with probabilistic semantics) in the same logic, we may want to adapt the classical understanding of truth, and perhaps consider the recognition of partial truths.

Additionally, even if the issues of semantic status have been resolved, we need to be aware of the different mechanics underlying probabilities on the one hand and truth values on the other hand. As we saw, the former are not compositional, which implies a relationship between complex formulas and their constituent components that differs

strongly from the one we encounter in classical logic, as well as in many compositional non-classical logics.

Finally, a philosophical concern to keep in mind is that the assignment of a logical truth value to a statement that also has a probabilistic valuation may imply semantic realism. Imagine, for instance, that we interpret probabilities as degrees of belief, which epistemic agents hold with respect to propositions. If, by independent logical semantics, these propositions are also assigned a truth value, it is possible for formulas to be true yet unknown. This entails semantic realism.

3 Epistemic utility theory

3.1 The basics

Epistemic utility theory is a relatively new field of research, which aims to justify the norms that govern agents' degrees of credences on purely epistemic grounds. In this it differs from other projects concerned with norms of rational credence, which have often appealed to pragmatic considerations (cf. the Dutch Book argument; [Fin31]). Among the authors who have been most active in its recent developments are Pettigrew, Leitgeb and Joyce.

The epistemic norm in which we will be most interested, is probabilism. This is to say that our credences should satisfy the axioms of probability theory (i.e.: non-negativity, normalization and finite additivity; see section 2.1). Key to epistemic utility is that it treats epistemic states as epistemic acts, which allow reasoning about epistemic scenarios by means of standard utility theory. The following functions, sets and variables are defined:

- A: a proposition (set of possible worlds)
- W: the (finite) set of possible worlds about which an agent has an opinion, with its power set P(W) the set of all propositions about which the agent has an opinion
- $b(A)$: an agent's credence function, expressing the agent's degree of credence in A, with domain $P(W)$ and codomain $[0, 1]$
- $\chi_A(w)$: the truth value of a proposition A at a world w
- **B**: the set of possible credence functions on $P(W)$: all functions mapping $P(W)$ to $[0, 1]$
- **P**: the set of credence functions satisfying the axioms of finitely additive probability (Pettigrew 2011, 1-2)

So, in order to establish probabilism, it must be shown that b is in **P**. Several arguments have been provided in order to prove that this is the case. This paper will largely be based on one of the most recent and noteworthy such arguments, namely the one formulated by Leitgeb and Pettigrew in 2010. The argument will be discussed in more detail in the next section.

Leitgeb and Pettigrew's argument, as well as several other important arguments in epistemic utility, relies heavily on the notion of *accuracy*. Accuracy expresses the closeness of an agent's degree of belief to the truth value of the propositions with respect to which she holds such beliefs. Nowhere in epistemic utility theory do credences and logical semantics interact more intimately than in the concept of accuracy.

As this paper will be grounded in the work that has already been done by Leitgeb and Pettigrew, we assume that probabilism has been established, and that we may therefore consider rational credences to be probabilistic under the premises of their argument. Thus, accuracy becomes a highly significant notion also in the context of this paper, because we may treat credences as probabilities, so that the interaction between credences and truth values becomes an interaction between probabilities and truth values. Because it is this interaction that we would like to study, the next section will discuss Leitgeb and Pettigrew's argument with a particular focus on the importance of accuracy.

3.2 The role of accuracy in epistemic utility: Leitgeb and Pettigrew's argument

Before we can investigate how accuracy is instrumental to establishing probabilism in Leitgeb and Pettigrew's argument, we should first examine the technical definition of accuracy to which they appeal.

In order to prevent confusion, it should first be pointed out that accuracy can denote two different things, namely:

1. the *norm* that an epistemic agent ought to approximate the truth - that she ought to minimise her inaccuracy ([LP10a], 202)

2. the descriptive notion of accuracy on which the norm 1 is based

We will see that 1 and 2 are very much intertwined: there are different ways of defining accuracy measures, all of which give rise to different norms.

As Leitgeb and Pettigrew point out, what is relevant is what Joyce ([Joy98], [Joy09]) called 'gradational accuracy', which 'depends only on the truth values of propositions at worlds and on the agent's belief function.'[2] (203) Thus, accuracy must define some (quantitative) relationship between truth and credence. As many such relationships are possible, it is useful to first distinguish between a local and a global approach, which leads to the following concepts:

- local inaccuracy measure: a function $I(A, w, x)$ that measures the distance of a credence x in a proposition A from the truth value $\chi_A(w)$ of A at world w.

- global inaccuracy measure: a function $G(w, b)$ that measures the distance between a credence function b, consisting of particular degrees of belief with respect to individual propositions A, and a world , consisting of particular truth values of individual propositions A. b and w are represented by vectors.[3]

From these different inaccuracy measures, we can derive corresponding norms, which aim at minimising their particular kinds of inaccuracy.

The mentioned local and global inaccuracy measures do not yet specify the particular worlds with respect to which they should be calculated. Leitgeb and Pettigrew argue that from the internalist perspective that they adopt, this cannot be the actual world, because an agent may not even know which world is the actual one. Instead, the inaccuracy measures should be evaluated with respect to the set E of all epistemically possible worlds in order to find the expected inaccuracy. This yields the following definitions.

Definition 3.1. Expected local inaccuracy. For a local inaccuracy measure I, a credence function b, a credence x, and propositions $A, E \subseteq W$, the expected local inaccuracy of x in A by the lights of b, with respect to I, and over the set E of epistemically possible worlds is defined by

$$\mathrm{LExp}_b(I, A, E, x) = \sum_{w \in E} b(\{w\}) I(A, w, x)$$

Definition 3.2. Expected global inaccuracy. For a global inaccuracy measure G, credence function b and b', and a propositions $E \subseteq W$, the expected global inaccuracy of b' by the lights of b, with respect to G, and over the set E of epistemically possible worlds is defined by

$$\mathrm{GExp}_b(G, E, b') = \sum_{w \in E} b(\{w\}) G(w, b')$$

[2] This understanding is somewhat reminiscent of Popper's idea of verisimilitude, but not quite the same. Indeed, the inadequacy of verisimilitude for the epistemic utility project follows from earlier papers, e.g. [Mil74]

[3] It would also have been possible to design accuracy measures, rather than inaccuracy measures, but Leitgeb and Pettigrew defend the use of inaccuracy measures on grounds of their convenient geometry. (202-3)

Again, from these measures we can derive corresponding norms: agents ought to minimise the expected local inaccuracy of their degrees of credence in all propositions $A \subseteq W$ relative to a local inaccuracy measure, or minimise the expected global inaccuracy of their credence function relative to a global inaccuracy measure.

Leitgeb and Pettigrew argue that these norms still have to be specified according to the credence function 'by the lights of which we assess expected local or global inaccuracy' (207). This leads to a distinction between synchronic and diachronic accuracy. As we will mainly be concerned with probabilism, which is a synchronic norm, we will only refer to synchronic expected local and synchronic expected global inaccuracy. These norms add to the expected local and global norms that an agent should minimise inaccuracy 'by the lights of her current belief function' (207).

The required norms have now been specified. What still needs to be made precise are which exact inaccuracy measures will be used. These measures have to express the distance between credences and truth values, and it is shown that they have to be *quadratic*.[4] Leitgeb and Pettigrew impose basic conditions on legitimate inaccuracy measures, namely local and global normality and dominance, comparability and minimum inaccuracy. Moreover, they discuss epistemic dilemmas that they wish to avoid by imposing the following additional constraints on the inaccuracy measures: agreement on inaccuracy, separability of global inaccuracy and agreement on directed urgency, together with continuous differentiability. Their Theorems 3-5 show that quadratic inaccuracy measures form the only set of functions that satisfy all these conditions. (219-229)

Leitgeb and Pettigrew elaborate on the assumptions underlying the argument. Most importantly, they mention the particular geometric presuppositions that are made: a Euclidean space is assumed, where credences and truth values have become 'comparable' (212). The assumption of Euclidean space indeed seems to facilitate a very specific treatment of truth values in relation to credence functions. In the set framework, these are geometrically equivalent notions, which enables the calculation of distances on which inaccuracy measures are based.

Now that the relevant accuracy-related norms and definitions have been formulated, we can see how accuracy is instrumental to the justification of probabilism that Leitgeb and Pettigrew provide in their second 2010 article ([LP10b]). While discussing the argument, it may be useful to think of the 3-step procedure that Pettigrew described during the 2012 Round Table on Coherence at the Munich Center of Mathematical Philosophy, see [Pet+12]. This procedure explains how coherence requirements in epistemic reasoning may be justified in three steps:

1. Define a vindicated set of beliefs $\mathbf{B}_v$.
2. Define a notion of distance $d(\mathbf{B}_v, \mathbf{B})$ between $\mathbf{B}_v$ and an agent's set of beliefs $\mathbf{B}$.
3. Adopt a fundamental principle or norm that uses $d(\mathbf{B}_v, \mathbf{B})$ to ground a coherence requirement for $\mathbf{B}$.

C.q., the coherence requirement that we wish to ground is the epistemic norm of probabilism. Leitgeb and Pettigrew's argument follows the structure of this 3-step procedure, but instead of defining a vindicated set of beliefs, the actual truth values of propositions are taken into account. Thus, for $\mathbf{B}_v$ we should read the set of truth values of believed propositions. The distance between these truth values and true beliefs is zero, so numerically, there is no difference between the set

[4] Quadratic inaccuracy measures, which are Brier scores, express local inaccuracy in terms of the squared difference (distance) between credences and truth values, and global inaccuracy in terms of the squared distance between the vectors representing credences and truth values. Distance denotes Euclidean distance. See section 3.3.

of truth values and the set of true (vindicated) beliefs. The notion of distance between an agent's set of beliefs and the set of truth values is expressed by (quadratic) inaccuracy measures, which quantify the distance between credence and truth. As for step 3, the norm that Leitgeb and Pettigrew choose to adopt is the above-mentioned diachronic and synchronic expected local and global accuracy norm.

We are interested in probabilism, and thereby in the synchronic norm. As synchronic expected global accuracy follows from synchronic expected local accuracy (242), we only need synchronic expected local accuracy: 'An agent ought to minimise the expected local inaccuracy of her degrees of credence in all propositions $A \subseteq W$ by the lights of her current belief function, relative to a legitimate local inaccuracy measure and over the set of worlds that are currently epistemically possible for her.' (241)

Leitgeb and Pettigrew then prove the following theorem:

Theorem 1. Suppose b is a credence function, $E \subseteq W$, $\sum_{w \in E} b(\{w\}) \neq 0$ and I is a quadratic local inaccuracy measure. Then the following two propositions are equivalent:

(i) For all $A \subseteq W$ and any $x \in [0, 1]$,

$$\mathrm{LExp}_b(I, A, E, b(A)) = \sum_{w \in E} b(\{w\}) I(A, w, b(A)) \leq \sum_{w \in E} b(\{w\}) I(A, w, x) = \mathrm{LExp}_b(I, A, E, x)$$

(b) Credence function b is a probability function with $b(E) = 1$.

By taking synchronic expected local accuracy as a fundamental principle, (i) of Theorem 1 is satisfied: the expected local inaccuracy of credence function b, with respect to all propositions $A \subseteq W$, by the lights of b, relative to the legitimate local inaccuracy measure I, and over the set of worlds E that are currently epistemically possible for her, is minimal. As (i) is equivalent to (ii) it follows that rational credence functions must be probabilistic.

Reasoning from Pettigrew's 3-step procedure, we see that accuracy plays a crucial role at several stages of the argument. It serves as the distance measure between sets of credences and the set of truth values (which coincides with the set of true beliefs), and it also forms the key component of the fundamental principle in which the argument is grounded. Hence, the entire argument depends on the geometric assumptions that are made in the definition of (in)accuracy, which allow us to express the difference between credences and truth values in terms of Euclidean distance. It is this presupposition that we would like to investigate in more (semantic) detail.

3.3 Accuracy: the geometry

Before expanding on the underlying assumptions, let us first have a closer look at the geometry of accuracy. As inaccuracy measures are defined on a particular framework, we will approach this geometry from the perspective of local and global inaccuracy measures.

Leitgeb and Pettigrew use Figure 1 to illustrate local inaccuracy measures I. As described before, a local inaccuracy measure is a function $I(A, w, x)$ that measures the distance of a credence x in an individual proposition A from the truth value $\chi_A(w)$ of A at world w. Credences range between 0 and 1, and truth values are treated classically, so they can either be 0 or 1. Hence, credences $b(A)$ can be plotted on the real axis together with truth values $\chi_A(w)$. Inaccuracy depends to the one-dimensional (Euclidean) distance between $b(A)$ and $\chi_A(w)$, which is the absolute difference $\mid b(A) - \chi_A(w) \mid$.

We recall that the global inaccuracy measure $G(w, b)$ is a function that measures the distance between a (global) credence function b, consisting of particular degrees of belief with respect to

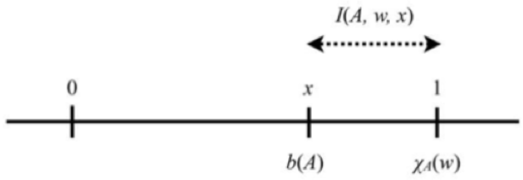

Figure 1: Local inaccuracy, from [LP10b], 204.

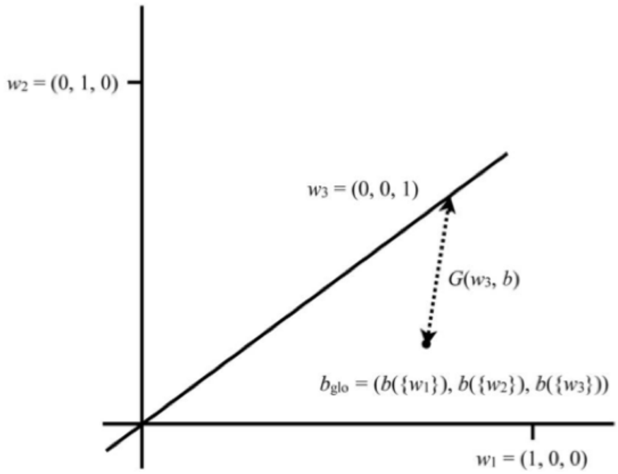

Figure 2: Global inaccuracy, from [LP10b], 205.

propositions $A \subseteq W$, and a world w consisting of particular truth values of these A. G could be considered to generalise I over the entire set of propositions with respect to which an agent has credences b, and which have truth values depending on the world w. As we are no longer dealing with single propositions, but with sets, b and w have to be represented as vectors. b and w are assumed to have the same geometric behaviour, which facilitates a treatment of credences and truth values in a single n-dimensional Euclidean space, where n is the number of propositional atoms (so that the number of propositions is 2^n). For $n = 3$, Leitgeb and Pettigrew provide Figure 2.

Given this geometric framework, global inaccuracy can be expressed as the Euclidean distance between a credence b and a world w. As b and w are n-dimensional vectors, this distance takes the form

$$|| \, b - w \, ||= \sqrt{(b(A_1) - \chi_{A_1}(w))^2 + \cdots + (b(A_n) - \chi_{A_n}(w))^2}$$

In general, we see that inaccuracy measures are capable of expressing the distance between credences and truth values by situating both concepts in the same Euclidean space. This space is one-dimensional if we look at single propositions with the local inaccuracy measure, and n-dimensional if we look at sets of n atoms with the global inaccuracy measure, but the basic presupposition that we are allowed to position credences and truth values in the same geometric framework is the same for all cases. In the next section we will investigate the basis of this presupposition.

3.4 Accuracy: the assumptions

We have seen how the treatment of accuracy in epistemic utility appeals to a Euclidean geometry, where credences and truth values behave analogously. This is a perspective that is justified by underlying assumptions on which we will now elaborate in greater detail. Moreover, we will return to the issues raised in section 2.3, and investigate how epistemic utility manages to overcome these issues by making these assumptions.

Our interest here will really be the fundamental notion of accuracy, rather than functions, such as the quadratic inaccuracy measures, that use accuracy to perform further computations. We approach accuracy from the perspective of local and global inaccuracy measures, as described in the preceding section.

We first want to be more explicit about the assumptions that allow for the described treatment of credences and truth values. Key is that credence functions b and worlds w, be it locally or globally, are treated as having identical geometric behaviour. As Leitgeb and Pettigrew formulate it:

> $(\cdots)$ we will demand that every $(\cdots)$ notion of the geometrical distance between two points supervenes on - is functionally dependent on - the Euclidean distance between these points. In this sense, measuring closeness [between credence and truth] will always amount to a geometrical, and indeed Euclidean, procedure in the context of this article. (211)

The assumption that credences and truth values can be situated in the same Euclidean space, and that the computation of the Euclidean distance between them is a legitimate operation, is justified as follows:

> $(\cdots)$ we regard rational agents to be aiming at distributing their degrees of belief in such a way that every such degree $b(A)$ approximates the truth value of the proposition A $(\cdots)$ Hence, a rational agent's degree of belief for a proposition is nothing but the agent's best possible estimate or 'simulation' of the truth value of that proposition, given her epistemic situation. (211)

Moreover:

> Since truth and falsity have been represented by real numbers, too, degrees of belief and truth values are comparable - they occupy the same quantitative or geometric scale. (212)

The perspective on credences as approximated truth values, together with the coincidence of the quantitative scale of credences (0, 1) and truth values $\{0, 1\}$, semantically justifies how credences and truth values are positioned in the same Euclidean space. Both locally and globally, it legitimises the expression of inaccuracy measures in terms of the Euclidean distance between truth and (global) credence.

We now continue reasoning from the perspective of the accuracy-based argument by Leitgeb and Pettigrew that was discussed in section 3.2. We saw how this argument established classical probabilism. Hence, if we wish to limit ourselves to rational credences, we do no longer have to consider non-probabilistic credence functions. Excluding all non-probabilistic credence functions from consideration effectively turns inaccuracy into a measure of the distance between truth values and probabilities. This is what causes accuracy to draw our attention to the interaction between truth and probability in epistemic utility, and the issues that such interactions may raise, as described in section 2.3. Now that we have discussed the justification of the treatment of truth and credence - and thereby probability - in the same Euclidean framework, we are in a position to assess how epistemic utility deals with these issues.

The first problem that we identified was the potential conflict between logical semantics on the one hand, and probabilistic semantics on the other hand. Leitgeb and Pettigrew have resolved this issue by interpreting (rational, probabilistic) credences as approximated truth values. Thereby, semantic compatibility is established between probabilistic credences and logical truth values.

The second issue, concerning the status of logical truth, is not really problematic in epistemic utility. Logical truth is not marginalised or relativised in any way by the inclusion of probabilistic credences, but maintains its traditional status. Although both concepts occupy the same geometric scale, credences do not substitute truth values. They are defined as functions over the set of all propositions, thus supplementing the truth values of these propositions, which are assigned independently.

The compositionality issue does not seem to be relevant to epistemic utility either. We have credence functions that assign degrees of belief, and characteristic functions that assign truth values to propositions at particular worlds. Credences and truth values are both involved in (in)accuracy computations, but never in such a way that the non-compositionality of probabilistic credences may conflict with the compositionality of the logic.

Finally, the fact that propositions are assigned truth values independent of agents' degrees of belief does not imply semantic realism in the case of epistemic utility. The reason for this is that epistemic utility does not postulate particular characteristic functions, but considers a set of possible worlds, all of which may assign different truth values to the available propositions. Thus, the truth values of propositions are not fixed and 'real', but differ from world to world.

We conclude that epistemic utility tackles the issues raised in section 2.3, and that the interaction between truth and probability is unproblematic, insofar as we adhere to the presuppositions made by the theory. We have seen that the definition of (in)accuracy in terms of Euclidean distance, and the resulting dynamics between credence and truth, are supported by the semantic consideration that credences can be regarded as expected truth values, together with the coincidence of their respective geometric or quantitative scales. This last point is essential: for an appropriate notion of accuracy, and thereby for an unproblematic relation between truth and probabilistic credence, we depend on the quantitative analogy between truth and belief. This analogy is present in epistemic utility as discussed in the above, by virtue of its logical semantics being *classical*.

The logical semantics of epistemic utility as we know it are bivalent - characteristic functions $\chi_A(w)$ only assign the values 0 and 1 - and behave classically. The comfortable fact that the traditional truth values 0 and 1 coincide with the lower and upper bound of the range along which credences can run, facilitates the geometrically identical treatment of truth and credence. It is this restriction to classical semantics that allows us to situate truth and credence in the same Euclidean space, and to perform the accuracy-based argument for probabilism. Moreover, it is on the basis of this assumption that we have been able to conclude that the interactions between truth and probability in epistemic utility are unproblematic.

Constraining the logical semantics of epistemic utility to a classical framework, we see, has permitted the development of an idealised accuracy conception. Credences and truth values can nicely be represented by vectors in the same Euclidean space, and the mathematical operations by which probabilism is established constitute no difficulties. Yet, there are numerous, well-known reasons for considering semantics other than the classical one. What if we want to abandon bivalence, and adopt, for instance, many-valued semantics? Epistemic utility in its current form does not guarantee that we will then be able to develop a notion of accuracy similar to the one proposed for classical semantics, nor do we know whether probabilism can then be established at all. And even if probabilism could still be justified, we would have to reevaluate the interaction between logical truth and probabilistic credence.

3.5 Generalisation to non-classical semantics

We will now investigate what happens, and what can be done, if we drop the assumption of semantic classicality. We will examine whether a generalised Leitgeb and Pettigrew-style justification of probabilism is possible, and how the notion of accuracy has to be adapted to facilitate this potential generalisation to non-classical semantics. If such a generalisation is indeed realisable, we want to investigate if the relation between truth and probabilistic credence can remain unproblematic, even under non-classical circumstances and with the adapted concept of accuracy. This will happen in the next section.

First, let us point out that probabilism of rational credences is something to aspire to, also under non-classical circumstances. Essentially, what probabilism says is that an agent ought not to believe a proposition more strongly than any of its logical consequences. In a rational setting, we understand that this is a norm worth justifying, irrespective of the particular logical semantics we use. Yet, as semantics tend to come with their own notion of logical consequence, we should be careful to investigate how the justification of probabilism in a classical setting could possibly relate to non-classical frameworks.

Given that we wish to establish probabilism in the case of non-classical semantics as well, and that we preferably want to do so by means of a generalised form of the Leitgeb and Pettigrew argument, what we are working towards is a new form of accuracy, which would display new interactions between probabilistic credence and logical truth. These interactions we would then like to assess.

We first consider a project that may be useful to our own, namely Williams' recent generalisation of Joyce's 1998 argument for probabilism ([Joy98]). In his 2012 paper, [Wil12], Williams shows how Joyce's conclusions do not depend on the background assumption of a classical logic and semantics. We should stress that although Joyce's original argument is a controversial one (cf. Maher), its mechanics rely on the same form of gradational accuracy that we saw was central to the Leitgeb and Pettigrew argument. Hence, Williams' focus on the notion of gradational accuracy in his Joyce generalisation can still be useful to the context of this paper.

Williams motivates his generalisation of Joyce's argument by pointing out:

> There are many who argue that this assumption [of classical semantics] fails, in general or for certain subject matters. For example, perhaps cases of presupposition failure ('The King of France is bald') are neither true nor false. Perhaps observational predicates obey an intuitionistic, rather than classical, calculus. Perhaps it's neither true nor false to say that Newtonian mass is relativistic mass. Perhaps the law of excluded middle should be given up in order to get a satisfying take on semantic paradoxes. Perhaps borderline cases of vague predicates require us to think about degrees of truth, or buy into a supervaluational framework. (515)

These are exactly the considerations that also inspire us to reconsider the conclusions of the Leitgeb and Pettigrew argument from a non-classical perspective, such as LP, supervaluationism or intuitionistic logic. Let us see how this works out for Joyce.

Williams argues that Joyce's notion of gradational accuracy does not depend on the use of classical logic, but that it can easily be adapted to other semantics, provided that we carefully specify how to deal with the designated truth values of such semantics. In the classical case, gradational accuracy, as closeness to truth value, is a relatively simple notion. We have the two truth values 0 and 1, with 1 designated, and accuracy demands that an agent's credences approach the designated truth value 1 for true propositions, and the non-designated 0 for false propositions. Without classical bivalence, however, we can no longer rely on the distinction between one designated and one non-designated truth value for our notion of accuracy. In LP, for example, we have two designated truth values, which distorts the geometry that made accuracy such a nice concept in the classical case.

Williams wants to solve this issue by specifying 'ideal aims for credence', or 'truth-values', for each semantics separately. (516) For Kleene gaps, LP gluts, (finite) fuzzy logic, fuzzy gaps, (degree) supervaluationism and intuitionism, he describes how truth 'statuses' (e.g., for LP, 0, 1, both) can be mapped to [0,1] in order to serve as such aims for belief. By resituating the mentioned non-classical semantics in the [0,1] interval, the geometric analogy between credences and truth values that we encountered in the classical setting is reconstructed and generalised to non-classical semantics, which thus obtain their own notion of gradational accuracy.[5]

The projection of non-classical truth statuses onto [0,1] is not altogether uncontroversial, however, as Williams points out:

> $(\cdots)$ even for those who agree on a certain kind of non-classical style of semantics, there may well be controversy over the doxastic role that the projection onto truth-values brings with it. (519)

If we leave the classical framework, it is not always straightforward how truth statuses should give rise to aims for credences. This is the doxastic role to which Williams refers in the above quotation: non-classical semantics do not always specify how their truth values relate to doxastic patterns and objectives. This complicates the design of inaccuracy measures for non-classical semantics. In the case of supervaluationism, for instance, Williams explains that 'it's far from clear what sort of doxastic role, or projection onto truth values,' (520) is intended. The interpretation of intuitionistic logic is also a delicate operation, and Williams' own version is 'deprecated by some prominent intuitionists.' (520)

However, if we agree on a particular mapping from truth statuses to the interval [0,1] for a semantics, we do have an analogue to gradational accuracy as we know it from the classical case. Williams continues by showing that this new notion of accuracy indeed facilitates a generalisation of Joyce's probabilism argument to non-classical semantics. As we are interested in Williams' perspective on accuracy as such, rather than the Joyce argument, we will not elaborate on the precise generalisation. Instead, we will determine whether Williams' treatment of non-classical, gradational accuracy is also applicable to the argument made by Leitgeb and Pettigrew.

We argue that, in fact, Williams' perspective on accuracy can be incorporated in the Leitgeb and Pettigrew argument, and that it can be used to generalise the argument to certain non-classical semantics. We will show this by running the argument for alternative semantics, *assuming* that a Wiliams style mapping from the truth statuses of non-classical semantics into [0,1] is possible.

[5] One could perhaps argue that Williams' mapping does not reconstruct the geometric analogy between credences and truth values, but that it sets up a new framework where we encounter no interactions between credences and truth values, but between credences and *ideal* credences. In this case, it would be somewhat misleading that the projected truth statuses are still called 'truth-values'. Moreover, according to this interpretation, Williams' mapping would reduce the role of truth statuses to a purely doxastic one. Of course, the doxastic function is essential in the given context, but I do not believe that it is Williams' objective to condense the semantics to an exclusively doxastic enterprise. Therefore, I would like to treat the mapped truth statuses as truth values with a doxastic authority, and not as pure credences. This background consideration is something to keep in mind, yet it will have no influence on our later arguments.

As discussed in section 3.2, Letigeb and Pettigrew first prove that the local and global inaccuracy measures have to be quadratic. The actual probabilism argument, which generates Theorem 1, relies on the quadratic nature of the inaccuracy measures. Therefore, we first have to show that non-classical semantics also support the use of such quadratic inaccuracy measures.

In 2.2 we described how the class of legitimate inaccuracy measures is restricted by local normality and dominance, global normality and dominance, local and global comparability and minimum inaccuracy. In order to avoid the discursive dilemma, conflicting inaccuracy values and uncertainty about the (directed) urgency with which agents should update their beliefs, Leitgeb and Pettigrew also impose the following conditions: agreement on inaccuracy, separability of global inaccuracy and agreement on directed urgency, together with the additional constraint of continuous differentiability. Based on these constraints and conditions, and the assumption of classical semantics, it can be shown that the legitimate inaccuracy measures are quadratic.

Given that we have a Williams style mapping for non-classical semantics, we can run the same argument in order to establish the quadratic nature of legitimate inaccuracy measures for general semantics. Let us first revisit all initial constraints, and show that we can impose these on inaccuracy measures for general semantics as well:

- *local normality and dominance*: the Williams mapping resituates truth statuses in the interval [0,1], so that $\chi_A(w)$ only takes values from [0,1] and the Euclidean geometry implied by normality and dominance is maintained.
- *global normality and dominance*: given a mapping from truth statuses to [0,1], we can create a Euclidean n-space for any n propositional atoms, where inaccuracy is interpreted as the Euclidean distance between the vector of all resituated truth values $\chi_A(w) \in [0,1]$ (a world), and the vector of all degrees of belief (a global credence function). Thereby, global normality and dominance are still supported.
- l*ocal and global comparability*: the Williams mapping allows us to generalise the geometric interpretation of inaccuracy as the distance from the truth, where distance is independent of dimension, and expressed by a 'strictly increasing function of the Euclidean metric' (220). Thus, we can still demand local and global comparability to hold.
- *minimum inaccuracy*: also for non-classical semantics, we only want inaccuracy measures that output zero for a zero distance between truth and credence or between a world and a global credence function. The Williams mapping supports this constraint.

The role of the Williams mapping is obvious: by projecting non-classical truth statuses onto [0,1], the same geometry is reconstructed that allowed Leitgeb and Pettigrew's reasoning about inaccuracy measures and probabilism for classical semantics. In a sense, the Leitgeb and Pettigrew argument does not even depend on the logical semantics being classical, but on the quantitative analogy between truth values and credences. Williams' mappings restore this analogy for non-classical semantics.[6]

The notions of expected local and global inaccuracy also generalise. Given a Williams mapping, we saw that the constraints imposed on (classical) inaccuracy measures can also be imposed on non-classical inaccuracy measures. Thus, general inaccuracy measures take the same geometric shape as the classical ones, so the definitions of expected local and global inaccuracy remain applicable.

[6]Of course, using the Williams mappings actually shifts the real difficulty from the epistemic utility argument to the compatibility between non-classical truth statuses and the [0,1] interval, as pointed out before. This is a later concern. We first show that supposing the existence of legitimate Williams mappings allows a rather straightforward generalisation of the Leitgeb and Pettigrew argument.

We now wish to require agreement on inaccuracy, separability of global inaccuracy and agreement on directed urgency, together with the additional constraint of continuous differentiability, on the basis of which we can now prove their Theorems 3, 4 and 5 for general inaccuracy measures. The proof of these theorems is exactly as provided by Leitgeb and Pettigrew, because we have seen how general inaccuracy measures can obey the same constraints as classical ones, and how the definition of expected inaccuracy does not have to be modified. Proving these theorems establishes that general inaccuracy measures have to be quadratic, too.

We are now ready to look at the actual probabilism argument, which centres around the proof of Theorem 1, which we wish to generalise as well. The proof of Theorem 1 depends on Lemma 2, which Leitgeb and Pettigrew state and prove as shown below.

Lemma 2. Suppose $I(A,w,x) = \lambda(\chi_A(w) - x)^2$. Suppose W is finite, b and b' are credence functions, $A, E \subseteq W$, and $\sum_{w \in E} b(\{w\}) \neq 0$. Then the following two propositions are equivalent:

(i) For all $A \subseteq W$ and $x \in [0,1]$,

$$\sum_{w \in E} b(\{w\}) I(A, w, b'(A)) \leq \sum_{w \in E} b(\{w\}) I(A, w, x).$$

(ii) For all $A \subseteq W$,

$$b'(A) = \frac{\sum_{w \in A \cap E} b(\{w\})}{\sum_{w \in E} b(\{w\})}.$$

Proof. By definition, $\sum_{w \in E} b(\{w\}) I(A, w, x) = \sum_{w \in E} b(\{w\}) \lambda(\chi_A(w) - x)^2$, so

$$\begin{aligned} \frac{d}{dx} \sum_{w \in E} b(\{w\}) I(A, w, x) &= 2\lambda(x \sum_{w \in E} b(\{w\}) - \sum_{w \in E} b(\{w\}) \chi_A(w)) \\ &= 0 \end{aligned}$$

if and only if

$$\begin{aligned} x &= \frac{\sum_{w \in E} b(\{w\}) \chi_A(w)}{\sum_{w \in E} b(\{w\})} \\ &= \frac{\sum_{w \in A \cap E} b(\{w\})}{\sum_{w \in E} b(\{w\})} \end{aligned}$$

As $\sum_{w \in E} b(\{w\}) I(A, w, x)$ is a positive quadratic in the variable x, this extremum is a minimum, as required. (263-264)

□

We have just observed how Williams style mappings translate non-classical truth statuses into points in the $[0,1]$ interval, so that the geometric behaviour of non-classical $\chi_a(w)$ is identical to that of their classical counterparts, and the same definition of expected local inaccuracy can be used. We have shown how this enables us to establish the legitimacy of quadratic inaccuracy measures, also for non-classical semantics. Thus, we see how the proof of Lemma 2, which appeals to the quadratic local inaccuracy measure $I(A, w, x)$, generalises to non-classical semantics up until the final identity:

$$x = \frac{\sum_{w \in E} b(\{w\})\chi_A(w)}{\sum_{w \in E} b(\{w\})} = \frac{\sum_{w \in A \cap E} b(\{w\})}{\sum_{w \in E} b(\{w\})}$$

This conclusion of the proof requires us to limit the applicability of Lemma 2 to a slightly restricted class of non-classical logics. What justified this identity for classical semantics is that the indicator function $\chi_a(w)$, which assigns 1 to worlds where A is true and 0 to worlds where A is false, simultaneously serves as an indicator function for those beliefs that apply to worlds where A is true: it effectively cancels all beliefs concerned with worlds where A is false, thus eliminating such beliefs from the summation $\sum_{w \in E} b(\{w\})\chi_A(w)$, and only preserving those beliefs $b(w)$ that apply to worlds w where A is true. $\chi_a(w)$ can be considered as a coefficient depending on the truth of A at w: if A is false at w, this coefficient is 0, so $b(w)$ is multiplied by 0 and cancels out, and if A is true at w, this coefficient is 1, so $b(w)$ is multiplied by 1, preserved and added to all other beliefs applying to the worlds where A is true. This is how $\sum_{w \in E} b(\{w\})\chi_A(w)$ can be said to equal $\sum_{w \in A \cap E} b(\{w\})$.

As a consequence, the proviso that we have to impose on the generalisation of Lemma 2 is that the Williams mappings for alternative semantics may only assign 0 or 1 to non-classical truth statuses. Fortunately, this does not dramatically affect our argument, as most of Williams' proposed mappings only send truth statuses to either 0 or 1. The only semantics that we have to exclude from our generalisation are fuzzy logics and degree supervaluationism. For all other semantics treated by Williams, the final identity of the proof of Lemma 6 remains applicable, which establishes this lemma for Kleene logics, LP, supervaluationism and intuitionism.

Almost trivially, the proof of Theorem 2 then also holds for non-classical truth statuses, as projected onto [0,1], for the specified logics. Supposing the legitimacy of the Williams mappings, we have been able to show that expected inaccuracy can be defined as it was for the classical case, that inaccuracy measures generally have to be quadratic, and that Lemma 2 holds. This is enough for Theorem 1 to hold as well.

We would like to be somewhat cautious in our final step, which is to conclude that the generalised proof of Theorem 1 also generally entails probabilism. Let us recall the logical axiomatisation of probability formulated in section 2.1. (P1)-(P3) are the logical equivalents of (K1)-(K3), provided that a logic satisfies (i)-(iii). It is easily checked that Kleene logics, LP, supervaluationism and intuitionism all satisfy (i)-(iii), given that we base ourselves on the relevant mappings ([Wil12], 517-519). Thus, if we can establish (P1)-(P3) for the mentioned logics, this is equivalent to establishing (K1)-(K3). As the generalisation of Theorem 2 proves (P1)-(P3) for rational credences for the given background logics, it follows that (K1)-(K3) also apply to these cases, and that we can indeed conclude probabilism.

3.6 Accuracy and non-classical semantics

Consequently, we can show that the interactions between probabilistic credence and truth that take place in the adapted notion of gradational accuracy remain unproblematic, by the same considerations that justified the dynamics between truth and credence for the classical case.

For whatever semantics we decide to use, we are always justified in interpreting credences as expected truth values. Smith in [Smi10] argues that in general, degrees of belief are best interpreted in this way, so also if we adopt non-classical semantics, it makes sense for an agents' belief in a statement to reflect the anticipated, possibly non-classical truth value of such a proposition. Considering credences as approximated truth values, as we saw before, resolves the potential conflict between logical and probabilistic semantics.

Now, any technical difficulties in the interactions between logical truth and probabilistic credence are defused by the reconstructed Euclidean space where, again, we can situate credences and truth values together. This does, of course, require that we agree on suitable Williams mappings, which make it possible for such a space to exist for non-classical semantics as well. As for now, the restored geometric analogy between credences and truth values seems to depend on the existence of such mappings.

Although the described generalisation implies no *technical* changes to gradational accuracy and the proofs of Lemma 6 and Theorem 2 as such, we claim that the required construction of a Williams style mapping may in fact have serious implications from a *semantic* point of view. In other words, we wish to investigate the assumption on which general gradational accuracy, and thereby the generalised Leitgeb and Pettigrew argument, is based, namely the existence and legitimacy of Williams' mappings. We will show how projecting truth statuses onto the interval [0,1] may not at all be a straightforward procedure, and how underlying semantics may have to be manipulated or damaged for such mappings to exist. To create a clear understanding of what may happen if we force non-classical truth statuses into [0,1], we will take LP as a case study, and investigate how possible Williams style mappings may disrespect the particular semantic characteristics of LP.

4 Case study: the logic of paradox

This final section will provide a critical assessment of the Williams style generalisation of the Leitgeb and Pettigrew argument. By considering LP, we will illustrate how vital aspects of non-classical semantics may be damaged by the constraints that Williams' truth mappings impose on logical truth. We will evaluate what is at stake, and how we may have to revise the generalisation of gradational accuracy in order for the probabilism argument to still go through. By suggesting three alternatives, we will attempt to establish probabilism for LP by other means than Williams' mapping, and examine the influence of these alternatives on the logical-probabilistic dynamics.

4.1 Relevance

LP was introduced by Priest in 1979 ([Pri79]). The most characteristic property of its semantics is the use of a truth glut: the set of truth values is $\{0, 1, \mathsf{both}\}$, with both assigned to formulas that are both true and false. Priest motivated this semantics by stressing the multitude of attempted and failed solutions to logical paradoxes, and by arguing that, perhaps, the best way to approach paradoxes is not by trying to solve them, but by simply accepting that there exist sentences that are true and false at the same time: 'Suppose we stop banging our heads against a brick wall trying to find a solution, and accept the paradoxes as brute facts.' (220) This is the dialetheist view underlying the introduction of both.

In LP, 1 and both are the designated truth values. Thus, in LP valuations, contradictions can take the value both, which is a designated truth value, so that ex falso no longer holds. This means that LP, contrary to classical logic, does not trivialise due to contradictions: it is a paraconsistent logic. Dialetheism and paraconsistency make the LP semantics suitable for reasoning about paradoxes in a way that most other logics do not tolerate. This makes LP worth considering in the context of epistemic utility for at least two reasons, on which we will now elaborate.

First, epistemic utility deals with agents who *believe* formulas to a certain degree. There are all kinds of things that can prevent them from *knowing* these formulas: conflicting pieces of evidence, uncertainty or ambiguity in the formulas themselves. If there is one type of formula that can be called ambiguous, it is the logical paradox. In the case of paradoxes, systems of (classical) inference break down, and we have no analytic means to clearly and securely determine truth or falsity. Usually, it is right to say that we do not know how to evaluate such statements. How we treat them epistemically or doxastically, is eventually not a matter of knowledge, but *belief.* And often, in cases such as the Liar Sentence, our belief fluctuates between truth and falsity. The system does not allow us to derive definite knowledge about such contradictory statements, and all we can do is entertain a belief that balances between truth and falsity - a belief that takes a dialetheist form. In fact, immediate and intuitive responses to paradoxes are often characterised by the uncomfortable belief that these sentences are both true and false at the same time. Such dialetheist reasoning appears to be a natural reaction to paradoxes, and because we wish to consider the beliefs that agents can have with respect to any kind of propositions, we would also like to accommodate beliefs about paradoxes. As these beliefs often display dialetheist behaviour, we would like to have a semantics at our disposal that is also dialetheist. Therefore, LP can be of use.

Secondly, apart from its dialetheism, the paraconsistency of LP is also relevant to epistemic utility theory. Initially, we do not assume the agents to reason probabilistically or according to any inferential pattern. It is very well possible that they entertain all kinds of contradictory thoughts. Such contradictory thoughts need not be irrational. In the case of logical paradoxes, for example, we sometimes have no other option than to expect a proposition to be false as strongly as we expect it to be true. We expect, or believe, it to be both false and true, and as Priests argues, we may wish to adopt a semantics that allows the assignment of simultaneous truth and falsity in such cases. Yet, we do not want the entire system to trivialise due to our dialetheist approach to paradoxes: we want paraconsistent semantics. LP is such a paraconsistent semantics.

These aspects of the LP semantics make the logic interesting to consider in the context of epistemic utility theory. These are also the features that are at stake if we choose to apply Williams' mapping for LP.

4.2 Williams' mapping

Let us now see which mapping Williams specifies for LP semantics in [Wil12]. Generally, the Williams mappings are defined as $[[\bullet]] : S \to [0,1]$, where S is the set of truth statuses for a particular semantics, c.q. $\{0, 1, \mathsf{both}\}$. For LP semantics, which is referred to as 'LP gluts', Williams describes the mapping $[[\bullet]]$:

> $[[\bullet]]$ sends 1 (true), both to 1 and 0 (false) to 0. (517)

$[[\bullet]]$ maps the truth statuses to a 'truth value' in [0,1], and apparently, it sends the designated truth values 1 and both to the same point in the [0,1] interval, namely 1, and 0 to 0. We recall how the mapping $[[\bullet]]$ from truth statuses to [0, 1] is supposed to express how such 'statuses function as aims for credence' (516). Assuming that credences, as expected truth values, are better if they are closer to the actual truth values, this Williams mapping thus allows for an analogue of gradational accuracy for LP.

What draws our attention is Williams' proposal to map the two designated truth values to the same point in the [0,1] interval: ' $[[\bullet]]$ sends 1, both to 1'. This is a somewhat worrying aspect of the mapping, as it implies that 1 and both have the same doxastic role, and that there is no difference between the respective ways in which they should serve as aims for credences. In practice, this implies that the most characteristic feature of the LP semantics, namely the third truth value both, is eliminated, or made indiscernible from the classical truth value 1. Thus, in order for the LP truth statuses to be situated in the [0, 1] interval, and for the accuracy geometry to be maintained, Williams proposes a mapping that manipulates the LP semantics rather drastically.

We would like to be more careful, and respect the distinction between LP's designated truth values. This is to say that we would also like to differentiate between the respective doxastic roles that 1 and both should play. As the Williams mapping equates these roles, what it actually says is that in the context of belief, there is no difference between 1 and both: expecting a proposition to be true is exactly the same thing as expecting it to be both true and false. The value both is collapsed into 1. Apart from the reconstruction of the Euclidean accuracy frame, Williams offers no convincing arguments for this intervention.

A nice semantic geometry and a smooth generalisation of gradational accuracy do not seem to justify this manipulation of the LP semantics. We argue that 1 and both do have different doxastic roles, and that this difference should be taken into account when composing a Williams style mapping, or in general, when dealing with LP in a doxastic-epistemic setting. Indeed, we discussed how the use of both can be an important reason for implementing the LP semantics, so the way in which the proposed Williams mapping destroys the distinction between the two designated truth values defeats the very purpose of adopting LP.[7]

[7] We now see how the interpretative issue raised in footnote 5 does not affect our argument. Whether we treat the images of Williams' mapping as truth values or as ideal credences, we face the same problem: the distinct designated truth statuses are concatenated into a single point, which is in defiance of the philosophical and conceptual justification of the LP semantics.

As an illustration of the essential (doxastic) difference between the truth statuses 1 and both, we return to the issue of paradoxes. Evaluating a logical paradox as only true or as both true and false is not at all the same thing, nor is believing such sentences to be only true doxastically identical to the belief that they are both true and false. In the preceding section we explained how dialetheism is an important motivation for adopting LP semantics. It allows us to evaluate paradoxes as simultaneously true and false, which is something different than the evaluation of (non-paradoxical) statements as only true. Similarly, believing such paradoxical formulas to be both true and false is radically different from believing them to be only true. Yet, Williams' LP mapping does not distinguish between the doxastic roles of both truth statuses, and biases the true side of both, thereby breaking the symmetry of the truth glut, which we want to emphasise as being true *as well as* false. Thus, the mapping does not manage to translate dialetheism to the [0, 1] scale. As this is one of the most vital aspects of the LP semantics, we may wish to reconsider the proposed mapping.

The mapping's disregard of dialetheism in fact reduces the LP semantics to regular, classical semantics. From the output of the mapping we could not infer that the adopted semantics are those of LP, because the resulting distribution of truth values behaves according to the classical Tarskian truth conditions. We conclude that the distinguishing properties of LP are damaged by the proposed Williams mapping, and that we would like to find an alternative that is at better terms with the LP semantics.

In the coming sections, we will consider how the inadequacy of the Williams mapping for LP semantics influences the generalised Leitgeb and Pettigrew argument, and explore some possible alternatives. Before continuing, however, let us emphasise that the issues raised in the above apply specifically to the LP mapping, but that very similar issues may emerge when we have a closer look at the mappings that Williams proposes for other non-classical semantics. The problem is that, whatever mapping we use, some semantics just do not combine nicely with the [0, 1] interval in any way. Certain aspects of the semantics will always have to be sacrificed for the sake of the [0, 1] geometry. For a semantics of Kleene gaps, for instance, Williams maps 1 to 1, and 0 and neither to 0, thus eliminating the third truth value neither. This is debatable on the same grounds as the elimination of both in LP, to which we objected in the above. Comparable arguments can be used against Williams' proposed mappings for other non-classical semantics. Hence, in our discussion of LP, we should bear in mind that similar considerations hold for other semantics as well.

4.3 Consequences for the probabilism argument

We would now like to assess how the findings of the preceding section influence the generalisation of the Leitgeb and Pettigrew argument for the case of LP semantics.

It is clear that we have to make a decision: either we sacrifice characteristic parts of the LP semantics in order for the Williams mapping to hold and the probabilism argument to go through, or we insist to preserve the LP semantics as they are, in which case we have to reject the Williams mapping. Now, as the main reason to adopt LP is the relevance of the characteristic aspects of its semantics, it seems counterintuitive to eliminate such aspects for the sake of a geometrically convenient mapping to the [0,1] interval. We therefore abandon the Williams mapping.

What does this entail for the probabilism argument? Our current position is an awkward one, as we no longer have a defined notion of gradational accuracy for LP semantics. We criticised the assumptions underlying Williams' treatment of the LP semantics, and rejected his mapping. Now we are back at the essential LP semantics, which as such are not compatible with the conception of gradational accuracy that forms the basis of the Leitgeb and Pettigrew argument. The bare set of LP truth statuses $\{0, 1, \mathsf{both}\}$ does not fit this notion, so we are unable to establish probabilism for LP semantics using Leitgeb and Pettigrew's accuracy based line of reasoning. Thus, the relation between logical truth and probability breaks down for LP.

Yet, we would still like to establish probabilism. As we stated in section 3.5, we generally seek to justify probabilism in rational credences, which includes the cases of non-classical logics such as LP. Also with the LP semantics, we do not want an agent to believe a proposition more strongly than any of its logical consequences. The introduction and interpretation of a third truth value does not change this. Moreover, we would still like to establish probabilism on epistemic rather than pragmatic grounds, which indicates that we should first explore the possible alternatives that can be found within epistemic utility theory and the framework that it already offers for such projects. Indeed, our use of Williams' mapping is only one possible attempt to generalise the Leitgeb and Pettigrew argument to semantics such as LP, and perhaps it would be best to also explore other ways in which we could generalise their accuracy based argument.

We should note that our wish to establish probabilism for LP semantics simultaneously amounts to the construction of new interactions between logical truth and probabilistic credence. In exploring the possibilities, we should remain conscious of the issues that were discussed at the beginning of the paper, and which we do not want to cause difficulties.

4.4 Suggested alternatives

4.4.1 A revised Williams mapping

Instead of using the exact same mapping that Williams proposes for LP in his paper, we could think of an adaptation of this mapping that inflicts less damage to the LP semantics. In section 3.5, we saw how the Leitgeb and Pettigrew argument can be generalised if we assume that there exists a legitimate Williams mapping for non-classical semantics. Thus, if we manage to find a suitable mapping for the LP truth statuses that does respect the LP semantics, we have a legitimate generalisation of the accuracy argument to LP. Following Williams, we refer to the required mapping as $[[\bullet]]$.

What would this legitimate Williams style mapping look like? Primarily, we should take care that the mapping distinguishes between the truth statuses 1 and both, unlike the mapping proposed by Williams himself. Semantically, the distinction between 1 and both is just as follows: statements that are both are not only true, but also false, whereas those that are 1 are only true. This is a crucial point that we require our mapping $[[\bullet]]$ to reflect. The most intuitive option would be to map both to the sum of the points on the real axis to which 0 and 1 are mapped. The original Williams mapping already does this: 0+1=1, yet without discriminating between the designated values. Thus, we want our mapping to satisfy the following two conditions: 1. $[[\bullet]]$ sends 1 and both to distinct points on the real axis 2. $[[\bullet]]$ sends both to the sum of the images of 0 and 1

Additionally, we would like to impose a third condition: 3. $[[\bullet]]$ sends 0 and 1 to distinct points such that the Euclidean distance between these points is 1

If we expand the codomain of the Williams mapping, the following function would satisfy conditions 1-3: $[[\bullet]]$ sends 1 to 1/2, 0 to -1/2 and both to 0.

In order to maintain the geometric analogy between truth values and credences, adopting this mapping would also demand that we translate credences $b_a(w)$ so that their scale matches the one of the mapped truth values. By condition 3, this is easily done by means of the following mapping, which we call $[[\star]]$:

$$[[\star]] \text{ sends credences } b_a(w) \in [0,1] \text{ to } b(w) = (b_a(w) - 1/2) \in [-1/2, 1/2]^8$$

[8] Note that this translation does not actually imply negative credences, but that it *maps* credences in [0, 1/2]

The relocated credences $b(w)$ can thus be interpreted as approximations of the truth values generated by the mapping $[[\bullet]]$.

Conditions 1 and 2 of the proposed mapping $[[\bullet]]$ guarantee that the characteristic properties of the LP semantics are preserved, so from the LP perspective, we are more satisfied with this mapping than with the one proposed by Williams. The next question is whether we are also satisfied with this mapping from a doxastic-epistemic point of view: can we use it to establish probabilism for LP, preferably by means of an adapted version of the Leitgeb and Pettigrew argument?

The implied geometry allows for the same Euclidean approach to inaccuracy measures. Moreover, by condition 3 we made sure that inaccuracy is situated on a quantitative scale analogous to the original one, so by an argument identical to that in 2.5 we can justify that the legitimate inaccuracy measures must be quadratic.

However, a difference with Leitgeb and Pettigrew's Lemma 2 is that the expected local inaccuracy $\sum_{w\in E} b(\{w\})I(A,w,x)$ is no longer a positive quadratic in the variable x, because b can now also take negative values. $I(A,w,x)$ is still positive, of course, so the absolute value of the expected local inaccuracy can be expressed as $\sum_{w\in E} \mid b(\{w\}) \mid I(A,w,x)$, which is a positive quadratic in the variable x. As the local inaccuracy is always positive, it makes sense for the expected local inaccuracy to also be a positive value, so we adopt the new expression. With this absolute expected local inaccuracy we check whether Lemma 2 still applies, and run through the proof:

By definition, $\sum_{w\in E} \mid b(\{w\}) \mid I(A,w,x) = \sum_{w\in E} \mid b(\{w\}) \mid \lambda(\chi_A(w)-x)^2$, so

$$\begin{aligned}\frac{d}{dx}\sum_{w\in E} \mid b(\{w\}) \mid I(A,w,x) &= 2\lambda(x\sum_{w\in E} \mid b(\{w\}) \mid - \sum_{w\in E} \mid b(\{w\}) \mid \chi_A(w))\\ &= 0,\end{aligned}$$

if and only if

$$x = \frac{\sum_{w\in E} \mid b(\{w\}) \mid \chi_A(w)}{\sum_{w\in E} \mid b(\{w\}) \mid}.$$

Now, we seem to run into the problem that was pointed out above. Because $[[\bullet]]$ does not only assign the values 0 and 1, we cannot directly prove the final identity of Lemma 2. Yet, we would like to think that our assignment of the values 1/2, -1/2 and 0 should create less of a problem than the continuous codomain of the Williams mapping for e.g. fuzzy logic. And indeed, we can think of a way to overcome this difficulty, which is inspired by the same reasoning that implicitly justifies the final identity of Lemma 2 for classical semantics. We recall that we made the following remark about the identity concerned:

> $\chi_A(w)$ can be considered as a coefficient depending on the truth of A at w: if A is false at w, this coefficient is 0, so $b(w)$ is multiplied by 0 and cancels out, and if A is true at w, this coefficient is 1, so $b(w)$ is multiplied by 1, preserved and added to all other beliefs applying to the worlds where A is true. This is how $\sum_{w\in E} b(\{w\})\chi_A(w)$ can be said to equal $\sum_{w\in A\cap E} b(\{w\})$.

to negative values. Because the mapped ehave like credences, but only on a different quantitative scale, we are still justified in calling them credence functions, but we should be aware that what we are referring to is really the image of the credences actually held. $b(w)$ should be regarded as 'dummy' credences, by means of which we will aim to establish probabilism for the $b_a(w)$. For the sake of typographic convenience in the upcoming calculations, we chose $b_a(w)$ to denote the actual credences, and $b(w)$ to denote those that result from our mapping $[[\star]]$.

With our mapping $[[\bullet]]$, $\sum_{w\in E} \mid b(\{w\}) \mid \chi_A(w)$ does not equal$\sum_{w\in A\cap E} \mid b(\{w\})$. However, we can replace the final identity of the proof of Lemma 2 by deriving the following, new identity for our particular mapping $[[\bullet]]$:

$$\begin{aligned} x &= \frac{\sum_{w\in E} \mid b(\{w\}) \mid \chi_A(w)}{\sum_{w\in E} \mid b(\{w\}) \mid} \\ &= \frac{1}{2}\frac{\sum_{w\in A_T\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} - \frac{1}{2}\frac{\sum_{w\in A_F\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} \end{aligned}$$

Although it appears to be quite different from the conclusion of the proof of Lemma 6, the above identity was obtained by exactly the same reasoning. Here as well, the indicator function $\chi_A(w)$, c.q. the output of $[[\bullet]]$, is treated as a coefficient depending on the truth value of a proposition A. The difference with the classical case is that we do not assign the truth values 0 and 1, but -1/2, 0 and 1/2. We can partition the set of epistemically possible worlds E into the sets A_T, A_F and A_B. These denote, respectively, the set of worlds where proposition A is (only) true, the set where it is (only) false, and the set where it is both true and false. As $[[\bullet]]$ maps 1 (true) to 1/2, 0 (false) to -1/2 and both to 0, we can separate terms and obtain the above identity. Note that, because both is mapped to 0, $\chi_A(w)$ is 0 for all worlds where A is both true and false, so the term containing A_B cancels out. Lemma 2 thus becomes:

Lemma 3. Suppose $I(A, w, x) = \lambda(\chi_A(w) - x)^2$. Suppose W is finite, b and b' are credence functions, $A, E \subseteq W$, and $\sum_{w\in E} b(\{w\}) \neq 0$. Then the following two propositions are equivalent:

(i) For all $A \subseteq W$ and $x \in [-\frac{1}{2}, \frac{1}{2}]$,

$$\sum_{w\in E} \mid b(\{w\}) \mid I(A, w, b'(A)) \leq \sum_{w\in E} \mid b(\{w\}) \mid I(A, w, x).$$

(ii) For all $A \subseteq W$,

$$b'(A) = \frac{1}{2}\frac{\sum_{w\in A_T\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} - \frac{1}{2}\frac{\sum_{w\in A_F\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid}.$$

Note that proposition (ii) implies that also

$$\mid b'(A) \mid= \frac{1}{2}\frac{\sum_{w\in A_T\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} - \frac{1}{2}\frac{\sum_{w\in A_F\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid}.$$

Next, we wish to prove the following version of Theorem 1:

Theorem 4. Suppose b is a credence function, $E \subseteq W$, $\sum_{w\in E} b(\{w\}) \neq 0$ and I is a quadratic local inaccuracy measure. Then the following two propositions are equivalent:

(i) For all $A \subseteq W$ and $x \in [-\frac{1}{2}, \frac{1}{2}]$,

$$\sum_{w\in E} \mid b(\{w\}) \mid I(A, w, b(A)) \leq \sum_{w\in E} \mid b(\{w\}) \mid I(A, w, x)$$

(b) Credence function b_a that is mapped to b by $[[\star]]$ is a probability function.

By Lemma 3, it suffices to show that, for b a credence function (mapped according to $[[\star]]$) and $E \subseteq W$, with $\sum_{w \in E} b(\{w\}) \neq 0$,

$$b(A) = \frac{1}{2} \frac{\sum_{w \in A_T \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid} - \frac{1}{2} \frac{\sum_{w \in A_F \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid},$$

if and only if the function $b_a(w)$ that $[[\star]]$ maps to b is a probability function on the power set of W and $b(\{w\}) = 0$ for $w \notin E$.

Now, what we suggest is to adopt a new understanding of $b(A)$, belief in a proposition A. Initially, $b(A)$ was defined as $\sum_{w \in A \cap E} b(\{w\}) = \sum_{w \in E} b(\{w\}) \chi_A(w)$. In the classical case we have that

$$\sum_{w \in A} b(\{w\}) = \sum_{w \in A \cap E} b(\{w\}) = \sum_{w \in E} b(\{w\}) \chi_A(w)$$

In other words, $b(A)$ can be expressed as a weighted summation of the beliefs in all epistemically accessible worlds E, with $\chi_A(w)$ the weighting factor (or coefficient). Thanks to the convenience of bivalence, this is an immediate result for classical semantics. For semantics with other $\chi_A(w)$ than 0 and 1, such as LP, we saw that the same does not hold. We argue, however, that it is very well possible to redefine $b(A)$, and treat it as a weighted summation:

$$b(A) \equiv \sum_{w \in E} b(\{w\}) \chi_A(w)$$

According to this new definition, the belief in a proposition depends on the truth value of this proposition at different worlds. This may seem strange, but we already recognised how $b(A)$ equals $\sum_{w \in A} b(\{w\}) = \sum_{w \in E} b(\{w\}) \chi_A(w)$ for classical semantics, and have no strong conceptual arguments against the postulation of a definition by means of a weighted summation over E as more primitive than an unweighted summation over A. In a sense, our proposed redefinition amounts to little more than making explicit what was already implicit in the case of classical semantics. Moreover, the new definition may in fact be more appropriate for certain non-classical semantics than the original one, because it recognises that the set $\{w \mid w \in A\}$ may not always be well-defined if $\chi_A(w)$ takes other values than just 0 and 1. Hence, we decide to adopt the new definition, which implies the following for LP:

$$b(A) \equiv \sum_{w \in E} b(\{w\}) \chi_A(w) = \frac{1}{2} \sum_{w \in A_T \cap E} b(\{w\}) - \frac{1}{2} \sum_{w \in A_F \cap E} b(\{w\})$$

Proof. With this definition of $b(A)$, we can prove a version of Theorem 1 for LP semantics. For the 'if' direction, it is enough to prove that, if b_a is a probability function and $b(\{w\}) = 0$ for $w \notin E$, then for all $A \subseteq W$,

$$\mid b(A) \mid = \frac{1}{2} \frac{\sum_{w \in A_T \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid} - \frac{1}{2} \frac{\sum_{w \in A_F \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid}$$

We suppose that b is generated by mapping the probability function b_a according to $[[\star]]$, and $b(\{w\}) = 0$ for $w \notin E$. Then, $0 = b_a(\emptyset)$, and

$$\begin{aligned}
-\frac{1}{2} &= b(\emptyset) \\
&= \sum_{w \in E} b(\{w\})\chi_{\emptyset}(w) \\
&= \frac{1}{2} \sum_{w \in \emptyset_T \cap E} b(\{w\}) - \frac{1}{2} \sum_{w \in \emptyset_F \cap E} b(\{w\}) \\
&= \frac{1}{2} \sum_{w \in \emptyset} b(\{w\}) - \frac{1}{2} \sum_{w \in E} b(\{w\})
\end{aligned}$$

Thus, $\sum_{w \in E} b(\{w\}) = 1 = \sum_{w \in E} \mid b(\{w\}) \mid$ and

$$\begin{aligned}
\mid b(A) \mid &= \mid \sum_{w \in E} b(\{w\})\chi_A(w) \mid \\
&= \sum_{w \in E} \mid b(\{w\}) \mid \chi_A(w) \\
&= \frac{1}{2} \sum_{w \in A_T \cap E} \mid b(\{w\}) \mid - \frac{1}{2} \sum_{w \in A_F \cap E} \mid b(\{w\}) \mid \\
&= \frac{\frac{1}{2} \sum_{w \in A_T \cap E} \mid b(\{w\}) \mid - \frac{1}{2} \sum_{w \in A_F \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid} \\
&= \frac{1}{2} \frac{\sum_{w \in A_T \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid} - \frac{1}{2} \frac{\sum_{w \in A_F \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid},
\end{aligned}$$

as required. For the 'only if' direction, we have to show that, if b is a (mapped) credence function and, for all $A \subseteq W$,

$$b(A) = \frac{1}{2} \frac{\sum_{w \in A_T \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid} - \frac{1}{2} \frac{\sum_{w \in A_F \cap E} \mid b(\{w\}) \mid}{\sum_{w \in E} \mid b(\{w\}) \mid},$$

then it follows that the credence function b_a that $[[\star]]$ maps to b satisfies the Kolmogorov probability axioms.[9] We find the following:

(K1) If $A \subseteq W$, then $b(A) \geq -1/2$, because $b : P(W) \to [-\frac{1}{2}, \frac{1}{2}]$.

(K2) $b(\emptyset) = \frac{1}{2} \frac{\sum_{w \in \emptyset \cap E} |b(\{w\})|}{\sum_{w \in E} |b(\{w\})|} - \frac{1}{2} \frac{\sum_{w \in E \cap E} |b(\{w\})|}{\sum_{w \in E} |b(\{w\})|} = -\frac{1}{2}$, and

$b(W) = \frac{1}{2} \frac{\sum_{w \in E \cap E} |b(\{w\})|}{\sum_{w \in E} |b(\{w\})|} - \frac{1}{2} \frac{\sum_{w \in \emptyset \cap E} |b(\{w\})|}{\sum_{w \in E} |b(\{w\})|} = \frac{1}{2}$.

[9] We use the Kolmogorov axioms because conditions (i)-(iii) of section 2.1 are not satisfied, so (P1)-(P3) are not applicable. We are justified in doing so because our construction of the sets A_T, A_F and A_B, together with our redefinition of $b(A)$, allows us to apply the possible worlds approach of the Kolmogorov axiomatisation.

(K3) If $A, B \subseteq W$ are disjoint, then

$$\begin{aligned}
b(A \cup B) &= \frac{1}{2}\frac{\sum_{w\in(A\cup B)_T\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} - \frac{1}{2}\frac{\sum_{w\in(A\cup B)_F\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} \\
&= \frac{1}{2}\frac{\sum_{w\in A_T\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} + \frac{1}{2}\frac{\sum_{w\in B_T\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} \\
&\quad - \frac{1}{2}\frac{\sum_{w\in A_F\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} - \frac{1}{2}\frac{\sum_{w\in B_F\cap E} \mid b(\{w\}) \mid}{\sum_{w\in E} \mid b(\{w\}) \mid} \\
&= b(A) + b(B),
\end{aligned}$$

since $(A \cup B)_T \cap E = (A_T \cap E) \cup (B_T \cap E)$ and $(A \cup B)_F \cap E = (A_F \cap E) \cup (B_F \cap E)$.

It follows that b, as mapped by $[[\star]]$, is a probability measure that has been translated by -1/2 along the real axis. Recalling that $[[\star]]$ sends credences $b_a(w) \in [0, 1]$ to $(b(w)-1/2) \in [-1/2, 1/2]$, we can conclude that the actual (unmapped) credence function $b_a(w)$ satisfies the Kolmogorov axioms, and is a proper probability function. This completes the proof of (an analogue of) Theorem 1 for LP. We have now established probabilism for LP using an adapted version of the Leitgeb and Pettigrew argument, based on a Williams style mapping that is more considerate of the semantic particularities of LP.

□

Finally, now that we have established that rational credences must be probabilistic also when we use LP as our background logic, we must analyse how the suggested justification influences the interactions between truth and probability. We return again to the issues raised in section 2.3. As we left most of Leitgeb and Pettigrew's technical apparatus intact, we can discuss these issues rather briefly, and refer to the conclusions reached in section 3.4.

Our interpretation of credences as approximated truth values has not changed, so also in this new suggestion we do not have to fear any semantic incompatibility between logical truth and probabilistic credence. The interaction between both notions in inaccuracy measures is shaped by the same, justified Euclidean arithmetic that was also appropriate for classical semantics. As for the status of logical truth, we went at great lengths to guarantee that the characteristic properties of truth in LP are respected by our proposed mapping $[[\bullet]]$. In fact, the preservation of these properties was the main motivation to revise Williams' mapping, so we do not risk any unjustified mutations of LP truth in our framework. Also, we do not perform any computations that may give rise to a conflict between the compositionality of logical truth and the non- compositionality of probabilistic credences, and the issue concerning semantic realism is still as unproblematic as it was in 2.4. We conclude that the relation between truth and probability constitutes no difficulties in the proposed set-up.

4.4.2 A two-dimensional approach to truth values and credences

We could consider abandoning the one-dimensional approach to local inaccuracy altogether, and adopt a two-dimensional Euclidean frame in order to incorporate the dialetheism of LP.[10] We can then redefine local inaccuracy in such a way that both the Euclidean geometry and the dialetheist character of LP semantics are maintained. The great advantage of this method is that we no longer need a mapping for truth statuses, as we treat logical truth as a two-dimensional vector that supports immediate localisation in a two-dimensional plane. See Figure 3 for an illustration.

[10]Through personal communication with Richard Pettigrew I am aware that Williams has suggested a treatment of truth values as lattices, which appears similar in spirit to the two-dimensional approach to logical truth that I propose here. However, no publication has yet resulted from Williams' suggestion.

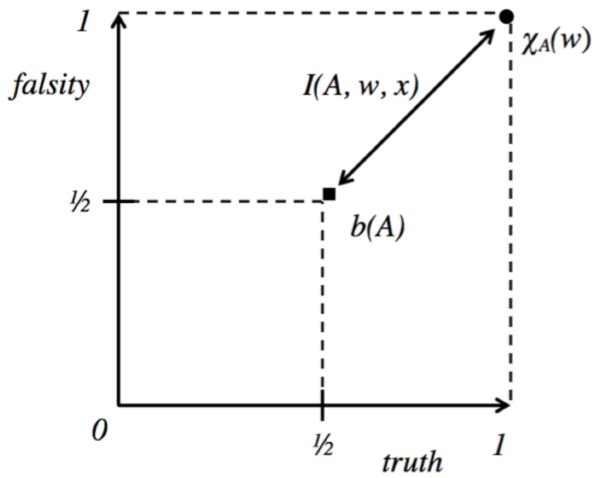

Figure 3: Two-dimensional truth values and credences.

The one-dimensional, real line along which degrees of belief $b(A)$ in propositions A and truth values $\chi_A(w)$ of such propositions at world w were plotted in the case of classical semantics (see Figure 1), is replaced by a two-dimensional Euclidean plane. The horizontal axis expresses the truth value of a proposition, and the vertical axis its 'falsity value'. The falsity value really is an analogue of the truth value: where truth values express the extent to which a formula is true, falsity values express the extent to which it is false. Both axes run from 0 to 1. For credences, the axes should be understood as 'expected truth' and 'expected falsity'. Thus, $\chi_A(w)$ and $b(A)$ become two-dimensional vectors:

$$\chi_A(w) = \begin{bmatrix} \tau(\chi_A(w)) \\ \varphi(\chi_A(w)) \end{bmatrix} \tag{1}$$

and

$$b(A) = \begin{bmatrix} \tau(b(A)) \\ \varphi(b(A)) \end{bmatrix}, \tag{2}$$

where $\tau(\chi_A(w))$ and $\varphi(\chi_A(w))$ respectively denote the logical truth and falsity values of a proposition A at a world w, while $\tau(b(A))$ and $\varphi(b(A))$ denote the expected truth and falsity values of A. We see how the LP semantics are respected: distinct (expected) designated truth values are kept distinct, neither the false nor the true part of both is biased, and dialetheism is not subordinated to the convenience of a particular geometry.

Suggesting this approach is the result of taking seriously the interpretation of credences as expected truth values, and the dialetheist perspective of LP. We exploit LP's tolerance towards simultaneous truth and falsity by allowing credences to anticipate not only a certain degree of truth, but also a certain degree of falsity. These degrees do not necessarily add up to 1: if $b(A)$ is situated in the upper right half of the plane, the sum of expected truth and falsity exceeds 1, and in the lower left half, it is lower than 1. In fact, the only region of the plane where expected truth and falsity do add up to 1 is the straight line from (0,1) to (1,0). Although the LP semantics cannot realise any other truth values than the ones situated at (0,1) (0), (1,0) (1) and (1,1) (both), this is not a problem. We remember that also in the one-dimensional Euclidean inaccuracy conception, expected truth values other than 0 and 1 were possible, even though classical semantics are bivalent. The fact that the plane theoretically tolerates unconventional expected truth values does not mean that these should also be anticipated in practice. Of course, one could isolate the subsets of the proposed plane that are considered 'admissible', and only take these into consideration, but such a restriction would be redundant. As it was unproblematic to let expected truth values run from 0 to 1 for classical local inaccuracy, it should also be possible to let expected truth and falsity values run from 0 to 1 for LP semantics, which generates the proposed plane. In general, however, we will assume that the expected truth and falsity either add up to 1, or to 2 if an agent expects both.

Figure 3 exemplifies how the local inaccuracy measure works in the two-dimensional LP plane. Both the expected truth value and the expected falsity value are 1/2, $\tau(b(A)) = \varphi(b(A)) = 1/2$, so that $b(A)$ is situated at the point (1/2, 1/2). The actual truth value of A at world w is both - the proposition is both true and false, so it has truth value 1 and falsity value 1. Thus, $\chi_A(w)$ is positioned at (1,1). The local inaccuracy depends on the Euclidean distance between $\chi_A(w)$ and $b(A)$:

$$\begin{aligned}
\| \chi_A(w) - b(A) \| &= \left\| \begin{bmatrix} \tau(\chi_A(w)) \\ \varphi(\chi_A(w)) \end{bmatrix} - \begin{bmatrix} \tau(b(A)) \\ \varphi(b(A)) \end{bmatrix} \right\| \\
&= \sqrt{(\tau(\chi_A(w)) - \tau(b(A)))^2 + (\varphi(\chi_A(w)) - \varphi(b(A)))^2} \\
&= \sqrt{(1 - \frac{1}{2})^2 + (1 - \frac{1}{2})^2} \\
&= \frac{1}{2}\sqrt{2}
\end{aligned}$$

We would like to establish that the appropriate inaccuracy measures are still quadratic. In section 3.2, we explained that Leitgeb and Pettigrew prove the legitimacy of quadratic inaccuracy measures by imposing local and global normality and dominance, comparability and minimum inaccuracy, while avoiding several epistemic dilemmas by also requiring agreement on inaccuracy, separability of global inaccuracy and agreement on directed urgency, together with continuous differentiability. On the basis of these constraints, Theorems 3-5 of their first 2010 paper can be proven. We would like to impose the same basic conditions on our inaccuracy measures for the newly defined LP space, and prevent the same epistemic dilemmas. As we changed the geometry, it seems that we cannot just apply the generalisation that we gave in 2.5, and by which we showed that non-classical semantics with a geometry-preserving Williams mapping also demand quadratic inaccuracy measures. However, our fundamental understanding of inaccuracy as dependent on the Euclidean distance between truth and credence remains unaltered. A close look at the proofs of Theorems 3-5 (230-234) teaches us that this is all we need for them to hold: there are no occurrences of isolated truth values or credences that could cause problems similar to the one concerning the final identity of Lemma 2, and the proposed two-dimensionality of individual truths and credences does not create any friction with the mathematical entities of the original proofs. We will not give a step-by-step justification of this claim, but refer to pages 230-234 of Leitgeb and Pettigrew's paper, where it is easily checked. Hence, we conclude that legitimate inaccuracy measures must still be quadratic. For the local inaccuracy measure $I(A, w, x)$, this implies that

$$I(A, w, x) = \lambda \mid\mid \chi_A(w) - x \mid\mid^2$$

The proposed redefinition of local inaccuracy relates to a new form of global inaccuracy in the same way that classical local inaccuracy related to global inaccuracy (see section 3.3). We still express global inaccuracy as dependent on the Euclidean distance between a credence b and a world w:

$$|| b - w ||= \sqrt{(b(A_1) - \chi_{A_1}(w))^2 + \cdots + (b(A_n) - \chi_{A_n}(w))^2}$$

As $\chi_A(w)$ and $b(A)$ are no longer scalars, we do have to modify our notion of multiplication. We will let $(\chi_A(w) - b(A))^2$ denote a dot product:

$$(\chi_{A_1}(w) - b(A))^2 = (\chi_{A_1}(w) - b(A)) \bullet (\chi_{A_1}(w) - b(A))$$

Furthermore, for every extra atomic proposition that we wish to take into account, the Euclidean space is not expanded by one, but by two dimensions (a truth and a falsity axis). It is difficult to visualise this new perspective on global inaccuracy, because for two atoms we already need four dimensions, and for one atom there is no difference with the local inaccuracy measure. Hence, we will not attempt such a visualisation, but emphasise that for any n dimensions, we can always calculate the Euclidean distance between a global credence function and a world, which gives us the global inaccuracy of such a function.

Furthermore, for every extra atomic proposition that we wish to take into account, the Euclidean space is not expanded by one, but by two dimensions (a truth and a falsity axis). It is difficult to visualise this new perspective on global inaccuracy, because for two atoms we already need four dimensions, and for one atom there is no difference with the local inaccuracy measure. Hence, we will not attempt such a visualisation, but emphasise that for any n dimensions, we can always calculate the Euclidean distance between a global credence function and a world, which gives us the global inaccuracy of such a function.

We are now ready to assess if we can prove Lemma 3 and 4, the analogues of Lemma 2 and Theorem 4 in order to establish probabilism for LP. For Lemma 2, we will prove the following version:

Lemma 5. Suppose $I(A, w, x) = \lambda \mid\mid \chi_A(w) - x \mid\mid^2$. Suppose W is finite, b and b' are credence functions, $A, E \subseteq W$, $\sum_{w \in E} \tau(b(\{w\})) \neq 0$ and $\sum_{w \in E} \varphi(b(\{w\})) \neq 0$. Then the following two propositions are equivalent:

(i) For all $A \subseteq W$ and $x \in [0,1] \times [0,1]$,

$$\sum_{w \in E} b(\{w\}) I(A, w, b'(A)) \leq \sum_{w \in E} b(\{w\}) I(A, w, x).$$

(ii) For all $A_T \subseteq W$, with $A_T = \{w \in E \mid \tau(\chi_A(w)) = 1\}$, $A_F = \{w \in E \mid \varphi(\chi_A(w)) = 1\}$ and $b(\{w\}) = \begin{bmatrix} \tau(b(\{w\})) \\ \varphi(b(\{w\})) \end{bmatrix}$,

$$b'(A) = \begin{bmatrix} \frac{\sum_{w \in A_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} \\ \frac{\sum_{w \in A_F \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} \end{bmatrix}.$$

Proof. We established that the local inaccuracy measure must be quadratic, so the expected local inaccuracy $\sum_{w \in E} b(\{w\}) I(A, w, x)$ becomes:

$$\begin{aligned}
\sum_{w\in E} b(\{w\})I(A,w,x) &= \sum_{w\in E} b(\{w\})\lambda \mid\mid \chi_A(w) - x \mid\mid^2 \\
&= \sum_{w\in E} b(\{w\})\lambda((\tau(\chi_A(w)) - \tau(x))^2 + (\varphi(\chi_A(w)) - \varphi(x))^2) \\
&= \sum_{w\in E} b(\{w\})\lambda(\tau(\chi_A(w)) - \tau(x)))^2 + \sum_{w\in E} b(\{w\})\lambda(\varphi(\chi_A(w)) - \varphi(x)))^2 \\
&= \sum_{w\in E} \begin{bmatrix}\tau(b(\{w\})) \\ \varphi(b(\{w\}))\end{bmatrix} \lambda(\tau(\chi_A(w)) - \tau(x)))^2 + \sum_{w\in E} \begin{bmatrix}\tau(b(\{w\})) \\ \varphi(b(\{w\}))\end{bmatrix} \lambda(\varphi(\chi_A(w)) - \varphi(x)))^2 \\
&= \lambda \sum_{w\in E} \begin{bmatrix}\tau(b(\{w\}))((\tau(\chi_A(w)) - \tau(x))^2 \\ \varphi(b(\{w\}))((\tau(\chi_A(w)) - \tau(x))^2\end{bmatrix} + \lambda \sum_{w\in E} \begin{bmatrix}\tau(b(\{w\}))((\varphi(\chi_A(w)) - \varphi(x))^2 \\ \varphi(b(\{w\}))((\varphi(\chi_A(w)) - \varphi(x))^2\end{bmatrix}
\end{aligned}$$

$x = \begin{bmatrix}\tau(x) \\ \varphi(x)\end{bmatrix}$, so $\frac{d}{dx} = \begin{bmatrix}d/d\tau(x) \\ d/d\varphi(x)\end{bmatrix}$, and

$$\begin{aligned}
\frac{d}{dx}\sum_{w\in E} b(\{w\})I(A,w,x) &= \frac{d}{dx}(\lambda \sum_{w\in E} \begin{bmatrix}\tau(b(\{w\}))((\tau(\chi_A(w)) - \tau(x))^2 \\ \varphi(b(\{w\}))((\tau(\chi_A(w)) - \tau(x))^2\end{bmatrix} \\
&\quad + \lambda \sum_{w\in E} \begin{bmatrix}\tau(b(\{w\}))((\varphi(\chi_A(w)) - \varphi(x))^2 \\ \varphi(b(\{w\}))((\varphi(\chi_A(w)) - \varphi(x))^2\end{bmatrix}) \\
&= \lambda \sum_{w\in E} \begin{bmatrix}d/d\tau(x)(\tau(b(\{w\}))((\tau(\chi_A(w)) - \tau(x))^2) \\ d/d\varphi(x)(\varphi(b(\{w\}))((\tau(\chi_A(w)) - \tau(x))^2)\end{bmatrix} \\
&\quad + \lambda \sum_{w\in E} \begin{bmatrix}d/d\tau(x)(\tau(b(\{w\}))((\varphi(\chi_A(w)) - \varphi(x))^2) \\ d/d\varphi(x)(\varphi(b(\{w\}))((\varphi(\chi_A(w)) - \varphi(x))^2)\end{bmatrix} \\
&= 2\lambda \begin{bmatrix}\tau(x)\sum_{w\in E}\tau(b(\{w\})) - \sum_{w\in E}\tau(b(\{w\}))\tau(\chi_A(w)) \\ \varphi(x)\sum_{w\in E}\varphi(b(\{w\})) - \sum_{w\in E}\varphi(b(\{w\}))\varphi(\chi_A(w))\end{bmatrix} \\
&= \begin{bmatrix}0 \\ 0\end{bmatrix},
\end{aligned}$$

if and only if

$$\tau(x)\sum_{w\in E}\tau(b(\{w\})) - \sum_{w\in E}\tau(b(\{w\}))\tau(\chi_A(w)) = 0$$

, and

$$\varphi(x)\sum_{w\in E}\varphi(b(\{w\})) - \sum_{w\in E}\varphi(b(\{w\}))\varphi(\chi_A(w)) = 0,$$

which is the case if and only if the following identities are true:

$$\begin{aligned}
\tau(x) &= \frac{\sum_{w\in E}\tau(b(\{w\}))\tau(\chi_A(w))}{\sum_{w\in E}\tau(b(\{w\}))} = \frac{\sum_{w\in A_T\cap E}\tau(b(\{w\}))\tau(\chi_A(w))}{\sum_{w\in E}\tau(b(\{w\}))} \\
\varphi(x) &= \frac{\sum_{w\in E}\varphi(b(\{w\}))\varphi(\chi_A(w))}{\sum_{w\in E}\varphi(b(\{w\}))} = \frac{\sum_{w\in A_T\cap E}\varphi(b(\{w\}))\varphi(\chi_A(w))}{\sum_{w\in E}\varphi(b(\{w\}))}
\end{aligned}$$

It follows that $x = \begin{bmatrix} \frac{\sum_{w \in A_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} \\ \frac{\sum_{w \in A_F \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} \end{bmatrix}$.

As $\sum_{w \in E} b(\{w\}) I(A, w, x)$ is a positive quadratic in the variables $\tau(x)$ and $\varphi(x)$, this extremum is a minimum for both $\tau(x)$ and $\varphi(x)$, which completes the proof.

□

Note that, in the above proof, we introduced the sets $A_T = \{w \in E \mid \tau(\chi_A(w)) = 1\}$ and $A_F = \{w \in E \mid \varphi(\chi_A(w)) = 1\}$, which are comparable to the sets A_T and A_F defined in 3.4.1. The difference is that we now let A_T denote the set of all worlds in E where A is true, which includes those worlds where A is both true and false. Likewise, A_F is the set of all worlds in E where A is false, which also includes the worlds where A is both true and false. Thus, A_T and A_F are not mutually exclusive. The set A_B of 3.4.1 is abolished, and all worlds where A is both true and false are contained by A_T as well as by A_F . It is the case that$E = A_T \cup A_F$.

Next, we would like to prove the following analogue of Theorem 1:

Theorem 6. Suppose $b = \begin{bmatrix} \tau(b) \\ \varphi(b) \end{bmatrix}$ is a credence function, $E \subseteq W$, $\sum_{w \in E} \tau(b(\{w\})) \neq 0$, $\sum_{w \in E} \varphi(b(\{w\})) \neq 0$ and I is a quadratic local inaccuracy measure. Then the following two propositions are equivalent:

(i) For all $A \subseteq W$ and $x \in [0,1] \times [0,1]$,

$$\sum_{w \in E} b(\{w\}) I(A, w, b'(A)) \leq \sum_{w \in E} b(\{w\}) I(A, w, x)$$

(b) Credence function b is a probability function.

What does it mean for b to be a probability function in the suggested scenario? One way to understand our two-dimensional approach to truth values and credences is as a combination of two separate systems: logical truth splits up into a logic of truth and a logic of falsity, and belief consists of a calculus of belief in propositional truth, and a separate calculus of belief in propositional falsity. We are working with LP, so logical truth and falsity are not complementary, which they were in the classical case, but may also occur together. Indeed, the LP semantics seem to invite us to discuss truth and falsity independently, which is nicely facilitated by the geometry set up in this suggestion. Thus, if we want to interpret what it means for b to be a probability function, it seems key to independently approach belief in truth and belief in falsity.

If we treat truth and falsity independently and recall the probability axiomatisations of section 2.1, we realise that b is a probability function if and only if both $\tau(b)$ and $\varphi(b)$ are probability functions. By virtue of the constructed sets A_T and A_F, it seems possible to translate the Kolmogorov axiomatisation to our case, yet, our application of the axioms has to be two-dimensional in order to fit our independent treatment of (expected) truth and falsity. In other words, rather than using the original axioms in terms of a one-dimensional credence b, we have to reformulate the axioms in terms of the component credences $\tau(b)$ and $\varphi(b)$. Understanding that together, $\tau(b)$ and $\varphi(b)$ constitute two-dimensional credences b, this will give us a two-dimensional analogue of (K1)-(K3) as mentioned in 1.1.

We use the following axiomatisation:

$(K1_\tau)$ $\tau(b(A)) \geq 0$ for all $A \subseteq W$

$(K2_\tau)$ $\tau(b(W)) = 1$

$(K3_\tau)$ $\tau(b(A \cup B)) = \tau(b(A)) + \tau(b(B))$ for disjoint $A, B \subseteq W$

$(K1_\varphi)$ $\varphi(b(A)) \geq 0$ for all $A \subseteq W$

$(K2_\varphi)$ $\varphi(b(W)) = 1$

$(K3_\varphi)$ $\varphi(b(A \cup B)) = \varphi(b(A)) + \varphi(b(B))$ for disjoint $A, B \subseteq W$

$(K1_\tau)$ - $(K3_\tau)$ are clear, as these equal (K1)-(K3) for traditional, scalar credences that only express belief in propositional truth. As we want two-dimensional credences b to behave probabilistically, we want both $\tau(b)$ and $\varphi(b)$ to be probability functions, which is why we also formulate $(K1_\varphi)$-$(K3_\varphi)$ for $\varphi(b)$ according to Kolmogorov.

We can now continue, and prove Theorem 6. By Lemma 5, it suffices to show that, for b a credence function and $E \subseteq W$, with $\sum_{w\in E} \tau(b(\{w\})) \neq 0$ and $\sum_{w\in E} \varphi(b(\{w\})) \neq 0$,

$$b(A) = \begin{bmatrix} \frac{\sum_{w\in A_T \cap E} \tau(b(\{w\}))}{\sum_{w\in E} \tau(b(\{w\}))} \\ \frac{\sum_{w\in A_F \cap E} \varphi(b(\{w\}))}{\sum_{w\in E} \varphi(b(\{w\}))} \end{bmatrix}$$

if and only if b is a probability function on the power set of W and $b(\{w\}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for $w \notin E$.

Proof. We start by proving the 'if' direction. We have to show that, if b is a probability function and $b(\{w\}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for $w \notin E$, then for all $A \subseteq W$,

$$b(A) = \begin{bmatrix} \frac{\sum_{w\in A_T \cap E} \tau(b(\{w\}))}{\sum_{w\in E} \tau(b(\{w\}))} \\ \frac{\sum_{w\in A_F \cap E} \varphi(b(\{w\}))}{\sum_{w\in E} \varphi(b(\{w\}))} \end{bmatrix}.$$

If b is a probability function and $b(\{w\}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for $w \notin E$, then

$$1 = \tau(b(\{w\})) = \sum_{w\in E} \tau(b(\{w\}))$$

and

$$1 = \varphi(b(\{w\})) = \sum_{w\in E} \varphi(b(\{w\})).$$

Hence:

$$\begin{aligned} b(A) &= \begin{bmatrix} \tau(b(A) \\ \varphi(b(A)) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{w\in A_T} \tau(b(\{w\})) \\ \sum_{w\in A_F} \varphi(b(\{w\})) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{w\in A_T \cap E} \tau(b(\{w\})) \\ \sum_{w\in A_F \cap E} \varphi(b(\{w\})) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\sum_{w\in A_T \cap E} \tau(b(\{w\}))}{\sum_{w\in E} \tau(b(\{w\}))} \\ \frac{\sum_{w\in A_F \cap E} \varphi(b(\{w\}))}{\sum_{w\in E} \varphi(b(\{w\}))} \end{bmatrix}, \end{aligned}$$

as required.
For the 'only if' direction, we have to show that, if b is a credence function and, for all $A \subseteq W$,

$$b(A) = \begin{bmatrix} \frac{\sum_{w \in A_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} \\ \frac{\sum_{w \in A_F \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} \end{bmatrix},$$

then it follows that b satisfies the axioms $(K1_\tau)$ - $(K3_\tau)$ and $(K1_\varphi)$-$(K3_\varphi)$.
We suppose that b is a credence function and that, for all $A \subseteq W$,

$$b(A) = \begin{bmatrix} \frac{\sum_{w \in A_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} \\ \frac{\sum_{w \in A_F \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} \end{bmatrix}.$$

Then:

$(K1_\tau)$ $\tau(b(A)) \geq 0$ for all $A \subseteq W$, because

$$b(A) = \begin{bmatrix} \tau(b(A)) \\ \varphi(b(A)) \end{bmatrix} : P(W) \to [0,1] \times [0,1]$$

$(K2_\tau)$ $\tau(b(W)) = \frac{\sum_{w \in W_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} = \frac{\sum_{w \in E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} = 1$

$(K3_\tau)$ If $A, B \subseteq W$ are disjoint, then

$$\begin{aligned} \tau(b(A \cup B)) &= \frac{\sum_{w \in (A \cup B)_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} \\ &= \frac{\sum_{w \in A_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} + \frac{\sum_{w \in B_T \cap E} \tau(b(\{w\}))}{\sum_{w \in E} \tau(b(\{w\}))} \\ \tau(b(A)) + \tau(b(B)), \end{aligned}$$

since $(A \cap B)_T \cup E = (A_T \cap E) \cup (B_T \cap E)$

$(K1_\varphi)$ $\varphi(b(A)) \geq 0$ for all $A \subseteq W$, because

$$b(A) = \begin{bmatrix} \tau(b(A)) \\ \varphi(b(A)) \end{bmatrix} : P(W) \to [0,1] \times [0,1]$$

$(K2_\varphi)$ $\varphi(b(W)) = \frac{\sum_{w \in W_T \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} = \frac{\sum_{w \in E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} = 1$

$(K3_\varphi)$ If $A, B \subseteq W$ are disjoint, then

$$\begin{aligned} \varphi(b(A \cup B)) &= \frac{\sum_{w \in (A \cup B)_T \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} \\ &= \frac{\sum_{w \in A_T \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} + \frac{\sum_{w \in B_T \cap E} \varphi(b(\{w\}))}{\sum_{w \in E} \varphi(b(\{w\}))} \\ \varphi(b(A)) + \varphi(b(B)), \end{aligned}$$

since $(A \cap B)_T \cup E = (A_T \cap E) \cup (B_T \cap E)$

We have thus shown that both $\tau(b)$ and $\varphi(b)$ satisfy the probability axioms. Hence, we conclude that b is a probability function, according to the understanding developed in the above. This completes the proof of our version of Theorem 2, by which we have justified probabilism for LP.

□

The structure of our argument is based on Leitgeb and Pettigrew's, but we manipulated the geometry in order to accommodate the LP semantics without requiring a Williams style mapping. We would like to check whether the interactions between truth and probability cause any trouble in this new framework, by assessing the problems identified in section 2.3.

First, it must be emphasised that the proposed geometry assumes no genuinely conceptual novelties. The philosophical interpretation of credences as expected or approximated truth values is as it was before, the only difference being that we now express them as two-dimensional vectors. Expected propositional truth and expected propositional falsity are inserted as the respective entries to these vectors, and by the same mechanics we replace logical truth by a vector comprising an expression of logical truth and one of logical falsity. It is clear how the adapted accuracy geometry should cause no semantic confusion.

Secondly, the status of logical truth is not endangered. As pointed out at the end of our discussion of suggestion 3.4.1, the reason for developing our argument was the desire to establish probabilism for the relevant semantics of LP, together with an awareness of the apparent inadequacy of Williams' mapping for LP. The construction of a framework with a form of accuracy that does justice to the LP perspective on truth is the most fundamental motivation and legitimisation of our suggestion, so from the beginning onwards we made sure to safeguard this status.

Finally, we neither encounter any problematic interactions between the compositionality of LP and the non-compositionality of probabilistic credences, nor do we run into trouble concerning semantic realism. In section 3.4 we already observed how these problems are prevented by the general set-up of epistemic utility theory. In brief, it turns out that the relation between logical (LP) truth and probabilistic credence creates no complications.

4.4.3 Measuring accuracy

As an indicative, completely different suggestion, we could assess the potential of defining and applying an accuracy measure rather than an inaccuracy measure. We only make brief notice of this option and will not explore it in technical detail, but point out that the implementation of an accuracy measure would drastically impact the mathematical form of an epistemic utility project, and that the involved changes may also enable us to derive new results.

Adhering to the understanding of credences as expected truth values, we could, for instance, approach accuracy as the agreement between credence and truth, expressed by a rational of the form $b(A)/\chi_A(w)$. We then have to impose the condition $\chi_A(w) \neq 0$, which would probably require us to compose another mapping for truth values. As accuracy is a somewhat more intuitive notion than inaccuracy, it would be interesting to investigate the merits of its possible measures in epistemic utility projects. We do, however, remember that Leitgeb and Pettigrew already touched upon the possible use of an accuracy measure, and argued against it:

> Why not maximise your accuracy instead [of minimising inaccuracy]? Because we like to think of inaccuracy as being given by a distance: the lesser the distance from the truth, the lesser the inaccuracy; the greater the distance from the truth, the greater the inaccuracy. Since distances from the truth are bounded from below, that is, by zero distance, they can be minimised, and that is what we are asking for in [the norm

> of] Accuracy. Using some means of transformation, these properties might perhaps be captured just as well by accuracy or ‘inverse distance’, but employing the notion of inaccuracy directly seems to give us a much more appealing way of stating our central epistemic goal. ([LP10a], 202-203)

Yet, our suggested rational expression for accuracy is not exactly an ‘inverse’ version of inaccuracy as we know it, nor is it even a distance. It is a fraction, a proportion, which would come with its own geometry, and its own conclusions. The interactions between truth and probability would have to be reassessed for this new framework, which would require us to have a more explicit understanding of its mathematical and technical properties.

As intriguing as it would be to study the power of such an accuracy measure, we must admit that there is a major drawback to this option, which is the definite departure that it implies from Leitgeb and Pettigrew’s argument. In the first two suggestions we could still rely on their well-grounded theoretical apparatus, whereas pursuing this third suggestion would require a completely new system, which may produce new insights, but which may also have little or no added value. We would have to find an altogether new method for establishing probabilism. Hence, we would first like to concentrate on the suggestions of 3.4.1 and 3.4.2.

5 Conclusion

Having finished our analysis of epistemic utility theory in terms of the relation between logical truth and probability, as well as our case study of LP, we can now revisit and assess the issues that motivated the paper. Let us recall how we compared and contrasted logic and probability theory in the first section, and identified a few general issues that a system combining both fields might cause. We pointed out how logical and probabilistic semantics might conflict, how the status of logical truth could be affected by the introduction of probabilities, how the interactions between compositional and non-compositional entities might be problematic, and how semantic realism could implicitly be endorsed.

As the issues mentioned take a form that differs from system to system, we studied epistemic utility and assessed how this particular theory deals with the problems concerned. Approaching epistemic utility from Leitgeb and Pettigrew's accuracy based reasoning, we showed how the issues are defused by the given conceptual set-up, and especially by the interpretation of probabilistic credences as expected truth values. More specifically, we saw that the theory's technical structure, and in particular its perspective on accuracy, entails mathematical interactions between both notions that rely on the legitimacy of Euclidean distance measures, which in turn require a quantitative analogy between truth and credence. It is essential for the Leitgeb and Pettigrew argument to hold that this Euclidean geometry is accepted, for otherwise the fundamental definition of gradational accuracy would break down. This, however, does implicitly require that the logical semantics be classical.

Moving beyond the assumption of classical logic, we explored the generalisation of Leitgeb and Pettigrew's argument to non-classical semantics by means of Williams' notion of generalised gradational accuracy. For a number of non-classical semantics, Williams defined mappings that resituate truth values on the same numerical scale as credences, thus reconstructing the Euclidean geometry where gradational accuracy is defined. We ran the argument according to this generalised form of gradational accuracy and found that, for Kleene logics, LP, supervaluationism and intuitionism, this method allows us to establish probabilism. However, this does require that we subscribe to the relevant Williams mappings, which involve some rather rough manipulations to the logical semantics for the sake of a reconstruction of the familiar Euclidean framework.

We investigated such manipulations more closely in our case study of LP. A critical evaluation of the Williams mapping for this logic taught us that it damages the vital, distinguishing aspect of the semantics, namely its dialetheism. By sending the two distinct designated truth values 1 and both to a single point on the real axis, Williams' proposed mapping disregards the basic differences between these values, and damages the LP semantics. For this reason, we abandoned the Williams mapping for LP, which also meant that we lost our justification of probabilism for this logic, and that the wanted relation between logic and probability broke down.

Aiming to re-establish probabilism, we proposed a few alternative solutions. Our first suggestion involved the design of a mapping that is based in the same reasoning as Williams' mappings, but that distinguishes between LP's two designated truth values, thus preserving dialetheism. We were able to integrate this new mapping in an adapted version of the Leitgeb and Pettigrew argument, and derive probabilism. As a second suggestion, we advanced a treatment of truth values and credences as vectors in a two-dimensional Euclidean plane. Interpreting probability functions as two-dimensionally structured as well, this method allowed the inference of probabilism too. A third, indicative suggestion was to introduce an accuracy measure to replace the conventional inaccuracy measure. This would radically change the geometry, so that an altogether new argument would have to be developed.

Comparing both suggestions, an advantage of the first one is that it maintains the geometry of Leitgeb and Pettigrew's original argument, while the second one requires twice as many dimensions. On the other hand, the first proposal depends on a mapping for truth values, whereas the second one directly includes the actual truth values. Either way, both options offer a framework that facilitates a Euclidean treatment of accuracy, and together with the unaffected understanding of credences as expected truth values, we saw that this resolves the identified issues of interacting logic and probability. Both conceptually and technically, the basis of the arguments supports peaceful and unambiguous dynamics. Yet, let us not lose sight of the machinery that had to be developed in order to show that the dynamics between logic and probability take place at all. Our suggested solutions make their own assumptions, which may profit from further, critical investigation.

Contextualising our case study of LP, we wonder whether comparable results can be obtained for other non-classical logics. What immediately comes to mind is how LP's use of a truth glut finds a direct parallel in the use of a truth gap in strong Kleene logic. Actually, we could formulate objections to Williams' proposed mapping for strong Kleene logic that are very similar to the ones we used against his LP mapping. We thus expect that suggestions comparable to the ones that were proposed for LP can also be developed for strong Kleene logic, and presumably for other three-valued logics as well. Moreover, as we saw that the use of Williams' mappings allows the justification of probabilism for a range of non-classical logics, we could consider repeating the strategy of our first suggestion for other logics. That is, we could manipulate the mappings in such a way that they do respect the relevant semantics, and use these revised versions in a generalised Leitgeb and Pettigrew argument. Of course, we could also try to find differently structured generalisations, but from our study it seems clear that a well-defined, Euclidean notion of accuracy, together with the interpretation of credences as approximated truth values, is required for the resulting interactions between truth and probability to be unproblematic.

Concluding on epistemic utility, it seems that accuracy based probabilism arguments should present no challenges to the dynamics between logic and probability, provided that we (a) construct a Euclidean geometry where accuracy is well-defined in terms of the distance between truth and credence, and (b) adhere to the interpretation of credences as expected truth values. For all discussed probabilism arguments, we observed that these two factors guarantee the resolution of any trouble that these dynamics may cause. Thus, even though we should stress that the presuppositions supporting our suggestions may well be reassessed from a different conceptual or philosophical perspective, we seem to have identified the issues that are most critical to the interaction between truth and probability in epistemic utility.

Finally, we want to relate our findings for epistemic utility to other systems that lie at the interface of logic and probability theory, but as we can imagine a great variety of such systems, we will discuss this as generally as possible. Obviously, the particular technical and mathematical circumstances that we encountered in epistemic utility theory will probably be relevant to a rather limited class of systems. However, while investigating the accuracy based Leitgeb and Pettigrew argument and its suggested adaptations, and especially while discussing these in terms of our general concerns, we did not only analyse the relation between logic and probability as instantiated in epistemic utility, but also developed an intuition of certain considerations that seem relevant to any system combining both fields. The most important such considerations are inspired by the conditions (a) and (b) of the preceding paragraph. Rephrasing (a) in more general terms, our continuous alertness with respect to legitimate accuracy measures made us aware of the importance of a clear, formal specification of the technical or mathematical relation between truth and probability. Combining both in a single system requires a strict definition of the adequate geometry, and an unambiguous demarcation of their respective statuses. Otherwise, we risk difficulties such as conflicting logical compositionality and probabilistic non-compositionality. The role of (b) convinces us that it is necessary to critically address and coordinate the semantic status of truth and probability before proposing a combined system, because otherwise semantic incoherence and confusion may occur. This was clearly illustrated by the geometry of epistemic utility, where probabilistic credence had to be interpreted as expected truth in order to semantically legitimise the accuracy definition. Hence, we generally wish to guarantee a conceptually justified form of agreement between logical and probabilistic semantics.

Evidently, any system that includes both logic and probability will come with its own theoretical characteristics and issues. Although such issues are best studied in their particular contexts, we hope that our discussion of epistemic utility has clarified the interactions and challenges that a system of combined logic and probability is likely to present, and that our conclusions may prove useful as a starting point for the investigation of other such systems in terms of the problems that we identified.

6 Bibliography

[Fin31] Bruno de Finetti. “Sul significato soggettivo della probabilita.” In: *Fundamenta Mathematicae* 17 (1931), pp. 298–329.

[Gib07] Allan Gibbard. “Rational credence and the value of truth”. In: *Oxford studies in epistemology* 2 (2007), pp. 143–164.

[Haj01] Alan Hajek. “Probability, logic, and probability logic”. In: *The Blackwell guide to philosophical logic* (2001), pp. 362–384.

[Joy98] James M Joyce. “A nonpragmatic vindication of probabilism”. In: *Philosophy of science* 65.4 (1998), pp. 575–603.

[Joy09] James M Joyce. “Accuracy and coherence: Prospects for an alethic epistemology of partial belief”. In: *Degrees of belief* (2009), pp. 263–297.

[LP10a] Hannes Leitgeb and Richard Pettigrew. “An objective justification of Bayesianism I: Measuring inaccuracy”. In: *Philosophy of Science* 77.2 (2010), pp. 201–235.

[LP10b] Hannes Leitgeb and Richard Pettigrew. “An objective justification of Bayesianism II: The consequences of minimizing inaccuracy”. In: *Philosophy of Science* 77.2 (2010), pp. 236–272.

[Mah02] Patrick Maher. “Joyce’s argument for probabilism”. In: *Philosophy of Science* 69.1 (2002), pp. 73–81.

[Mil74] David Miller. “Popper’s qualitative theory of verisimilitude”. In: *The British Journal for the Philosophy of Science* 25.2 (1974), pp. 166–177.

[Par01] Jeff B Paris. “A note on the Dutch Book method.” In: *ISIPTA*. Vol. 1. 2001.

[Pet12] Richard Pettigrew. “An improper introduction to epistemic utility theory”. In: *EPSA Philosophy of Science: Amsterdam 2009*. Springer, 2012, pp. 287–301.

[Pet+12] Richard Pettigrew et al. *‘Round Table on Coherence’*. July 2012.

[Pri79] Graham Priest. “The logic of paradox”. In: *Journal of Philosophical logic* 8.1 (1979), pp. 219–241.

[Ram31] Frank P Ramsey. “Truth and probability (1926)”. In: *The foundations of mathematics and other logical essays* (1931), pp. 156–198.

[Sch00] Stephen Schiffer. “Vagueness and partial belief”. In: *Philosophical Issues* 10.1 (2000), pp. 220–257.

[Smi10] Nicholas JJ. Smith. “Vagueness and degrees of truth. Oxford University Press, Oxford, 2008, viii+ 341 pp.” In: *Bulletin of Symbolic Logic* 16.4 (2010), pp. 533–535.

[Sza10] Zoltan Gendler Szabo. “Compositionality”. In: *Stanford encyclopedia of philosophy* (2010).

[Wil12] J Robert G Williams. “Gradational accuracy and nonclassical semantics”. In: *The Review of Symbolic Logic* 5.4 (2012), pp. 513–537.