A Note on the Temporal Sure Preference Principle and the Updating of Lower Previsions

Matthias C. M. Troffaes
Durham University, UK
matthias.troffaes@gmail.com

Michael Goldstein
Durham University, UK
michael.goldstein@durham.ac.uk

Abstract

This paper reviews the temporal sure preference principle as a basis for inference over time. We reformulate the principle in terms of desirability, and explore its implications for lower previsions. We report some initial results. Specifically, we present a simple condition for consistency of the temporal sure preference principle with any given collection of assessments, and we derive various bounds on the natural extension. We also discuss some of the technical difficulties encountered.

Keywords. updating, inference, temporal coherence, desirability, lower prevision

1 Introduction

Probabilistic inference has two components, one static and one dynamic. The static component is a description of probabilistic judgements now, where we are free to make any allocations of uncertainty that we consider to be appropriate, expressed, for example, through buying and selling prices on appropriate gambles, subject only to the constraints imposed by coherence over the collection of uncertainty judgements, precise or imprecise, that we choose now to make. The dynamic component describes how these uncertainty statements may change over time, as we receive further information, reflect further on the information that is currently available to us, and so forth.

Aspects of the dynamic component are expressed within the static component, for example through conditioning statements, which express our current buying and selling prices given various called-off bets which describe conditions under which the bets will or will not take place. Such conditioning is informative for our future judgements, but does not determine them, partly as our future experiences will not be summarisable as the observation of membership of a partition that we could specify in advance of our inferences, partly because we are always free to reflect further on the information that we have already received and change our judgements to those that we feel are in closer accord with the prior evidence, and partly because, in any case, there is nothing in the usual probabilistic formalism that forces an equivalence between current views on certain called-off bets, and actual future uncertainty assessments about the relevant quantities. This should not be seen as a failure of conditional reasoning itself—indeed, conditional reasoning is still a perfectly valid and extremely useful formalism for embedding the dynamic features of inference strictly within our current static judgements as to how such an inference might proceed.

At this point, perhaps we should note that one might indeed not care about modelling future beliefs, and take the stance that all future decisions are fully determined solely by current beliefs about those random variables that affect these decisions. For example, normal form decision making is precisely concerned with such scenario: if a subject makes all future decisions right now, only his current beliefs count, and his future beliefs are completely irrelevant. In practice however, beliefs are revised over time, and it is rarely the case that future beliefs, which will determine future decisions, are determined solely on the basis of called-off bets with respect to current beliefs, say through repeated application of Bayes theorem. Analyzing our current beliefs about our future beliefs, as in this paper, is thus important if we now wish to know how we will act in the future based on the actual, but now still uncertain, beliefs that we will hold in the future.

Temporal coherence is concerned with the careful description of the relationships between the static and dynamic features of probabilistic reasoning. We do not know what our future uncertainty judgements will be, but we may now express views about them. These views are, themselves, probabilistic. The basic questions that we must ask are:

(i) Are there any constraints that are reasonable to impose on our current judgements about our future judgements?
(ii) How may such constraints be exploited within the general approach to inference?
(iii) How does the conventional approach to probabilistic

reasoning, via conditioning, fit into the actual temporal evolution of beliefs?

This paper is a modest initial exploration of a particular development of such temporal reasoning, based on the work of [3, 4] and summarised in [5]. In particular, we discuss some of the implications of temporal reasoning for inference with coherent lower previsions. We will only explore questions (i) and (ii).

The key concept in studying temporal coherence is the so-called *temporal sure preference principle*, which establishes a link between certain future preferences and current preferences, thereby allowing us to say something now about our future beliefs.

In imprecise probability theory, preferences come about as a very natural way of modelling beliefs, and it has been argued that the concept of desirability, that is, which gambles we (possibly marginally) prefer to the zero gamble, forms one of the most elegant mathematical and philosophical foundations for imprecise probability [10, 11, 9].

The traditional way of looking at updating in the subjective approach to imprecise probability goes by means of conditioning, that is, looking at called-off gambles. For instance, very recently, Zaffalon and Miranda [12] provided a justification for conditioning and conglomerability, through temporal reasoning, in a setting where future beliefs are assumed to be fixed now.

However, in practice, future subjective beliefs rarely reflect past called-off gambles, and in fact there is no compelling reason for this to be so, simply because there is no compelling reason for them to be fixed now. Indeed, it seems far more natural to start out from the premise that future beliefs are inherently random, which leads to a more general theory, but of course we also risk it to be far less tractable—interestingly, in the precise case, the generality gained leads to updating rules which are far more efficient than computing with called-off gambles, particularly for large scale problems (for instance, see [1]). Having preference, in the form of desirability, at its foundations, imprecise probability is a natural candidate for temporal coherence. We hope it might lead us, as in the precise case, to say something meaningful now about future beliefs, in a way that updating is more flexible, more realistic, and potentially also numerically easier, than the traditional called-off gamble approach (i.e. the generalized Bayes rule [9, Sec. 6.4]).

This paper is organised as follows. Section 2 briefly summarises the main results that we need for lower previsions and desirability. Section 3 reviews temporal coherence and its main implications for previsions. Section 4 explores an approach to temporal coherence for lower previsions. We conclude in Section 5.

2 Lower Previsions and Desirability

Let Ω denote a possibility space. A *gamble* is simply a bounded random quantity, and is mathematically represented by a real-valued function on Ω. We will denote gambles by capital letters X, Y, $\dots$. The set of all gambles on Ω is denoted by $\mathcal{L}(\Omega)$. The set of all gambles on Ω that are constant on elements of some partition $\mathcal{A}$ is denoted by $\mathcal{L}(\mathcal{A})$.

As mentioned in the introduction, we will take desirability to be the basic concept, and will use it for studying the implications of temporal coherence on lower previsions. To keep the treatment as simple as possible, however, we will restrict ourselves to sets of almost-desirable gambles induced by lower previsions.

The following serves to fix the notation and conventions used in the paper. It is assumed that the reader is familiar with lower previsions and desirability. We refer to [9] for much more information on the topic. In particular, throughout the paper, we will use the properties of coherent lower previsions extensively [9, Sec. 2.6.1].

Specifically, let $\underline{\mathrm{E}}$ be a coherent lower prevision on $\mathcal{L}(\Omega)$ (without loss of generality, through natural extension [9, Sec. 3.1]), that is, $\underline{\mathrm{E}}$ satisfies:

C1 $\underline{\mathrm{E}}(X) \geq \inf X$

C2 $\underline{\mathrm{E}}(X+Y) \geq \underline{\mathrm{E}}(X) + \underline{\mathrm{E}}(Y)$

C3 $\underline{\mathrm{E}}(\lambda X) = \lambda \underline{\mathrm{E}}(X)$

for all X, $Y \in \mathcal{L}(\Omega)$ and all $\lambda \geq 0$. The upper prevision $\overline{\mathrm{E}}$ corresponding to $\underline{\mathrm{E}}$ is defined as:

$$\overline{\mathrm{E}}(X) = -\underline{\mathrm{E}}(-X). \tag{1}$$

By $\mathfrak{P}(\Omega)$ we denote the set of all coherent lower previsions on $\mathcal{L}(\Omega)$.

With $\underline{\mathrm{E}}$ we can then associate a *set of (almost) desirable gambles*:

$$\mathcal{D} := \{X \in \mathcal{L} \colon \underline{\mathrm{E}}(X) \geq 0\}. \tag{2}$$

For simplicity of exposition, when in the following we say desirable, we really mean almost-desirable. The following conditions are satisfied:

D1 if $X \geq 0$ then $X \in \mathcal{D}$,

D2 if $\sup X < 0$ then $X \notin \mathcal{D}$,

D3 if $X \in \mathcal{D}$ and $Y \in \mathcal{D}$ then $X + Y \in \mathcal{D}$,

D4 if $\lambda \geq 0$ and $X \in \mathcal{D}$ then $\lambda X \in \mathcal{D}$, and

D5 if $X + \epsilon \in \mathcal{D}$ for all $\epsilon > 0$, then $X \in \mathcal{D}$.

Note that we can recover $\underline{\mathrm{E}}$ from $\mathcal{D}$ through:

$$\underline{\mathrm{E}}(X) = \sup\{a \in \mathbb{R} \colon X - a \in \mathcal{D}\} \tag{3}$$

so in the following, we can use $\underline{\mathrm{E}}$ and $\mathcal{D}$ interchangeably.

A lower prevision is called a prevision when it is self-conjugate, that is, when $\underline{\mathrm{E}} = \overline{\mathrm{E}}$, in which case we simply denote it by E. It is well known that previsions correspond to expectation operators, and lower previsions correspond to lower envelopes of expectation operators.

We will consider lower previsions at different points in time—in fact, at just two points in time, 0 and $t > 0$.

By Ω we denote the possibility space at time 0: it represents our subjective judgement, now, about what events are possible. Because Ω will include events involving future beliefs, which we do not know, we emphasize that, in general, a full specification of Ω is not possible.

We assume that we are able to specify a partition $\mathcal{A}$ of Ω which generates all events relevant to the problem domain at hand. Unlike Ω, the partition $\mathcal{A}$ is explicitly modelled, and hence, represents the operational part of Ω. We assume that our current assessments about the problem domain only involve gambles that are constant on the elements of $\mathcal{A}$, and we need not make any further assessments about any other gambles, that is, we can specify a lower prevision $\underline{P}^{\mathcal{A}}$ defined on some subset of $\mathcal{L}(\mathcal{A})$. In particular, we need not make any direct assessments about our future beliefs—those will come in later through the temporal sure preference principle.

We also assume the existence of a partition $\mathcal{B}_t$ of Ω, such that exactly one of the elements of this partition will occur at time t. We will not make any assumption about $\mathcal{B}_t$, in fact, operationally, it is usually impossible identify now what $\mathcal{B}_t$ ought to be. One could consider an element of $\mathcal{B}_t$ to be a possible possibility space at time t, thus elements of $\mathcal{B}_t$ will be denoted by Ω_t. For any $\omega \in \Omega$, by $[\omega]_t$ we denote the unique element Ω_t of $\mathcal{B}_t$ that contains ω. Perhaps we need to emphasize that we do not assume any relationship between $\mathcal{A}$ and $\mathcal{B}_t$. In particular, we do not assume that, say, $\mathcal{B}_t$ refines $\mathcal{A}$: this would mean that, at time t, we would know which element A of $\mathcal{A}$ obtains, and generally, of course this will not be the case.

By $\underline{\mathrm{E}}_t$ we denote our coherent lower prevision at time t—its value is known to us at time t. So, $\underline{\mathrm{E}}_0$ is our current lower prevision, and embodies both our current assessments $\underline{P}^{\mathcal{A}}$ concerning the problem domain at hand, as well as any further principles taken into account, such as for instance the temporal sure preference principle, which we will discuss in detail later. However, $\underline{\mathrm{E}}_t$ is in fact a random lower prevision now:[1]

$$\underline{\mathrm{E}}_t(\Omega_t) \in \mathfrak{P}(\Omega_t) \text{ for any } \Omega_t \in \mathcal{B}_t, \tag{4}$$

whose value is only realised at time t.

When comparing gambles, as we will need to do further in the paper, it is convenient that those gambles are expressed with respect to the same possibility space. For this reason, it is more convenient to consider $\underline{\mathrm{E}}_t$ as a mapping from Ω to $\mathfrak{P}(\Omega)$:

$$\underline{\mathrm{E}}_t(\omega)(X) \coloneqq \underline{\mathrm{E}}_t([\omega]_t)(X|_{[\omega]_t}), \tag{5}$$

for any $\omega \in \Omega$ and $X \in \mathcal{L}(\Omega)$. We will follow this convenient notation for the remainder of the paper. Note that one may think of $\mathcal{B}_t$ as the partition generated by $\underline{\mathrm{E}}_t$.

For any gamble $X \in \mathcal{L}(\Omega)$, by $\underline{\mathrm{E}}_t(X)$ we denote the random lower prevision of X at time t:

$$\underline{\mathrm{E}}_t(X)(\omega) \coloneqq \underline{\mathrm{E}}_t(\omega)(X). \tag{6}$$

Clearly, $\underline{\mathrm{E}}_t(X) \in \mathcal{L}(\Omega)$, and it is constant on the elements of $\mathcal{B}_t$.

Similarly, we write $\mathcal{D}_t$ for the set of desirable gambles corresponding to $\underline{\mathrm{E}}_t$. So, $\mathcal{D}_t$ is a random set of gambles:

$$\mathcal{D}_t \colon \Omega \to \wp(\mathcal{L}(\Omega)) \tag{7}$$

where

$$\mathcal{D}_t(\omega) \coloneqq \{X \in \mathcal{L}(\Omega) \colon \underline{\mathrm{E}}_t(\omega)(X) \geq 0\}, \tag{8}$$

and as with $\underline{\mathrm{E}}_t$, the value of $\mathcal{D}_t$ is only realised at time t.[2] Clearly, $\mathcal{D}_t$ is constant on the elements of $\mathcal{B}_t$.

3 Temporal Coherence for Previsions

In this section, we review the existing theory of temporal coherence for previsions.

3.1 Beliefs and Updating

By X, we denote a gamble whose value is unknown to us. Of course, we may have present beliefs about X. We assume that X is constant on the elements of the partition $\mathcal{A}$. Our present expectation for X is denoted by $\mathrm{E}_0(X)$, and our present variance for X is denoted by $\mathrm{var}_0(X)$. The subscript in E_0 and var_0 denotes time, where time 0 corresponds to the present.

As X is unknown, we may try to learn about X by observing another random quantity, which we denote by Y: say we actually observe the value of Y at time $t > 0$, whilst X remains unknown to us at time t. Again, we assume that Y is constant on the elements of the partition $\mathcal{A}$—because

[1] Remember that $\mathfrak{P}(\Omega_t)$ denotes the set of all coherent lower previsions on $\mathcal{L}(\Omega_t)$.

[2] We should note that $\mathcal{D}_t(\omega)$, when defined as a subset of $\mathcal{L}(\Omega)$ as in Eq. (8), may not satisfy D5, however of course $\mathcal{D}_t(\omega)$ will satisfy D5 as a subset of $\mathcal{L}([\omega]_t)$. Also note that $X \in \mathcal{D}_t(\omega)$ if and only if $I_{[\omega_t]}X \in \mathcal{D}_t(\omega)$.

we assumed that Y is known at time t, it will also be constant on the elements of $\mathcal{B}_t$. Here too, we may have present beliefs about Y, such as its present expectation $\mathrm{E}_0(Y)$ and present variance $\mathrm{var}_0(Y)$. In fact, we may hold present beliefs about X and Y jointly, such as for instance the present covariance between X and Y, which we denote by $\mathrm{cov}_0(X, Y)$.

As mentioned, whilst the value of Y will be known at time t, X remains unknown. Consequently, we may also consider, now, our future beliefs about X. However, because the future has yet to obtain, those future beliefs are uncertain in themselves. In other words, $\mathrm{E}_t(X)$ and $\mathrm{var}_t(X)$, the actual expectation and variance of X which represents our beliefs about X at time t, are gambles in themselves, whose values are only known to us at time t:

$$\mathrm{E}_t(X)\colon \Omega \to \mathbb{R}, \qquad \mathrm{var}_t(X)\colon \Omega \to \mathbb{R}. \tag{9}$$

For example, we can think about our current beliefs about our future beliefs, and could consider for instance our present expectation and variance of these gambles: $\mathrm{E}_0(\mathrm{E}_t(X))$, $\mathrm{E}_0(\mathrm{var}_t(X))$, $\mathrm{var}_0(\mathrm{E}_t(X))$, and $\mathrm{var}_0(\mathrm{var}_t(X))$.

The general problem of updating might then be concerned with answering the following questions. First, what should be the relationship between:

- our current beliefs about $\mathrm{E}_t(X)$,
- our current beliefs about X, and
- our current beliefs about Y?

More challengingly, what should be the relationship between:

- our actual beliefs $\mathrm{E}_t(X)$ about X at time t, and
- any updating rule for X as a function of Y?

3.2 The Temporal Sure Preference Principle

In order to establish relationships between current and future beliefs, we must impose conditions that go beyond coherence at a single time point. These conditions should be sufficiently weak and compelling to be widely applicable, while leading to a meaningful account of inference.

Any principle which asserts that beliefs now are compelling for beliefs in the future is, by its nature, unconvincing, as we cannot know what future information we may receive or what the outcome of our future reflections may be. The converse, however, is that we may often view our future beliefs as compelling for our current beliefs, as all such future reflections and information will be taken into account in such future judgements. In order for future judgements to influence our current judgements, we must know what such future judgements are. We therefore introduce the notion of a sure prefererence, at a future time, as one which we are now sure that we will hold at that time.

It may seem unreasonable, now, to think that we hold any such sure preferences. However, it so happens that we do indeed hold many such, and recognising them explicitly, and formalising their implications for our current judgements, provides a natural account of temporal reasoning. For this reason, Goldstein introduced the following principle (see [3], [4], [5, Sec. 3.5]):

Principle 1 (The Temporal Sure Preference Principle I)**.** *For any gambles $U \in \mathcal{L}(\Omega)$ and $W \in \mathcal{L}(\Omega)$, if you have a* ***sure*** *preference for U over W at future time t, then you should not have a strict preference for W over U now.*

It is useful to briefly reflect on what it means to have a sure preference for U over W at future time t. Remember, at future time t, an element Ω_t of $\mathcal{B}_t$ obtains, and we hold beliefs $\mathrm{E}_t(\Omega_t) \in \mathfrak{P}(\Omega_t)$—for now these beliefs are assumed to be precise. A *sure* preference means a preference regardless of the outcome Ω_t in $\mathcal{B}_t$. So in other words, we are sure to prefer U to W at time t whenever

$$\mathrm{E}_t(\Omega_t)(U|_{\Omega_t}) \geq \mathrm{E}_t(\Omega_t)(W|_{\Omega_t}) \text{ for all } \Omega_t \in \mathcal{B}_t, \tag{10}$$

or equivalently, whenever

$$\mathrm{E}_t(U)(\omega) \geq \mathrm{E}_t(W)(\omega) \text{ for all } \omega \in \Omega, \tag{11}$$

where we use the notation introduced earlier in Eqs. (5) and (6).

The temporal sure preference principle should be considered as a prescription for a particular domain of discourse, rather than as a fundamental condition for rationality. There are various reasons why, in a particular application, it might not hold. For example, we might consider that, at the future time, we could undergo personality changes which render our future judgements suspect to us now (the Doctor Jekyll and Mister Hyde scenario). More prosaically, we might just recognise situations where our future judgements are likely to be less reliable than our current judgements (for example, the problem of forgetting). Therefore, the intention of the temporal sure preference principle is that it should be viewed as a very weak, and widely applicable principle, whose relevance we should consider for the problem at hand. If we consider the temporal sure preference principle applicable in our problem, then we may draw on the strong implications of the principle to provide an account of temporal coherence for this situation. We know of no weaker alternative principle that allows a similar account of the inferential process, for the many applications where we will be willing to assert temporal sure preference.

The aim of this section is to study this principle in terms of desirability [10, 11, 9], whilst at the same time reviewing the main well-known consequences of the temporal sure preference principle for previsions, in order to provide a good understanding of the ideas and techniques involved before we move on to lower previsions in Section 4.

If we take preference $U \succeq W$ to mean that $U - W + \epsilon$ is desirable for all $\epsilon > 0$ [9, Sec. 3.7.5, first paragraph], and $U \succ W$ to mean that $U - W - \epsilon$ is desirable for some $\epsilon > 0$ [9, Sec. 3.7.7, second paragraph], then it is a trivial exercise to reformulate the above principle in terms of desirability:[3]

Principle 2 (The Temporal Sure Preference Principle II)**.** *For any gamble $U \in \mathcal{L}(\Omega)$, if, for all $\epsilon > 0$, $U + \epsilon$ is sure to be desirable for us at future time t, then, for all $\epsilon > 0$, $-U - \epsilon$ should not be desirable for us now.*

Perhaps it is useful to note already here that many variations of Principle 2 are possible. We will consider some of those variations, which are all equivalent for previsions, but which are no longer equivalent for lower previsions.

We give a quick proof of equivalence, which holds generally—not just for sets of desirable gambles corresponding to previsions, but for arbitrary sets of desirable gambles; we do not even need to rely on coherence.

Proposition 3. *Principles 1 and 2 are equivalent.*

Proof. Suppose Principle 1 is satisfied. Suppose that, for all $\epsilon > 0$, $U + \epsilon$ is sure to be desirable to us at future time t. This means that, surely, $U \succeq_t 0$ at time t. Consequently, by Principle 1, $0 \not\succ_0 U$ now, or in other words, $0 - U - \epsilon$ is not desirable now for any $\epsilon > 0$. In other words, Principle 2 is satisfied.

Conversely, suppose that Principle 2 is satisfied. Suppose that, surely, $U \succeq_t W$ at time t. This means that, for all $\epsilon > 0$, $U - W + \epsilon$ is surely desirable at time t. Consequently, by Principle 2, for all ϵ, $-U + W - \epsilon$ is not desirable now. But this means precisely that $W \not\succ_0 U$, now. In other words, Principle 1 is satisfied. □

An obvious question at this point is: what kind of gambles can be surely desirable at some future time t? Obviously, any positive constant gamble would be, but that is hardly useful, as we already know that these are desirable to us now. For more interesting examples, consider cases where U is a function of E_t. For example, at time t, surely, the gamble $\mathrm{E}_t(X) - X + \epsilon$ is desirable for all $\epsilon > 0$ (note that at time t, $\mathrm{E}_t(X)$ is a constant, whilst X is still a gamble). The temporal sure preference principle then tells us that the gamble $-\mathrm{E}_t(X) + X - \epsilon$ is not desirable to us now.

3.3 Implications

The next proposition, due to Goldstein [4, Theorem 1], forms the basis for linking future beliefs about expectation and variance to current beliefs about expectation and variance. The proof is short, and provides an excellent example of how the temporal sure preference principle can be invoked to make non-trivial statements about $\mathrm{E}_t(X)$, so we reproduce it below.

Proposition 4. *If Principle 2 is satisfied, then it must hold that*

$$\mathrm{E}_0((X - \mathrm{E}_t(X))^2) \le \mathrm{E}_0((X - Y)^2). \quad (12)$$

where Y is surely known by time t.

Proof. Note that, for previsions, $U \preceq_t W$ precisely when $\mathrm{E}_t(U)(\omega) \le \mathrm{E}_t(W)(\omega)$ for all $\omega \in \Omega$, and $U \not\succ_0 W$ precisely when $\mathrm{E}_0(U) \le \mathrm{E}_0(W)$. Also, note that $\mathrm{E}_t(U)(\omega) = U(\omega)$ for any gamble U that is constant on the elements of $\mathcal{B}_t$, such as $\mathrm{E}_t(X)$ and Y.

Consequently, for any $\omega \in \Omega$,

$$\begin{aligned} &\mathrm{E}_t((X - Y)^2 - (X - \mathrm{E}_t(X))^2)(\omega) && (13)\\ &= \mathrm{E}_t(-2XY + Y^2 + 2X\mathrm{E}_t(X) - \mathrm{E}_t(X)^2)(\omega) && (14)\\ &= -2\mathrm{E}_t(X)(\omega)Y(\omega) + Y^2(\omega) + \mathrm{E}_t(X)^2(\omega) && (15)\\ &= (\mathrm{E}_t(X)(\omega) - Y(\omega))^2 \ge 0 && (16) \end{aligned}$$

where we have used the linearity of $\mathrm{E}_t(\omega)$. So, at time t,[4]

$$(X - \mathrm{E}_t(X))^2 \preceq_t (X - Y)^2. \quad (17)$$

Whence, by Principle 2, now,

$$(X - \mathrm{E}_t(X))^2 \not\succ_0 (X - Y)^2, \quad (18)$$

which yields the desired inequality. □

Those readers familiar with the usual called-off argument for conditional previsions may fear that we have, inadvertedly, relied on conglomerability of E_0 to complete the above argument. Perhaps, it is instructive to try follow this misinterpretation to put such fears at rest. Indeed, in the proof, we first show that, effectively,

$$(X - Y)^2 - (X - \mathrm{E}_t(X))^2 \quad (19)$$

is desirable at time t. One might correctly, but confusingly, understand that this means that the called-off gamble

$$I_{\Omega_t}\left((X - Y)^2 - (X - \mathrm{E}_t(X))^2\right) \quad (20)$$

is now desirable. In fact, it is sure to be desirable at time t—if Ω_t does not obtain, then it is zero and thus desirable, and if Ω_t does obtain, then the reasoning in the proof can be used to show that it is desirable as well—thus, by the temporal sure preference, indeed, the called-off gamble defined in Eq. (20) is desirable now. Then, assuming conglomerability, we can glue all these called-off gambles together to prove that the gamble in Eq. (19) is desirable now. We simply emphasize here that the actual proof works quite differently. In particular, the temporal sure preference principle is only applied once, namely on the gamble in Eq. (19): called-off gambles are never considered.

[3] The attentive reader will note that in Principle 2, we can actually take desirability to be actual desirability, rather than almost-desirability.

[4] It is interesting to compare Eq. (17) with the operational definition of expectation of de Finetti [2], in which Eq. (17) is the definition of $\mathrm{E}_t(X)$, rather than a derived property.

The proof of Proposition 4, and the above discussion, already hint at a slightly simpler version of the temporal sure preference principle:

Principle 5 (The Temporal Sure Preference Principle III)**.** *For any gamble $U \in \mathcal{L}(\Omega)$, if U is sure to be desirable for us at future time t, then U should be desirable for us now:*

$$\bigcap_{\Omega_t \in \mathcal{B}_t} \mathcal{D}_t(\Omega_t) = \bigcap_{\omega \in \Omega} \mathcal{D}_t(\omega) \subseteq \mathcal{D}_0. \tag{21}$$

Proposition 6. *Principle 5 implies Principle 2.*

Proof. Assume that Principle 5 holds. If $U + \epsilon$ is sure to be desirable at time t, for all $\epsilon > 0$, then consequently, $U + \epsilon$ is desirable now, for all $\epsilon > 0$. If $-U - \delta$ would be desirable for us now for some $\delta > 0$, then $U + \delta/2 - U - \delta = -\delta/2$ would be desirable as well, which would lead us to incur a sure loss, so $-U - \delta$ cannot be desirable now for any $\delta > 0$. In other words, Principle 2 holds. □

Proposition 7. *If our set of desirable gambles corresponds to a prevision, that is, if*

$$\mathcal{D}_0 = \{U \colon \mathrm{E}_0(U) \geq 0\} \tag{22}$$

for some prevision E_0, then Principle 5 is equivalent to Principle 2.

Proof. Assume Principle 2 holds. If U is sure to be desirable at time t, then obviously $U + \epsilon$ is also sure to be desirable at time t, for all $\epsilon > 0$. Consequently, $-U - \epsilon$ is not desirable now, for all $\epsilon > 0$, or in other words, $\mathrm{E}_0(-U) - \epsilon < 0$ for all $\epsilon > 0$. This means that $\mathrm{E}_0(U) \geq 0$, so U is desirable to us now. □

In other words, for the remainder of this section, where we are concerned with previsions only, we can assume Principle 5 without loss of generality. We will thus assume that desirability is as in Eq. (22).

Proposition 4 has a number of very interesting consequences:

Corollary 8. *If Principle 5 is satisfied, then*

$$\mathrm{E}_0(X - \mathrm{E}_t(X)) = 0. \tag{23}$$

Proof. In Proposition 4, let $Y := \mathrm{E}_t(X) + b$ where $b \in \mathbb{R}$, and take the minimum over b. □

Note that Eq. (23) is very similar to the usual definition of conglomerability as in for instance [9, p. 305, (C15)], so it is worth emphasizing that Eq. (23) is *not* your usual conglomerability, because $\mathrm{E}_t(X)$ is not necessarily obtained through conditioning.

We can also say something about the expected future variance, that is, $\mathrm{E}_0(\mathrm{var}_t(X))$.

Corollary 9 (Adjusted Variance)**.** *If Principle 5 is satisfied, then it holds that:*

$$\mathrm{var}_0(X - \mathrm{E}_t(X)) = \mathrm{E}_0(\mathrm{var}_t(X)) \leq \mathrm{var}_Y(X), \tag{24}$$

with

$$\mathrm{var}_Y(X) := \mathrm{var}_0(X) - \frac{\mathrm{cov}_0(X, Y)^2}{\mathrm{var}_0(Y)}, \tag{25}$$

where Y is surely known by time t.

Proof. To prove the inequality in Eq. (24), take $a + bY$ for Y in Proposition 4, and minimize over a and b.

Note that the usual formulation uses $\mathrm{var}_0(X - \mathrm{E}_t(X))$ only. It is easy to see that this is $\mathrm{E}_0(\mathrm{var}_t(X))$, which seems easier to interpret, and is also relevant for what comes later:

$$\begin{aligned}\mathrm{var}_0(X - \mathrm{E}_t(X)) = \mathrm{E}_0((X - \mathrm{E}_t(X) \\ - \mathrm{E}_0(X - \mathrm{E}_t(X)))^2)\end{aligned} \tag{26}$$

and by Eq. (23) $\mathrm{E}_0(X - \mathrm{E}_t(X)) = 0$, so

$$= \mathrm{E}_0((X - \mathrm{E}_t(X))^2) \tag{27}$$

and again by Eq. (23) $\mathrm{E}_0(\cdot) = \mathrm{E}_0(\mathrm{E}_t(\cdot))$, so

$$= \mathrm{E}_0(\mathrm{E}_t((X - \mathrm{E}_t(X))^2)) \tag{28}$$
$$= \mathrm{E}_0(\mathrm{var}_t(X)). \tag{29}$$

□

So, the temporal sure preference principle allows us to quantify uncertainty about future variance.

In the proof of Corollary 9, the value for $a + bY$ where the minimum is achieved is precisely the *adjusted expectation*:

Corollary 10 (Adjusted Expectation)**.** *If Principle 5 is satisfied, then*

$$\mathrm{E}_t(X) = \mathrm{E}_Y(X) + S_t(X), \tag{30}$$

where

$$\mathrm{E}_Y(X) := \mathrm{E}_0(X) + \frac{\mathrm{cov}_0(Y, X)}{\mathrm{var}_0(Y)}(Y - \mathrm{E}_0(Y)), \tag{31}$$

and

$$\mathrm{E}_0(S_t(X)) = 0, \quad \mathrm{cov}_0(S_t(X), \mathrm{E}_Y(X)) = 0. \tag{32}$$

Proof. Take $a + bY + c\mathrm{E}_t(X)$ for Y in Proposition 4, and do the usual magic. □

In other words, the temporal sure preference principle also allows us to quantify a linear connection between observations and future beliefs.

If Y is the indicator of some event E, then $\mathrm{E}_Y(X) = E(X|E)$, that is, adjusted expectation coincides with conditional expectation. So, Eq. (30) also provides an interpretation of the relation between conditioning and our actual posterior expectation.

The above results are only an initial tasting of the realm of possibilities. Of considerable interest is that the above treatment generalises almost trivially to the multivariate case.

4 Temporal Coherence for Lower Previsions

Let us now investigate the implications of the temporal sure preference principle for lower previsions.

4.1 The Temporal Sure Preference Principle for Lower Previsions

In the context of desirability, it makes sense to adopt Principle 5, for at least two reasons:

1. The principle seems reasonably compelling. Indeed, if U is sure to be desirable for us at time t, then it does not matter whether we accept it already now, or whether we accept it only at time t: the gamble has the same outcome either way.
2. We may use it as a production rule in natural extension.

By the second point, we mean the following. As mentioned in the introduction, we assume a partition $\mathcal{A}$ which represents what we could call the operational part of Ω. Specifically, all direct assessments of lower previsions $\underline{P}_0^{\mathcal{A}}(Y)$, which represent our beliefs now, concern gambles $Y \in \mathcal{L}(\mathcal{A})$. In other words, our initial assessments are embodied by a lower prevision $\underline{P}_0^{\mathcal{A}}$ which is defined on a subset of $\mathcal{L}(\mathcal{A})$. We can then consider the natural extension of $\underline{\mathrm{E}}_0^{\mathcal{A}}$ to all gambles $\mathcal{L}(\mathcal{A})$; let us denote that natural extension by $\underline{\mathrm{E}}_0^{\mathcal{A}}$. It is different from $\underline{\mathrm{E}}_0$, which embodies our beliefs about $\underline{P}_0^{\mathcal{A}}$ but also those implied by the temporal sure preference principle. Indeed, under Principle 5, all gambles V for which

$$\underline{\mathrm{E}}_t(V)(\omega) \geq 0 \text{ for all } \omega \in \Omega, \tag{33}$$

or briefly, for which $\underline{\mathrm{E}}_t(V) \geq 0$, are desirable now. Consequently,

$$\underline{\mathrm{E}}_0(U) = \sup_{\substack{\alpha\in\mathbb{R} \\ Y\in\mathcal{L}(\mathcal{A})\colon \underline{\mathrm{E}}_0^{\mathcal{A}}(Y)\geq 0 \\ V\in\mathcal{L}(\Omega)\colon \underline{\mathrm{E}}_t(V)\geq 0}} \{\alpha\colon U - \alpha \geq Y + V\} \tag{34}$$

for any gamble $U \in \mathcal{L}(\Omega)$.

Before we proceed investigating actual inferences from the above expression for natural extension, we need to address a few concerns. First, there is no guarantee that Principle 5 is consistent with our initial assessments $\underline{P}_0^{\mathcal{A}}$. Eq. (34) provides us with a means to verify this: we merely have to check that $\underline{\mathrm{E}}_0(0) < +\infty$ [9, p. 123, ll. 4–7]. Secondly, there is no guarantee that Principle 5 does not modify $\underline{\mathrm{E}}_0^{\mathcal{A}}$ on $\mathcal{L}(\mathcal{A})$. Thirdly, this form of natural extension is inherently non-constructive: it involves an operator $\underline{\mathrm{E}}_t$ about which we have not specified much at all. The next proposition answers the first two concerns. The last concern of course remains, but nevertheless, we will show that we still can derive something non-trivial about $\underline{\mathrm{E}}_t$, just as in the precise case discussed earlier.

Proposition 11. *If, for every $A \in \mathcal{A}$, there is an $\Omega_t^A \in \mathcal{B}_t$ such that*

$$\underline{\mathrm{E}}_t(\Omega_t^A)(A) = 1, \tag{35}$$

then Principle 5 is consistent with $\underline{P}_0^{\mathcal{A}}$, and, for all $X \in \mathcal{L}(\mathcal{A})$,

$$\underline{\mathrm{E}}_0(X) = \underline{\mathrm{E}}_0^{\mathcal{A}}(X). \tag{36}$$

Proof. If we can prove Eq. (36), then consistency follows immediately.

Consider any $X \in \mathcal{L}(\mathcal{A})$. Clearly, $\underline{\mathrm{E}}_0(X) \geq \underline{\mathrm{E}}_0^{\mathcal{A}}(X)$. We now prove the converse inequality. Indeed,

$$\underline{\mathrm{E}}_0(X) = \sup_{\substack{\alpha\in\mathbb{R} \\ Y\in\mathcal{L}(\mathcal{A})\colon \underline{\mathrm{E}}_0^{\mathcal{A}}(Y)\geq 0 \\ V\in\mathcal{L}(\Omega)\colon \underline{\mathrm{E}}_t(V)\geq 0}} \{\alpha\colon X - \alpha \geq Y + V\} \tag{37}$$

$$= \sup_{\substack{\alpha\in\mathbb{R} \\ Y\in\mathcal{L}(\mathcal{A})\colon \underline{\mathrm{E}}_0^{\mathcal{A}}(Y)\geq 0 \\ V\in\mathcal{L}(\Omega)\colon \underline{\mathrm{E}}_t(V)\geq 0}} \Bigg\{\alpha\colon (\forall A \in \mathcal{A}) \tag{38}$$

$$\left(X(A) - \alpha \geq Y(A) + \sup_{\omega\in A} V(\omega)\right)\Bigg\} \tag{39}$$

so, if we can show that $\sup_{\omega\in A} V(\omega) \geq 0$ whenever $\underline{\mathrm{E}}_t(V) \geq 0$, then

$$\leq \sup_{\substack{\alpha\in\mathbb{R} \\ Y\in\mathcal{L}(\mathcal{A})\colon \underline{\mathrm{E}}_0^{\mathcal{A}}(Y)\geq 0}} \{\alpha\colon X - \alpha \geq Y\} \tag{40}$$

$$= \underline{\mathrm{E}}_0^{\mathcal{A}}(X). \tag{41}$$

We are left to show that $\sup_{\omega\in A} V(\omega) \geq 0$ whenever $\underline{\mathrm{E}}_t(V) \geq 0$. In fact, we will show that $\sup_{\omega\in A\cap\Omega_t^A} V(\omega) \geq 0$, by contraposition. Note that Eq. (35) already implies that $A \cap \Omega_t^A$ is non-empty.

Suppose that $\sup_{\omega\in A\cap\Omega_t^A} V(\omega) < 0$, then there would be an $\epsilon > 0$ such that for all $\omega \in A \cap \Omega_t^A$,

$$V(\omega) < -\epsilon. \tag{42}$$

Therefore, necessarily, also

$$\underline{\mathrm{E}}_t(\Omega_t^A)(V) \leq \underline{\mathrm{E}}_t(\Omega_t^A)(I_{A^c}V) + \overline{\mathrm{E}}_t(\Omega_t^A)(-I_A\epsilon) = -\epsilon \tag{43}$$

because $\overline{\mathrm{E}}_t(\Omega_t^A)(A^c) = 0$, so $\underline{\mathrm{E}}_t(\Omega_t^A)(I_{A^c}V) = 0$, and $\underline{\mathrm{E}}_t(\Omega_t^A)(A) = 1$, so $\overline{\mathrm{E}}_t(\Omega_t^A)(-I_A\epsilon) = -\epsilon$. But Eq. (43) contradicts the assumption that $\underline{\mathrm{E}}_t(V) \geq 0$. □

The consistency condition in Eq. (35) has a simple interpretation: for every $A \in \mathcal{A}$, we must allow for the possibility that at time t, we will be certain that A has obtained. Note that we only must logically allow for this possibility—it may well have zero probability—so the condition is really very weak.

We also immediately have the following important result, which effectively reformulates Principle 5 in terms of lower previsions:[5]

Proposition 12. *Principle 5 holds if and only if, for every gamble $U \in \mathcal{L}(\Omega)$,*

$$\inf_{\omega\in\Omega} \underline{E}_t(U)(\omega) \le \underline{E}_0(U). \tag{44}$$

Proof. "only if". Suppose Principle 5 holds. We could rely on our expression for natural extension, Eq. (34), however it is instructive to use only Principle 5 in the proof.

For any $\epsilon > 0$, simply note that

$$U - \underline{E}_t(U) + \epsilon \le U - \inf_{\omega\in\Omega} \underline{E}_t(U)(\omega) + \epsilon \tag{45}$$

so $U - \inf_{\omega\in\Omega} \underline{E}_t(U)(\omega) + \epsilon$ is sure to be desirable at time t, because $U - \underline{E}_t(U) + \epsilon$ is. Consequently, we have that

$$\underline{E}_0\left(U - \inf_{\omega\in\Omega} \underline{E}_t(U)(\omega) + \epsilon\right) \ge 0 \tag{46}$$

and because this holds for all $\epsilon > 0$, we arrive at Eq. (44), after using the constant additivity of $\underline{E}_0$.

"if". Suppose Eq. (44) holds. Consider any gamble $U \in \mathcal{L}(\Omega)$. If U is sure to be desirable at time t, then $\underline{E}_t(U)(\omega) \ge 0$ for all $\omega \in \Omega$. Consequently, by Eq. (44),

$$\underline{E}_0(U) \ge \inf_{\omega\in\Omega} \underline{E}_t(U)(\omega) \ge 0 \tag{47}$$

so U is desirable now. Principle 5 follows. □

4.2 Implications

The treatment for previsions relied on the scoring definition of expectation, via Proposition 4. However, no proper scoring rules exist for lower previsions [7]. We try to generalise Proposition 4 anyway. We do so in two ways: first without scoring, and secondly using the relationship between expressions of the form $(X - a)^2$, and lower and upper variance—which is the closest notion to scoring we have for lower previsions. There are certainly more ways to go about it, but for this introductory paper, we will stick to these two.

First, we derive the following imprecise counterpart of Corollary 8.

Corollary 13. *If Principle 5 is satisfied, then*

$$\underline{E}_0(X - \underline{E}_t(X)) \ge 0. \tag{48}$$

[5] The attentive reader will note that the 'only if' part of the proof of Proposition 12 remains valid, if in Principle 5, we take desirability to be actual desirability, rather than almost-desirability.

Proof. By Eq. (44):

$$\inf_{\omega\in\Omega} \underline{E}_t(X - \underline{E}_t(X))(\omega) \le \underline{E}_0(X - \underline{E}_t(X)). \tag{49}$$

Now note that $\underline{E}_t(X - \underline{E}_t(X))(\omega) = 0$ for all $\omega \in \Omega$, by coherence of $\underline{E}_t(\omega)$. □

Clearly, if we were to impose a conditioning interpretation, Eq. (48) corresponds to one of Walley's conditions for coherence [9, p. 303, (C11)].

Corollary 13 has a number of interesting immediate consequences:

Corollary 14. *If Principle 5 is satisfied, then*

$$\overline{E}_0(X - \overline{E}_t(X)) \le 0, \tag{50}$$
$$\underline{E}_0(\underline{E}_t(X)) \le \underline{E}_0(X) \le \underline{E}_0(\overline{E}_t(X)), \tag{51}$$
$$\overline{E}_0(\underline{E}_t(X)) \le \overline{E}_0(X) \le \overline{E}_0(\overline{E}_t(X)). \tag{52}$$

Proof. The first inequality holds by:

$$0 \le \underline{E}_0(-X - \underline{E}_t(-X)) = -\overline{E}_0(X - \overline{E}_t(X)). \tag{53}$$

The second inequality holds because

$$\underline{E}_0(X - \underline{E}_t(X)) \ge 0 \tag{54}$$
$$\implies \underline{E}_0(X) + \overline{E}_0(-\underline{E}_t(X)) \ge 0 \tag{55}$$
$$\implies \underline{E}_0(X) \ge \underline{E}_0(\underline{E}_t(X)) \tag{56}$$

and

$$\underline{E}_0(-X - \underline{E}_t(-X)) \ge 0 \tag{57}$$
$$\implies \overline{E}_0(-X) + \underline{E}_0(\overline{E}_t(X)) \ge 0 \tag{58}$$
$$\implies \underline{E}_0(\overline{E}_t(X)) \ge \underline{E}_0(X). \tag{59}$$

The thrid one is proved similarly. □

We can derive neither a lower bound on $\underline{E}_0(\underline{E}_t(X))$, nor an upper bound on $\overline{E}_0(\overline{E}_t(X))$, for example, due to the possibility of dilation [8].

Finally, let us see how far we can get with lower and upper variance. We need the following lemma [9, p. 618, G2]:

Lemma 15. *For every gamble X, there are previsions E_1 and E_2 in the credal set of $\underline{E}$, such that for all $a \in \mathbb{R}$:*

$$\underline{\mathrm{var}}(X) := \underline{E}((X - E_1(X))^2) \le \underline{E}((X - a)^2), \tag{60}$$
$$\overline{\mathrm{var}}(X) := \overline{E}((X - E_2(X))^2) \le \overline{E}((X - a)^2). \tag{61}$$

In particular, for all $a \in \mathbb{R}$, $(X - a)^2 - \underline{\mathrm{var}}(X)$ is desirable. Note that $\overline{\mathrm{var}}(X) - (X - a)^2 - \epsilon$ is non-desirable, however this does not help us very much—in fact, this led us to investigate temporal sure preference also for non-desirability, yet the resulting principle seems not very compelling, and leads to serious issues.

Proposition 16. *If Principle 5 is satisfied, then*

$$\underline{\mathrm{E}}_0(\underline{\mathrm{var}}_t(X)) \leq \underline{\mathrm{E}}_0((X-Y)^2), \tag{62}$$
$$\overline{\mathrm{E}}_0(\underline{\mathrm{var}}_t(X)) \leq \overline{\mathrm{E}}_0((X-Y)^2). \tag{63}$$

Proof. By definition of variance,

$$\underline{\mathrm{E}}_t\left((X-Y)^2 - \underline{\mathrm{var}}_t(X)\right) \geq 0 \tag{64}$$

(remember that Y is a known constant at time t). Whence, by Eq. (44), also

$$\underline{\mathrm{E}}_0((X-Y)^2 - \underline{\mathrm{var}}_t(X)) \geq 0. \tag{65}$$

Concluding, by coherence,

$$\underline{\mathrm{E}}_0((X-Y)^2) \geq \underline{\mathrm{E}}_0(\underline{\mathrm{var}}_t(X)), \tag{66}$$

and

$$\overline{\mathrm{E}}_0((X-Y)^2) \geq \overline{\mathrm{E}}_0(\underline{\mathrm{var}}_t(X)). \tag{67}$$

□

Again, we cannot say anything about, say, $\overline{\mathrm{E}}_0(\overline{\mathrm{var}}_t(X))$.

As for adjusted lower expectation, if we are happy to bound, say, the upper expectation of the future lower variance, by Eq. (63), any function Y of observed quantities at time t which aims to minimize $\overline{\mathrm{E}}_0((X-Y)^2)$ could be a candidate. A good choice of function of course depends on the optimisation problem, and an obvious stumbling block is that even already for a simple linear form, say $a + bY$, $\overline{\mathrm{E}}_0((X-(a+bY))^2)$ cannot be written as a function of the imprecise expectation and imprecise variance of X and Y. In other words, at this point, we seem to get stuck, although there might be interesting and feasible solutions for specific cases, for instance, using techniques from imprecise regression.

5 Conclusion

We have discussed the temporal sure preference principle in the context of desirability and lower previsions. We found more than one way to generalise the temporal sure preference principle to lower previsions, so we used the simplest version, related directly to desirability.

We have identified an expression for natural extension under the suggested temporal sure preference principle. We then derived a simple condition, which guarantees consistency of the temporal sure preference principle with prior specifications, and which also guarantees that those prior specifications are not modified by adopting the temporal sure preference principle, so we can still use the usual (non-temporal) form of natural extension for gambles as far as our current beliefs are concerned.

We have also derived a host of bounds on lower and upper expectations of future lower and upper expecations and variances. In this initial investigation, a particular challenge which remains is to provide lower and upper bounds on *all* future lower and upper expectations and variances.

An obvious next step would be to investigate possible updating rules implied by the temporal sure preference principle, for example using ideas from imprecise regression. The optimisation problems involved do not appear to have nice closed solutions in general, essentially due to the non-linearity of the lower and upper previsions. It would be very interesting to find non-trivial imprecise instances of lower previsions where such updating rules could be calculated explicitly. In this paper, we had an initial look at linear updating rules and lower and upper variance, but of course there might be many more ways to go about it.

Temporal reasoning without conditioning also raises interesting questions about the need for a possibility space. In fact, it is one of the premises of temporal reasoning that we cannot specify in advance what the possibility space ought to be. In the current paper, it serves only as a mathematical construct to establish a clear link with Walley's [9] approach to lower previsions and desirability. We might be better off simply ignoring the possibility space entirely, and instead working with random quantities directly, following the approach of de Finetti [2] and Williams [10, 11].

Finally, one might wonder, why not also introduce a principle for temporal coherence concerning non-desirability: say, if a gamble is surely non-desirable at a future time t, should it also be non-desirable now? One can show that, for previsions, this principle is equivalent to the usual temporal sure preference principle.

However, for lower previsions, this is no longer so, and it leads us to infer additional constraints. In fact, it leads to additional constraints that are usually not satisfied in the standard theory when updating is taken to be conditioning. We simply note here that temporal reasoning on non-desirability seems far less compelling, certainly so under the standard interpretation that non-desirability merely means that we do not say whether we accept a gamble or not. Here, a reject-accept approach to desirability [6] might lead to a better treatment.

Acknowledgements

The first author would like to thank Enrique Miranda and Gert de Cooman for interesting discussions on the topic of temporal coherence, which fed the desire to write this paper. We are particularly indebted to Gert de Cooman for making many useful suggestions which improved the clarity of the paper. We also thank all reviewers, whose many scrutinizing yet very constructive and supportive comments have helped us clarifying numerous technical aspects of the paper.

References

[1] Peter S. Craig, Michael Goldstein, Jonathan C. Rougier, and Allan H. Seheult. Bayesian forecasting for complex systems using computer simulators. *Journal of the American Statistical Association*, 96(454):717–729, June 2001. URL: http://www.jstor.org/stable/2670309, doi:10.1198/016214501753168370.

[2] Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*. Wiley, New York, 1974–5. Two volumes.

[3] Michael Goldstein. Temporal coherence. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, *Bayesian Statistics 2*. 1985.

[4] Michael Goldstein. Prior inferences for posterior judgements. In M. C. D. Chiara et al., editors, *Structures and Norms in Science*, pages 55–71. Kluwer, 1997.

[5] Michael Goldstein and David A. Wooff. *Bayes Linear Statistics: Theory and Methods*. Wiley, Chichester, 2007.

[6] Erik Quaeghebeur, Gert de Cooman, and Filip Hermans. Accept & reject statement-based uncertainty models. Submitted. arXiv:1208.4462.

[7] Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane. Forecasting with imprecise probabilities. In F. Coolen, G. de Cooman, T. Fetz, and M. Oberguggenberger, editors, *ISIPTA'11: Proceedings of the Seventh International Symposium on Imprecise Probability: Theories and Applications*, pages 317–326, Innsbruck, 2011. SIPTA.

[8] Teddy Seidenfeld and Larry Wasserman. Dilation for sets of probabilities. *The Annals of Statistics*, 21(3):1139–1154, 1993. doi:10.1214/aos/1176349254.

[9] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[10] Peter M. Williams. Notes on conditional previsions. Technical report, School of Math. and Phys. Sci., Univ. of Sussex, 1975.

[11] Peter M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44(3):366–383, 2007. doi:10.1016/j.ijar.2006.07.019.

[12] Marco Zaffalon and Enrique Miranda. Probability and time. *Artificial Intelligence*, 198:1–51, May 2013. doi:10.1016/j.artint.2013.02.005.