The Retrospective Account of Bayesian Updating[†]

Abstract. This paper argues for a new account of Bayesian updating by taking a retrospective approach to diachronic coherence. This approach says that an agent is diachronically coherent whenever the information she has revised her beliefs on satisfies whatever constraint we would want our evidence to satisfy. This approach opposes a common way of thinking about the Bayesian framework, according to which it treats evidence as a black box. The aim of this paper is to provide a different interpretation of Bayesianism's main updating constraint by filling in this black box with a Bayesian account of evidence.

1 Introduction

Say I come to you and I ask: How should I revise my beliefs when I learn something new? It's likely you would recommend one of two things. On the one hand, you might tell me that when I get evidence, I should revise my beliefs using the right updating procedure. I should update my beliefs using a rule that is likely to result in beliefs that are more accurate or that have some other valuable property. On the other hand, you might tell me that when I update my beliefs, I should make sure that I update on the right information: on information that is likely to result in beliefs that are more accurate or more valuable. This is similar advice given from two different perspectives. In the first case, you assume that I already have some evidence in hand, and you tell me what I should do with it. In the second case, you assume that I already have some updating rule in mind, and you tell me what it should be applied to. In both cases, you assume that what I am after is a more accurate, or valuable, set of beliefs.

Bayesians have traditionally adopted the first perspective. According to Bayesian orthodoxy, evidence is a black box that gives rise to the norm of conditionalization.[1] Whatever it means for an agent to get evidence, when this happens, it triggers the requirement to condition on it. This paper argues for a new account of Bayesian updating by defending what I will call "the retrospective approach" to diachronic coherence. The retrospective approach takes the second perspective described above. Instead of claiming that an agent is diachronically coherent when, assuming she has some piece of evidence, we can infer that she has revised her beliefs by conditioning on it, the retrospective approach says that an agent is diachronically coherent when, *assuming she has conditioned on some information*, we can infer that this information satisfies whatever constraint we would want our evidence to satisfy.[2]

[†] All acknowledgments of help have been temporarily removed in order to preserve blind review.

[1] For this description, see, for instance, Skyrms [1987].

[2] To the best of my knowledge, the only other reference to "retrospective conditioning" comes briefly in passing from Diaconis and Zabell [1982, p.822]. This discussion is an attempt to give this concept, from a statistics paper, a philosophical interpretation.

On the retrospective approach, then, whether or not an agent has updated at all, properly speaking, is determined by whether she has conditioned her beliefs on the right information. I argue that one natural interpretation of the right information is information that is not inconsistent. In conditioning on inconsistent information, we violate a norm I call "diachronic probabilism". Like other Bayesian norms of coherence, diachronic probabilism can be justified on accuracy grounds.

Before we begin, I want to note two assumptions that I will be making. The first assumption this discussion will make is that the best interpretation of Bayesian updating is something that is, in some sense, up for grabs. It's often noted that there are many different ways of articulating Bayesianism's main commitments.[3] Given this, a reasonable question to ask, though one that is not asked very often, is what formal properties a Bayesian would want her evidence to have. The aim of this paper is to offer an answer to this question. In so doing, this paper resolves a certain theoretical puzzle by making sense of the second perspective described above.

There's a more concrete puzzle that the second perspective described above allows us to resolve. It's well-known that Bayesian epistemology is often accused of placing unrealistic demands on the epistemic agent. Bayesian epistemology requires that the agent be logically omniscient: that she know all logical truths. Bayesian epistemology also requires the agent to have computational capacities that far outstrip those of any actual agent.[4] These problems have generated an entire literature of responses, which are well worth a paper of their own. But even without evaluating each of them, the general lesson is clear: we must either revise Bayesian epistemology's norms or we must revise the way we interpret these norms.

I propose that we do the latter. Rather than understand Bayesian epistemology as a set of action-guiding norms, I propose that we understand Bayesian epistemology as a set of evaluative norms. Just as we can say that a knife is good insofar as it's sharp and bad insofar as it isn't, without imposing any obligations on anyone to do anything, we can say that certain credence distributions, or sets of credence distributions, have certain features that make them good or bad, without imposing any obligations on anyone to do anything.

The evaluative approach, unlike the action-guiding approach, is compatible with a retrospective account of Bayesian updating. If I only know whether I have evidence *after* I have revised my beliefs, as the retrospective approach implies, there's no sense in which I can be offered any guidance about how to update. There's no sense in which you can tell me what credences I should update to if you do not know in advance what kind of evidence I have. But not knowing this in advance, it seems appropriate for you to retrospect over my belief revision and ask whether I have revised my beliefs in a way that is good or that is best—in a way that implies that the information I have revised my beliefs on has those formal properties that make it evidence after all.

Understanding Bayesian epistemology as a set of evaluative norms is also com-

[3] Famously, I.J. Good [1971] claimed that there are at least 46,656 varieties of Bayesians.

[4] For canonical expressions of this complaint, see Harman [1986, p.25-26] and Earman [1992, p. 56].

patible with the first approach described above. I can ask whether I have revised my beliefs using whatever updating procedure is good or best. But once we adopt the evaluative approach, it becomes less clear why this first approach should be favored. More strongly, taking the evaluative approach seriously might seem to recommend favoring the retrospective approach: cutting ties with the action-guiding approach is perhaps done most decisively by adopting an account of updating that is incompatible with it.

The retrospective approach implies new possibilities for understanding the structure of diachronic coherence. While we can't engage in more than one update at a time, we can retrospect over more than one update at a time. The retrospective approach is, then, compatible with the idea that diachronic coherence may be a relation that is defined over *sets* of belief revisions. We'll also see that the retrospective approach has some weird consequences, ones that are at odds with our intuitions about what the agent ought to do. However, in rejecting the idea that the Bayesian framework generates these kinds of 'oughts' in the first place, we are free to embrace these weird consequences. In embracing these consequences, my retrospective account tests the limits of the evaluative approach to diachronic coherence.

There's one more nice feature of the retrospective approach that I'd like to point out. On the first, more orthodox approach to diachronic coherence described above, we must always update in response to some question. To be defensible on value-based grounds, the orthodox account must assume that the evidence partition—which might be taken to represent the subject matter being inquired about—is known in advance of the agent's update. More precisely, it is known in advance of her getting the evidence she updates on.[5] This feature of the orthodox account means that it cannot accommodate cases of "reasoning in the wild". While I sometimes update my beliefs in response to some particular question that I am trying to settle, at other times, I am simply struck by information. I go outside, not intending to settle the question of whether it is raining, and yet nevertheless come to discover that it is and, therefore, adjust my beliefs accordingly. The retrospective approach can accommodate these kinds of cases. As we will see in just a moment, while the retrospective approach implies that there is some question that is answered by an agent's update, in virtue of this update implying a certain evidence partition, it takes no stand on whether this question was ever asked. By neither requiring, nor precluding that we always reason in response to a particular question we have in mind in advance, the retrospective approach accommodates a variety of forms of reasoning.

The first assumption this discussion will make, then, is that an alternative to the orthodox interpretation of Bayesian updating is both possible and desirable. A second assumption this discussion will make is that, when deciding on the best interpretation of Bayesian updating, we should consider Bayesian updating in its most general form. Although most discussions about diachronic coherence begin from the assumption that we get propositional evidence, or certain evidence, I'm going to assume, for illustrative purposes at least, that Bayesian epistemology includes norms that govern

[5] See Skyrms [1987]. For some more recent discussions, see also Pettigrew [2016] and Schoenfield [2017]. More on this in §4 and §5.

cases where the agent gets *uncertain* evidence as well. In short, I'm going to assume that something like "Jeffrey conditioning" is a legitimate updating rule.[6] This doesn't seem unreasonable. Arguably, the situations we find ourselves in most often are cases where we do seem to have uncertain evidence. I assume that the reason we idealize away from the sorts of cases that Jeffrey conditioning covers isn't because they aren't important, but because we assume they will be covered in the same way as regular Bayesian conditioning. It's good practice to abstract away from the uncertainty of evidence when this detail isn't important for whatever purpose the framework is being put towards. However, when the problem at issue is how we should interpret the Bayesian framework itself, the structure of uncertain evidence is surely a detail worth considering.

If we assume that beliefs come in degrees, then, it seems arbitrary to withhold this assumption from our evidence. As Jeffrey [1992, p.11] put it, opinion, no matter how it comes to us, should be probabilities "all the way down, to the roots." Jeffrey's rule represents a particularly strong way of reflecting the idea that what makes some view Bayesian is its commitment to probabilism. The account of diachronic coherence defended in this paper does this as well; on my account of it, diachronic coherence is probabilism all the way down. However, though my account is best articulated from a perspective that assumes that we sometimes get uncertain evidence, the results of this discussion don't depend upon this assumption. One can reject this assumption and still accept my proposal.

Here's how the discussion will go. In §2, I distinguish two accounts of diachronic coherence and go on to introduce a third approach. On my retrospective approach, an agent is diachronically coherent whenever the information she has conditioned her beliefs on satisfies whatever constraint we would want our evidence to satisfy. In §3, I argue that one constraint we would want our evidence to satisfy is that it not be inconsistent over time. Since this constraint is equivalent to the requirement that the evidence an agent has accumulated over time be a probability function, I call this way of filling out the retrospective approach "diachronic probabilism". In §4, I argue that we can use the resources of epistemic utility theory to give an accuracy argument for diachronic probabilism. Finally, I conclude in §5 by briefly comparing diachronic probabilism to some other Bayesian updating rules.

2 A Retrospective Approach to Diachronic Coherence

Standard Bayesianism assumes that an agent's degrees of confidence, or credences, in the propositions she entertains can be represented as an assignment of real numbers to those propositions. It further assumes that the following synchronic norm governs this assignment:

[6] I use the term "Jeffrey conditioning" colloquially here to refer to updates on uncertain evidence. The correct interpretation and formulation of Jeffrey conditioning is a complicated matter, made all the more complicated by the fact that Jeffrey's own views about his rule seem to have changed over time. This interpretive question won't be addressed here.

Probabilism: At each time, an agent's credences should satisfy the probability axioms.

Most Bayesians also take an agent's credences to satisfy a diachronic constraint. This constraint tells us how we should revise our beliefs when we get some new evidence. It tells us that we should be guided by our conditional probabilities. Normally this is interpreted as the rule that our current unconditional probabilities should equal our prior probabilities conditional on the evidence that we've gotten. However, as Blackwell and Girshick [1979, p.221] show, this rule falls out of the more general requirement that our conditional probabilities remain fixed, or "rigid", on the evidence that we've gotten in between this transition.[7] Some proposition, A, is rigid on E, for the transition from p to p', just in case the following holds:

Rigidity: $p(\mathrm{A} \,|\, \mathrm{E}) = p'(\mathrm{A} \,|\, \mathrm{E})$, where $p(\mathrm{E}), p'(\mathrm{E}) > 0$.

If our conditional probabilities are to guide us in our updates, they must be the same before and after these updates. That's what rigidity says. While any Bayesian account of diachronic coherence must be committed to rigidity, there are different accounts one can give of how the latter mathematical relation gives rise to the former normative relation. The aim of this paper is to tell a new kind of story about how this happens.

Before moving on, a quick word about terminology. In what follows, I will assume that "diachronic coherence" is synonymous with "updating" and, also, with "conditioning or revising on one's evidence". I will assume that these are all success terms, in contrast to "conditioning or revising on one's information", though I will sometimes violate these linguistic stipulations when context would make it too awkward to do otherwise.

The literature provides us with two options, when it comes to understanding the relation between rigidity and diachronic coherence. On the one hand, we might adopt a deflationary account of diachronic coherence and assume that there is nothing more to being diachronically coherent than to remaining probabilistically coherent over time. How does this get us rigidity? Famously Blackwell and Girshick [1979] show that rigidity follows as a consequence of the probability axioms. For any probabilistic belief transition, there is *some* partition that is sufficiently fine-grained to represent the transition in question as rigid on this partition.[8] An account of diachronic

[7] For discussion of this result, see Diaconis and Zabell [1982, p.824]. For other discussions that emphasize this guiding feature of our conditional probabilities, see Jeffrey [1965], Jeffrey [1970], Skyrms [2006] and, more recently, Weisberg [2014].

[8] See also Diaconis and Zabell [1982, p.824] and Jeffrey [1992, p.122-128] for discussion of this result and related issues. As Diaconis and Zabell note, there will be cases where our conditional probabilities are undefined for some partition—namely, where we assign a member of our partition a credence of zero. However, their result still holds for all updates, if we take a sufficient partition to be a partition that is sufficient to represent a probabilistic belief transition as an update that is conditional on every proposition in this partition for which a conditional probability is defined (*Cf.* Blackwell and Girshick [1979, §8.4.3]).

coherence might, then, take our evidence to be the "sufficient partition" for a probabilistic belief revision: the coarsest partition that is sufficiently fine-grained to represent this revision as rigid with respect to it.[9] An account of diachronic coherence that involves this way of understanding evidence can be formulated in the following way:

A Deflationary Account of Diachronic Coherence:
Where $\mathcal{S}=\{B_1, ..., B_n\}$ is a set of beliefs that form a partition, and where, p, is the agent's prior probability distribution, and, p', is the agent's posterior probability distribution, an agent is diachronically coherent just in case the following holds:

$\forall p \forall p' \exists \mathcal{S} (\forall B_i \in \mathcal{S})\ \forall A (p(A \left| B_i) = p'(A \right| B_i)$), if defined.

Since the deflationary account entails that we reverse-engineer an evidence partition out of an agent's update, it does not imply anything about when an agent has gotten the evidence that she updates on. Nor does it imply anything about whether she knew in advance what partition this evidence would come from. It simply says that our update will be consistent with this partition and with the evidential values this partition implies.

Of course, a more familiar way of understanding the relation between getting evidence and revising in accordance with rigidity takes these to be two different steps in the updating process. We get some piece of evidence *first* and then the only right way to revise one's beliefs is in a way that is rigid on this *particular* piece of evidence.[10] According to this more orthodox approach to updating, the process of getting evidence is the *exogenous* part of an update: it is the assumption we use to determine whether our updating norm has been satisfied. By contrast, the process of revising one's beliefs on this evidence in accordance with rigidity is the *endogenous* part of an update. It is the norm-governed part of the update.[11] Whatever it means for an agent to get evidence, when this happens, it triggers the requirement to condition on it. This picture of updating can be captured by the following conditional norm:

The Orthodox Account of Diachronic Coherence:
Where $\mathcal{S}=\{B_1, ..., B_n\}$ is a set of beliefs that form a partition whose values have

[9] For example, say that I update from the probability distribution, $p(EF)$=.25, $p(E\overline{F})$=.25, $p(\overline{E}F)$=.25, $p(\overline{EF})$=.25 to the probability distribution, $p'(EF) = .3, p'(E\overline{F}) = .2, p'(\overline{E}F) = .3, p'(\overline{EF}) = .2$. The coarsest grained partition capable of representing this as a transition by the agent's conditional probabilities is $\{F_i\}$, since this transition can be represented as an update on $p(F)$=.6, $p(\overline{F})$=.4. But we can also represent this transition as an update on the partition $\{EF_i\}$, where the members of this partition are assigned the values that correspond to the agent's posterior probability distribution.

[10] For some canonical discussions of this way of understanding Bayesian updating, see Jeffrey [1970], Jeffrey [1992], Diaconis and Zabell [1982], Skyrms [1987] and Howson and Urbach [2006].

[11] The endogenous/exogenous distinction was introduced by Howson and Urbach [2006]. Note that exogenous here need not mean causal. I am simply taking it to mean 'outside the domain of the normative'.

changed non-inferentially, and where, p, is the agent's prior probability distribution, and, p', is the agent's posterior probability distribution, an agent is diachronically coherent just in case the following holds:

$\forall p \forall p' (\exists \mathcal{S} \rightarrow \forall (\mathrm{B_i} \in \mathcal{S}) \forall A (p(\mathrm{A} \big| \mathrm{B_i}) = p'(\mathrm{A} \big| \mathrm{B_i})))$, if defined.

The orthodox account says that an agent is diachronically coherent whenever she gets evidence and revises her beliefs in a way that is rigid on this evidence. Since a probabilistic agent might fail to be coherent in this way—since she might fail to revise her beliefs on the particular piece of evidence she has in hand—the orthodox account of updating is stronger than probabilism, which seems like something we would want.

I want to propose a way of understanding diachronic coherence that reverses the relation just described. Instead of requiring that rigidity apply, *conditional* on us having some evidence, I propose that we require that the information we have revised our beliefs on satisfy some evidential constraint, *conditional* on us having revised by rigidity. More carefully, I want to propose that some probabilistic belief revision constitutes an update just in case, (1) there is a sufficient partition for that transition and, (2) that sufficient partition satisfies whatever constraint we would want our evidence to satisfy:

A Retrospective Approach to Diachronic Coherence:
Where $\mathcal{S}$=$\{\mathrm{B_1}, ..., \mathrm{B_n}\}$ is a set of beliefs that form a partition, and where, p, is the agent's prior probability distribution and, p', is the agent's posterior probability distribution, an agent is diachronically coherent just in case the following two conditions hold:

1) $\forall p \forall p' \exists S (\forall \mathrm{B_i} \in S)\ \forall A (p(\mathrm{A} \big| \mathrm{B_i}) = p'(\mathrm{A} \big| \mathrm{B_i})$), if defined, and

2) $\mathcal{S}_1, \mathcal{S}_2, ..., \mathcal{S}_\mathrm{n}$ satisfy our constraint on evidence.

Like the orthodox account, the retrospective approach also has an endogenous part and an exogenous part, but they are reversed. Instead of assuming that an agent has some piece of evidence, and going on to ask whether she has arrived at the right posterior credences by conditioning, the retrospective approach assumes that an agent has arrived at her posterior credences by conditioning, and goes on to ask whether she has conditioned on the right information—on evidence.

Summing up, then, there are two perspectives from which to develop an account of diachronic coherence that is stronger than probabilism. These perspectives correspond to two different questions one might ask. On the one hand, one might ask: *assuming that this is my evidence, have I revised my beliefs by conditioning on it*? If I have, then according to the orthodox account, I am diachronically coherent. Alternatively, one might ask: *assuming I have conditioned my beliefs on some information, does this information have those features that would make it evidence*? If it does, then according to the retrospective approach, I am diachronically coherent.

3 A Retrospective Account of Diachronic Coherence

3.1 Diachronic probabilism

We've considered three ways of understanding what it could mean for a Bayesian agent to be diachronically coherent. One of these ways is not like the others though. Unlike the deflationary account and the orthodox account, the retrospective approach is a schema. It leaves open what features some information must have in order to be evidence.

In this section, I propose and defend a way of filling out the retrospective approach. My aim in doing this is to provide a proof of concept for the retrospective approach; to show that there is a plausible way of developing this approach into an account. While there are likely other such ways, it goes beyond this discussion to consider and evaluate all these possibilities. Though it will take a little while to develop, the idea that motivates my account is a simple one: an account of diachronic coherence should not take a stand on *how* an agent incorporates her evidence into her posterior credence distribution. Instead it should yield the result that two agents who begin from the same priors and get the same evidence end up with the same posterior credence distribution. I'll argue that we can extract a substantive account of evidence from this very weak and very plausible constraint.

To see how we do this, let us start by distinguishing two ways that the Bayesian framework might interpret how our evidence interacts with our priors. On the *sequential model* of updating, we revise our beliefs at each time that we get some new piece of evidence. The sequential model of updating tells us to condition our current information on our current priors, where our current priors were generated by conditioning the information we had a moment ago on the priors that we had before those, and so on. Each of the approaches described in the previous section assumes that we update sequentially in this sort of way.

Some have argued that the sequential model of updating is too strong: we should be able to relax the requirement that an agent must update at each and every time she gets some new piece of information. If an agent has gotten information at several different times in between t_0 and t_6, she should (for instance) be able to condition on the information she has gotten in between t_0 to t_3, and then on the information that she has gotten in between t_3 to t_6. Or she should be able to condition on the evidence that she's gotten in between t_0 and t_6 all at once.[12] Since this model of updating takes as input the information an agent has accumulated over an interval of time, call this the *cumulative model* of updating.[13]

There's a stronger argument that sometimes gets made when discussing sequential and cumulative models of updating. This argument says that not only should we be able to update on the information we have accumulated over an interval of time, but that whether we engage in sequential or cumulative updating should not make a

[12] See, for instance, Skyrms [1983], Titelbaum [2013], Meacham [2015] and Meacham [2016].

[13] I borrow the language of "sequential" and "cumulative" from Meacham [2016], though I use these terms a little differently than he does.

difference to the end result.[14] If an agent has gotten information at several different times in between t_0 and t_6, she should (for instance) be able to condition on all of this information at once, or on each piece of this information in turn, without this making a difference to her posterior credence distribution. This idea is motivated by a kind of anti-arbitrariness principle: it's arbitrary to think that anything other than priors and evidence should determine an agent's credences. Where we assume the orthodox account of updating, this reasoning leads to the following principle:

> **Anti-Arbitrariness About Updating (ORT):** Our posterior credence distribution should be the same regardless of whether we update sequentially or cumulatively.

What would it look like for the retrospective approach to avoid arbitrariness? Recall this approach says that diachronic coherence is a matter, not of our arriving at the right posterior credence distribution, but of the information that our belief revision implies satisfying whatever constraint we would want our evidence to satisfy. Given this, one way (though perhaps not the only way) of articulating what anti-arbitrariness would amount to on the retrospective approach is to say that it amounts to the requirement that the information an agent has revised her beliefs on satisfy whatever constraint we would want our evidence to satisfy, regardless of whether this constraint is imposed on all of this information at once, or on each piece of this information in turn. Suppose that H is the proposition that *it will hail today*. And suppose that, at t_1, an agent revises her beliefs on the weighted partition $\{(H, .5), (\overline{H}, .5)\}$ and then, at t_2, she revises her beliefs on the weighted partition $\{(H, .6), (\overline{H}, .4)\}$. A plausible anti-arbitrariness principle might say that, if we require the latter two weighted partitions to satisfy some constraint on evidence, then the set $\{\{(H, .5), (\overline{H}, .5)\}, \{(H, .6), (\overline{H}, .4)\}\}$ should also satisfy this constraint:

> **Anti-Arbitrariness About Updating (RET):** We are governed by the same evidential constraint, regardless of whether we update sequentially or cumulatively.

Recall that the aim of this section is to fill out the retrospective approach by defending a constraint on evidence. While **Anti-Arbitrariness About Updating (RET)** does not provide us with an evidential constraint, it nudges us towards one. A condition that is trivially satisfied by all sequential accounts of updating is that our evidence be probabilistic. If we assume probabilism—if we assume that our credences should, at every moment, satisfy the probability axioms—then an account of updating that tells us to update at every moment that we get some new evidence will entail that our evidence satisfies probabilism. Not only does our evidence satisfy probabilism on all sequential accounts of updating, but the discussion of the previous section suggests that this is the *only* constraint that all such accounts have in common. Even the deflationary account satisfies this constraint.

[14] See, for instance, Meacham [2015].

If probabilism is the weakest constraint on sequential evidence, then by **Anti-Arbitrariness About Updating (RET)** it must also be a constraint on an agent's cumulative evidence: on the information she has accumulated over an interval of time. It's a bit trickier to say what it would mean for an agent's cumulative evidence to satisfy probabilism. However a plausible thought is that an agent's cumulative evidence satisfies probabilism whenever it is *representable* as a probability function, i.e., whenever there is a probabilistically weighted partition that represents all the constraints that this evidence imposes. An agent who gets $\{(E, .5), (\overline{E}, .5)\}$ at t_1, and then gets $\{(F, .5), (\overline{F}, .5)\}$ at t_2, has as evidence $\{(EF, .25), (\overline{E}\mathrm{F}, .25), (E\overline{F}, .25), (\overline{EF}, .25)\}$.[15] By contrast, an agent who, as in the example above, gets $\{(H, .5), (\overline{H}, .5)\}$ at t_1 and then gets $\{(H, .6), (\overline{H}, .4)\}$ at t_2 does not have evidence, properly speaking, since there is no way to represent this set of information as a probability function. In this case, the agent has *inconsistent information*: she has information that contains propositions that have been assigned more than one value.[16] With this in mind, we might take the following to be a non-trivial constraint on an agent's belief revisions over time:

> **Diachronic probabilism:** An agent's cumulative evidence should be representable as a probability function.

The aim of this section is to fill out the retrospective approach to diachronic coherence by providing a constraint on evidence. I think that diachronic probabilism is a good candidate for this constraint. Diachronic probabilism turns what has traditionally been assumed to be a trivial assumption about the nature of evidence—that it is probabilistic—into a substantive diachronic constraint, without appealing to anything stronger than a weak anti-arbitrariness principle. We can now fill in our schema from the previous section in a way that entails that diachronic coherence is a matter of the set of sufficient partitions of each of our belief transitions satisfying diachronic probabilism:

> **A Retrospective Account of Diachronic Coherence**:
> Where p and p' are probability distributions that are held sequentially, and where $\mathcal{S}=\{\mathrm{B}_1, ..., \mathrm{B}_n\}$ is a set of beliefs that form a partition, an agent is diachronically

[15] See Diaconis and Zabell [1982, §4.2] for a suggestion along these lines.

[16] This proposal contains an important ambiguity in what it means for information to be consistent. On the one hand, one might take an agent's evidence to be consistent just in case the values over the evidence partition in question never change, either directly in virtue of updating on the same partition more than once, or indirectly in virtue of making an update that changes the values along this partition: i.e., in virtue of updating over two weighted partitions that aren't probabilistically independent. On the other hand, one might maintain that evidence is consistent *even if* the second condition is violated: even if the updates are over two weighted partitions that aren't probabilistically independent.

To remain diachronically coherent in the first, stronger sense, an agent must continually update on finer grained partitions—on $\{\mathrm{E}_i\}$, and then on $\{\mathrm{EF}_j\}$, etc. While the first account of evidential consistency strikes me as the more plausible one, I won't take a stand on this question, since nothing in this paper turns on which of these accounts we adopt.

coherent just in case the following two conditions hold:

1) $\forall p \forall p' \exists \mathcal{S} (\forall \mathrm{B_i} \in \mathcal{S})\, \forall A (p(\mathrm{A} \big| \mathrm{B_i}) = p'(\mathrm{A} \,\big|\mathrm{B_i})$), if defined, and

2) $\{\mathcal{S}_1, \mathcal{S}_2, ..., \mathcal{S}_\mathrm{n}\}$ satisfies diachronic probabilism.[17]

It's tempting to think that we violate diachronic probabilism anytime we violate ordinary probabilism. However, I'm going to assume what all standard accounts of updating assume, which is that an updating rule is only a constraint on an agent who already conforms to ordinary probabilism. This means that the only way for an agent to violate diachronic probabilism is for her to have inconsistent information. It might seem odd to suggest that the kind of violation I've just described is the violation of a probabilistic norm. It's natural to think of probabilism as a constraint on a function. But an agent with inconsistent information has evidence that is not even a function at all.

One reason we should think of inconsistent information as the violation of a probabilistic norm is because of the way that inconsistent evidence patterns. If we take our evidence from before—$\{\{(H, .5), (\overline{H}, .5)\}, \{(H, .6), (\overline{H}, .4)\}\}$—we see that, while we are probabilistically coherent over H and $\overline{H}$ at each time, we are probabilistically incoherent over each of these propositions, paired with their logical complement from a different time: $\{\{(H, .5), (\overline{H}, .4)\}, \{(H, .6), (\overline{H}, .5)\}\}$. Even if we do assume, then, that any probabilistic constraint must be a constraint on a function, we can understand diachronic probabilism as imposing precisely this sort of constraint. Diachronic probabilism is a constraint on the function over the members of an evidence partition at two different times.

Later on I'll argue that we can say something even stronger in support of the idea that diachronic probabilism and ordinary probabilism are relevantly similar. We can extend the general reasoning that we offer for conforming to probabilism to diachronic probabilism. If we assume what seems reasonable—that our epistemic norms are defined by the reasons we have to conform to them—this will mean that diachronic probabilism and ordinary probabilism are norms of the very same kind.

3.2 Defending diachronic probabilism

We've just seen that violations of diachronic probabilism correspond to cases where an agent's evidence is inconsistent over time. I've argued that we should understand this inconsistency as a form of diachronic incoherence. While I've already provided some motivation for diachronic probabilism by showing that it falls out of a plausible anti-arbitrariness principle, in this section, I'll motivate diachronic probabilism in a different way by arguing that it fills a gap in the Bayesian literature. Diachronic probabilism allows us to accommodate externalist intuitions about the nature of evidence.

The key move my account makes is to take diachronic coherence to be a relation that is defined, not over belief revisions, but over *sets* of belief revisions. My account

[17] This is shorthand, of course, for the requirement that the values along these partitions satisfy diachronic probabilism.

takes diachronic coherence to be a relation that is defined over the information that underwrites these revisions. While thinking about diachronic coherence in this way goes against the grain, it does not seem like a crazy suggestion. There is precedence in the literature for accounts of Bayesian updating that don't take it to be a relation defined over individual credence functions that the agent has at two different times. Consider time-slice epistemology, which argues for synchronic surrogates of orthodox conditionalization.[18] Diachronic probabilism looks like the inverse of time-slice epistemology. Instead of taking the units over which our updating constraint is defined to be *less* temporally extended than the units assumed by the orthodox account, diachronic probabilism takes these units to be *more* temporally extended than the units assumed by the orthodox account. Instead of providing a synchronic surrogate for the diachronic constraint of conditionalization, diachronic probabilism provides a diachronic surrogate for the synchronic constraint of probabilism.

There's a more fundamental way that diachronic probabilism and synchronic conditionalization are inverses of one another. We've already seen that one reason for thinking it is a bad thing for updates to have been made on inconsistent evidence is that inconsistent evidence isn't probabilistic. But we can also defend diachronic probabilism on more intuitive grounds. A different reason we have for thinking that getting inconsistent evidence is a bad thing appeals to an intuition that I think is widely shared, namely, that unlike our credences, our evidence should not be up for revision. Our evidence should have a standing that our credences don't have. Consider accounts of evidence that take it to be knowledge, or to be factive, or—a bit closer to home—that take evidence to be that which we get with a credence of one, so that we can never lose it (more on this in a moment). My approach does not require that evidence be any of these things. But in maintaining that an agent is diachronically coherent just in case her evidence is consistent over time, it preserves, in a formal way, an important feature of each of these accounts. It preserves the *unrevisability* of our evidence. My account preserves the idea that our evidence is unrevisable by entailing that we are diachronically incoherent—that we do not have evidence at all—when the values over the partitions we have conditioned on have changed over time.

Some would take the unrevisability of evidence to be a bug, rather than a feature, of an updating framework built to accommodate it. But the fact that not everyone would endorse this picture of evidence is no mark against diachronic probabilism. Consider again its inverse, synchronic conditionalization. One of the problems that motivates synchronic conditionalization is that agents on the orthodox, sequential framework are unable to lose certain evidence. As Meacham [2010, p.94] notices, the problem of not being able to lose certain evidence is, more fundamentally, the problem of there being a mismatch between our intuitions about evidence and our updating framework. Cases where it feels as though we should be able to lose certain evidence—cases of memory loss, for instance—are motivated by internalist intuitions about the nature of evidence. They are motivated by the thought that our evidence should su-

[18] For two canonical examples of synchronic surrogates of conditionalization, see Meacham [2010] and Hedden [2015a,b].

pervene on the way that things seem to us at the present moment.[19] But the orthodox, sequential framework is committed to a form of externalism about evidence that we get with certainty, since it entails that anytime our current seemings conflict with the evidence that is encoded in our priors, the latter swamps the former. Synchronic surrogates of conditionalization resolve this mismatch by positing priors that don't encode an agent's previous evidence, thereby bringing the framework in line with evidential internalism.

Diachronic probabilism is a solution to a problem that is the mirror image of the problem of losing certainties: *the problem of retaining uncertainties*. Just as it's impossible to lose certain evidence on the orthodox, sequential framework, when we get conflicting evidence, it's impossible to retain uncertain evidence in this same situation. Anytime our current uncertain evidence conflicts with our previous uncertain evidence, the former swamps the latter. This swamping feature commits the orthodox, sequential framework to a form of internalism about uncertain evidence. It entails that our total evidence is determined by the way that things are for us at the present moment. To the extent that we are pulled by the externalist intuition that evidence is not the kind of thing that should be beholden to how things are for us at the present moment, diachronic probabilism resolves this mismatch between our intuitions and the framework by positing an account of evidence that entails that we are diachronically *incoherent* anytime our current evidence swamps our previous evidence—or, equivalently, anytime the values over the partitions we have conditioned on have changed over time.

From the point of view of the internalist, then, orthodox, sequential updating encodes too much memory, while from the point of view of the externalist, orthodox, sequential updating encodes not enough memory. Internalists resolve the former problem by updating a credence function that the agent has at each moment. I've argued that externalists can resolve the latter problem by making diachronic coherence a matter of updating on information that remains unchanged over time. Like synchronic conditionalization, diachronic probabilism is formally compatible with many different accounts of evidence. But it is most friendly to those who lean externalist, in the way that synchronic conditionalization is most friendly to those who lean internalist. It's no objection to diachronic probabilism, then, that it is not for everyone, any more than this is an objection to time-slice epistemology. The cost of a framework, like the orthodox, sequential framework, that is not biased is that it succumbs to counterexample—to agents who can't forget or to agents who can't remember—whenever it is paired with some substantive account of evidence. Diachronic conditionalization is a remedy for the second kind of counterexample, just as synchronic conditionalization is a remedy for the first.

Before closing this section, I want to address a final concern about diachronic probabilism. One might object that this norm encodes two ideas that are in tension: (1) that our evidence can be uncertain, and (2) that our evidence is unrevisable. Though the arbitrary swamping of uncertain evidence is surely a bad thing, being unable to *ever*

[19] See, again, Meacham [2010, p.94] and, also, Hedden [2015a,b].

revise uncertain evidence is surely just as bad. Isn't the point of being able to hold uncertain evidence that we can revise it later if we need to?

More generally, one might worry about the notion of inconsistency at the heart of this account. It's clear why we should not hold propositional evidence that is inconsistent. It's clearly a bad thing to hold as evidence both some proposition and its negation. But, notwithstanding the externalist intuitions my account seems to vindicate, one might ask whether it is really irrational to update on the same evidence partition more than once? And, if so, *why* is it irrational?

As I've suggested already, I think it's a mistake to try to defend the irrationality of uncertain, inconsistent evidence—not because such evidence is not irrational, but because the standing of such evidence as rational or irrational is irrelevant. It's the wrong question to be asking. It's true that my account has no intuitive story to tell about why an agent who updates on the proposition that it's raining outside with a credence of .6 is thereby precluded from updating on the same proposition with a credence of .7. But the orthodox account similarly has no story to tell about why it precludes certain behavior that is not, intuitively, irrational. The orthodox account has no story to tell about why an agent who gets evidence as the result of being hit on the head is precluded from ignoring this evidence, and is required, instead, to condition her beliefs on it. The orthodox account also has no story to tell about why an agent who is incapable of assigning probabilities to all of those propositions one would need to in order to be capable of conditioning is precluded from abstaining from this act that she is incapable of performing. Again, I want to say that, to the extent that our intuitions about how an agent ought to act closely track action-guidingness, or what it is rational for her to do, these intuitions cannot be trusted to give us the right evaluative account of Bayesian updating.[20] In the next section, I'll argue that what we *can* say is that updates by diachronic probabilism, like updates by the orthodox account, have some valuable features that can be spelled out in accuracy terms. We may not be able to give an account of the irrationality of updates that violate diachronic probabilism, but, like updates that violate orthodox conditionalization, we can give an account of their badness. The account outlined in this section is, then, in keeping with the aim of this paper, which is to test the limits of an evaluative approach to Bayesian norms.

4 Justifying Diachronic Probabilism

In the last section, I defended an account of diachronic coherence I called diachronic probabilism. In the section before that, I described the general approach to diachronic coherence that diachronic probabilism vindicates. No defense of any Bayesian norm is complete, however, before we have shown the value of conforming to it. In this section, I argue that diachronic probabilism can be grounded in the value of accuracy. An agent who avoids updating on inconsistent evidence by conforming to diachronic probabil-

[20] One might object that I am conflating action-guidingness with rationality. If one doesn't think that these two concepts overlap in the relevant way, one can simply substitute "action-guidingness" for "rationality" in this paragraph, without losing what I take to be the essential point being made.

ism does better, in an accuracy-related sense, than an agent who flouts this requirement. To see this, we first need a framework for understanding this kind of claim.

Epistemic utility theory is a decision theory that identifies the rationality of credal states with their tendency to promote a specifically epistemic form of utility, one that is typically characterized as *gradational accuracy*. In addition to a credence function, epistemic utility theory appeals to the notion of a *vindication function* and an *inaccuracy measure*. The kind of vindication function we are interested in here—an alethic vindication function—is a characteristic function of a possible world: it specifies, for every proposition, whether that proposition is true (1) or false (0) at that world. An inaccuracy measure is a function that meets certain constraints that make its measure of the distance between a credence function and a vindicator an appropriate measure of the inaccuracy of that credence function.[21] Finally, a local inaccuracy measure is a function from credence functions and propositions and vindicators to real numbers, representing the degree of inaccuracy a credence function lends to a proposition at a vindicator, while a global inaccuracy measure is a function from credence functions and vindicators to real numbers, representing the total degree of inaccuracy of a credence function at a vindicator.

The aim of accuracy-first epistemology is to vindicate various epistemic norms by showing that they are the means to the end of gradational accuracy. Accuracy-first epistemology provides us with theorems that show that, relative to some reasonable class of inaccuracy measures, conforming to probabilism and a form of Bayesian updating ensures that we will satisfy certain analogues of the decision-theoretic norms of dominance (namely, accuracy-dominance) and maximize expected utility (namely, minimize expected inaccuracy).[22] Accuracy arguments are believed to be an improvement over the representation theorem arguments and dutch book arguments that have traditionally been appealed to in order to justify Bayesian norms. While accuracy arguments appeal to many of the same mathematical resources as these more established arguments, the value they promote is a distinctly epistemic one.

To get an idea of what the accuracy-dominance argument for probabilism looks like, consider a weather forecaster who assigns a credence of .7 to the proposition that it will hail tomorrow and a credence of .5 to the proposition that it won't hail tomorrow. This agent has a probabilistically incoherent set of credences, since $.7+.5\neq 1$. It can be shown that, because this agent has a probabilistically incoherent set of credences, her credences are accuracy-dominated. According to a reasonable measure of inaccuracy, there exists a probabilistically coherent set of credences that will be less inaccurate than it is, no matter which state of the world comes about. In the case just described, these

[21] The two most standard constraints are Truth-Directedness, the constraint that accuracy be the overriding factor in determining how epistemically good our credences are, and Extentionality, the constraint that it be the *only* relevant factor. There are other reasonable, though slightly more contentious, constraints on inaccuracy measures. The inaccuracy measure that many take to satisfy all reasonable constraints is the Brier score and, more recently, the more general class of divergences outlined in Pettigrew [2016].

[22] For some of the most prominent of these arguments, see Greaves and Wallace [2006], Joyce [2009], Leitgeb and Pettigrew [2010a], Leitgeb and Pettigrew [2010b] and Pettigrew [2016].

probabilistically coherent credences are $\{(H, .6), (\overline{H}, .4)\}$.[23] This can be represented geometrically in the following way:

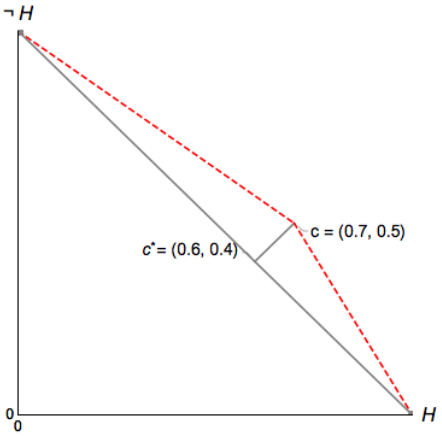

Figure 1: The argument for probabilism

By contrast, a probabilistic set of credences would *not* be accuracy-dominated: it is not the case that there is some set of credences that would do better than it does, no matter which state of the world comes about. To see this, notice that moving to a different probability function along the diagonal line between H and $\overline{H}$ above will always move one closer to one world, by virtue of moving one further away from another. Since non-probabilistic credence assignments are accuracy-dominated, but probabilistic credence assignments are not accuracy-dominated, a decision-theoretic agent with an epistemic utility function does better by conforming to probabilism.

Accuracy-first epistemology also provides theorems that show that the norm to plan to conditionalize—"plan conditionalization"—has certain accuracy advantages.[24] Both Greaves and Wallace [2006] and Pettigrew [2016] show that plan conditionalization minimizes expected inaccuracy, relative to any other updating policy one could

[23] If we assume the Brier score, for instance, we get that if it does hail, $\{(H, .7), (\overline{H}, .5)\}$ will be more inaccurate than $\{(H, .6), (\overline{H}, .4)\}$:

$$(1 - .6)^2 + (0 - .4)^2 < (1 - .7)^2 + (0 - .5)^2$$
$$.32 < .34$$

And if it doesn't hail, $\{(H, .7), (\overline{H}, .5)\}$ will still be more inaccurate than $\{(H, .6), (\overline{H}, .4)\}$:

$$(0 - .6)^2 + (1 - .4)^2 < (0 - .7)^2 + (1 - .5)^2$$
$$.72 < .74$$

[24] The term "plan conditionalization" is coined by Easwaran [2013].

have adopted instead.[25] Plan conditionalization is an updating norm that assumes that we know in advance the evidence partition we plan to update over. As many have noted, plan conditionalization just is the orthodox account, with this additional caveat made explicit.[26]

We'll return to consider plan conditionalization briefly again in the next section. In the remainder of this section, I will argue that diachronic probabilism can also be justified by an accuracy argument. Indeed, I will argue that diachronic probabilism can be justified by reasoning similar to the reasoning that justifies probabilism. To begin to see this, recall that diachronic probabilism says the following:

> **Diachronic probabilism:** An agent's cumulative evidence should be representable as a probability function.

Earlier we noted that an obstacle to thinking of diachronic probabilism as a probabilistic norm is that violations of diachronic probabilism—sets of inconsistent evidence—aren't functions at all. This is equally an obstacle to an accuracy-dominance argument for diachronic probabilism. Since inconsistent evidence is not a function, and since accuracy-dominance can only be defined over functions, inconsistent evidence cannot be claimed to be accuracy-dominated. Nevertheless, I think we can say that inconsistent evidence is defective in a closely related sense. It's tempting to want to say that what goes wrong, from a value perspective, in cases where our evidence is inconsistent, is that we have evidential value assignments that point us in different directions. Therefore, no matter how the world turns out, they could not both be pointing us in the right direction. This is similar to the reasoning that Kolodny [2007] appeals to in order to make sense of the inconsistency of believing a contradiction. Kolodny argues that the best way to interpret the requirement that gets violated when we violate the law of non-contradiction is not as the requirement to conform to the wide-scope norm to have coherent attitudes, but as the requirement to avoid violating a narrow-scope norm. Since either P is true or $\neg P$ is true, but not both, either it will be the case that one is required by reason not to believe P, or it will be the case that one is required by reason not to believe $\neg P$. Where we know this, and where we violate the law of non-contradiction, we can't help but violate some substantive requirement. Either we violate the requirement not to believe P based on our epistemic reasons, or we violate the requirement not to believe $\neg P$ based on our epistemic reasons. Kolodny takes our reasons to be something like our evidence.

Kolodny's argument doesn't apply to the sorts of cases we have been discussing. The sorts of value gluts he has in mind are truth-value gluts, and not gluts that represent inconsistent evidence. What's wrong with inconsistent evidence from an accuracy

[25] Not all have taken plan conditionalization to be justified in this way. For a way of justifying plan conditionalization that appeals to dominance reasoning, see Briggs and Pettigrew [forthcoming].

[26] A consequence of this is that the strongest version of the orthodox account that is defensible on value-based grounds makes this account a synchronic norm. For this point, see for instance Pettigrew [2016, p.188].

perspective then? I think that, analogous to Kolodny's suggestion, what goes wrong in cases where we have inconsistent evidence is that, no matter how the world turns out, the values we assign to our evidence partition will inevitably make it the case that some piece of evidence I hold will be more inaccurate than some other piece of evidence I hold. Consider again the case where I have as evidence two sets of probabilistic values over the partition $\{H, \overline{H}\}$. Perhaps these values are $\{\{(H, .5), (\overline{H}, .5)\}, \{(H, .6), (\overline{H}, .4)\}\}$. If it turns out that H, then $\{(H, .5), (\overline{H}, .5)\}$ will be more inaccurate than $\{(H, .6), (\overline{H}, .4)\}$. If it turns out that $\overline{H}$, then $\{(H, .6), (\overline{H}, .4)\}$ will be more inaccurate than $\{(H, .5), (\overline{H}, .5)\}$.[27] No matter how the world turns out, some piece of evidence I hold will inevitably be more inaccurate than some other piece of evidence I hold.[28,29]

As conceded already, this is no accuracy-dominance argument. An accuracy-dominance argument would show that violating diachronic probabilism ensures that *the* evidence an agent holds does worse in *every* state of the world. What my argument shows is that an agent who violates diachronic probabilism will have *some* evidence that does worse in *some* state of the world. The first claim is interesting; the second claim seems trivial. However, though my argument for diachronic probabilism is not an accuracy-dominance argument, it is clearly not trivial. There is a good-making property that agents who satisfy diachronic probabilism will have: the property of not having evidence partitions whose values are such that one of the weighted evidence parti-

[27] Again, assuming the Brier score, if it does hail, then $\{(H, .5), (\overline{H}, .5)\}$ will be more inaccurate than $\{(H, .6), (\overline{H}, .4)\}$:

$$(1 - .6)^2 + (0 - .4)^2 < (1 - .5)^2 + (0 - .5)^2$$
$$.32 < .5$$

And if it doesn't hail, then $\{(H, .6), (\overline{H}, .4)\}$ will be more inaccurate than $\{(H, .5), (\overline{H}, .5)\}$:

$$(0 - .5)^2 + (1 - .5)^2 < (0 - .6)^2 + (1 - .4)^2$$
$$.5 < .72$$

[28] One might object that the fact that some of our evidence will also be *less* inaccurate than some of our other evidence, no matter which state of the world comes about, means that the value of inconsistent evidence is inconclusive. Why not look at the glass as half full and be happy that some of our evidence will be less inaccurate than some of our of our other evidence, come what may. However, a crucial assumption made by Kolodny, and those who work within the accuracy framework, is that there is more disvalue to inaccurate credences/believing falsehoods than there is value to accurate credences/believing truths. My argument preserves the analogue of these asymmetries by assuming that the badness of being more inaccurate outweighs the goodness of being less inaccurate.

[29] One might object that we should interpret these gluts as imprecise evidence. In other words, we should assume that the values that have been assigned to P and $\neg P$ represent the upper and lower bounds of two sets of probability distributions. But I don't think that this is right. The sorts of considerations that motivate the view that imprecise credences can be rational don't seem to be at work in cases like the one described above. For some of these considerations, see White [2009].

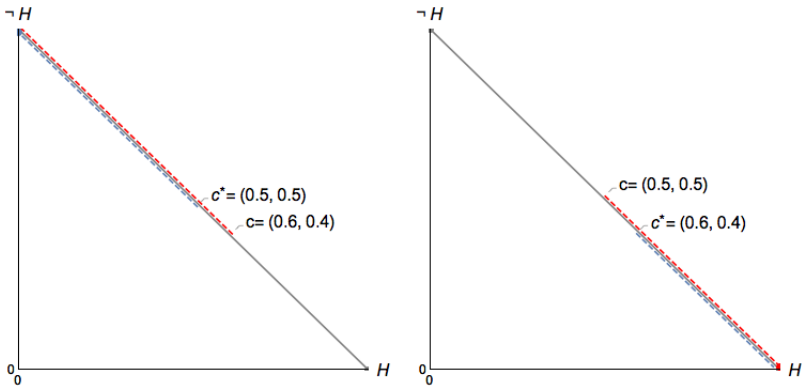

Figure 2: The argument for diachronic probabilism

tions these values give rise to is inevitably more inaccurate than the other.[30]

Conversely, an agent who violates diachronic probabilism has the bad-making property of inevitably holding some piece of evidence that *does* do worse than some other piece of evidence that they hold. No matter what happens, such an agent is better off not holding some of her evidence. If it hails tomorrow, then she would have done better to have discarded some of her evidence. Same thing if it doesn't hail tomorrow. Just as we know that some part of a contradiction will be false, no matter what happens, we know that some part of an agent's inconsistent evidence set will be more inaccurate than some other part of this set, no matter what happens.

While my argument for diachronic probabilism is not an accuracy-dominance argument, then, it generalizes the reasoning behind such arguments. Both the argument for diachronic probabilism and the accuracy-dominance argument for probabilism say that no matter which state of the world comes about, an agent does better by discarding some of her credences, in favor of some probabilistic credences. The argument for probabilism says, in addition, that the probabilistic credences the agent has reason to favor is the same, no matter which state of the world comes about. We can say, then, that we have the same general reason for conforming to diachronic probabilism that we have for conforming to probabilism: in order not to have evidence that we will inevitably have reason to discard.

Before closing, it's worth noting two important features of the argument. First, notice that the way the argument has been formulated makes sense of our earlier caveat that diachronic probabilism is only a constraint on an agent who is probabilistically coherent at a time. An agent with non-probabilistic, inconsistent evidence does not have evidence that she has reason to discard, no matter what happens. It's easy to see that a pair of evidence sets that does not lie on the diagonal line in between H and

[30] My argument is, then, in the tradition of others who have argued that we can have interesting accuracy arguments that are weaker than dominance arguments. See, for instance, Fitelson and Easwaran [2015].

$\neg H$—for instance, $\{\{(H, 1), (\overline{H}, 1)\}, \{(H, 0), (\overline{H}, 0)\}\}$—may be *equally* inaccurate, no matter which state of the world comes about.

This way of thinking of things also ensures that the argument does not over-generate in a different direction. It's almost trivial that, for many pairs of weighted evidence partitions, one will be more inaccurate than the other. However, recall that the retrospective approach takes our evidence to be the sufficient partition for our update. The target of the current argument is, then, not *weighted* evidence partitions, but evidence partitions simpliciter. And there's a general claim (the one we've just made) that holds of evidence partitions that does not hold of weighed evidence partitions. To reiterate: For any pair of evidence partitions we have updated over that violate diachronic probabilism, no matter what values these partitions get assigned, one of the weighted evidence partitions that results will inevitably be more inaccurate than the other. By contrast, it's *not* the case that, for any pair of weighted evidence partitions whatsoever, one of these will inevitably be more inaccurate than the other. For instance, this will not be true of $\{\{(E, .5), (\overline{E}, .5)\}, \{(F, .5), (\overline{F}, .5)\}\}$.

The existence of cases like the previous one show that there is a modal property that updates on evidence partitions that violate diachronic probabilism have that weighted evidence partitions, in general, don't have: the property of being such that they inevitably or *necessarily* give rise to weighed evidence partitions, such that one will be more inaccurate than the other in some state of the world. The only way to update over a partition more than once—in a way that assigns this partition inconsistent values—is for these values to change. If the same partition has been assigned different values, then one set of values will inevitably be more inaccurate than the other. That's what the accuracy argument for diachronic probabilism says.

One might object that to show that it is a defect to have some property necessarily entails that it is a defect to have that property in the actual world, thereby re-inviting the worry that the argument over-generates to weighted evidence partitions that don't violate diachronic probabilism. But this reasoning is clearly mistaken. Take the dominance argument for probabilism, which assumes that a credence function is defective if there is some other credence function that does better in every state of the world: i.e., if there is some credence function that necessarily does better than it does. It does not follow from this that a credence function is defective if there is some credence function that does better than it does in the actual world. This would be, in effect, to abandon probabilism by maintaining that we are only permitted to hold a credence of 0 or 1 in any proposition.

Finally, one might still be skeptical of the general strategy of defining an updating norm by arguing that certain evidence partitions have a modal property we would not want them to have. However, again, the dominance argument for probabilism also advances this sort of claim. So does any norm that invokes some dominance-like principle. It's true these modal properties are of little use in guiding the agent in the actual world. But that is just part and parcel of the evaluative approach we have adopted. More importantly, if our aim in appealing to the accuracy literature is to give an accuracy-based interpretation of the badness of some feature of our updates that is distinctive of

that very feature, then this is also just part and parcel of the accuracy-first approach.

5 Conclusion

Here are the results of our discussion:

- There is a plausible way of understanding diachronic coherence, according to which it is the requirement that the information we use to revise our beliefs have a certain property that makes it evidence (§2).
- There is a plausible way of understanding diachronic coherence, according to which the property that makes some information evidence is a relational property. Or, equivalently,
- There is a plausible way of understanding diachronic coherence, according to which it is a norm that governs sets of belief revisions, rather than individual belief revisions (§3).
- A plausible norm for sets of belief revisions is that they be made on consistent evidence (§3).
- The norm to have consistent evidence can be justified on accuracy grounds (§4).

Throughout this discussion, I've described several updating rules that are staples of the Bayesian literature. It would take a much longer paper to establish that diachronic probabilism should be preferred to these more familiar norms. However, before closing, I want to quickly head off the worry that diachronic probabilism may be strictly worse than these other norms. I've motivated diachronic probabilism in a number of different ways throughout this discussion—first, by noting that the retrospective approach that it vindicates is the characteristic form of the evaluative approach to diachronic coherence; and, second, by noting that diachronic probabilism can accommodate externalist intuitions about the nature of evidence. Together with these other grounds, I hope that my last few remarks here will provide enough of a glimpse of how a defense of diachronic probabilism would go to end on an optimistic note.

There are a couple of dimensions along which we might evaluate an updating norm. On the one hand, we would likely want to say that an updating norm that imposes stronger constraints in more epistemic situations is better than one that imposes weaker constraints in fewer epistemic situations. On the other hand, we would likely also want to say that an updating norm that yields more expected value is better than one that yields less expected value. When it comes to the first of these dimensions, diachronic probabilism lies in between the deflationary account and the orthodox account outlined at the beginning of this discussion. Diachronic probabilism imposes stronger constraints than the deflationary account: on the former account, but not the latter, an agent must do more than merely conform to ordinary probabilism in order to be diachronically coherent over a set of probabilistic belief transitions. But diachronic

probabilism imposes weaker constraints than the orthodox account: on the latter account, but not the former, an agent must do more than conform to ordinary probabilism in order to be diachronically coherent over a single probabilistic belief transition (for, recall, diachronic probabilism is only defined over sets of such transitions). Since the orthodox account imposes constraints on more belief transitions than diachronic probabilism, one might reason that we should prefer the orthodox account to diachronic probabilism.

However, I don't think that this follows. While diachronic probabilism can be justified by the accuracy argument I describe in the previous section, the accuracy literature has notoriously failed to provide us with an argument for the orthodox account. Leitgeb and Pettigrew [2010b] defend a diachronic Bayesian updating rule. However, as Pettigrew [2016] notes, this rule doesn't take updates to be defined by rigidity. While the orthodox account provides us with stronger constraints, then, there is no reason to think that it leaves us with a more valuable set of credences.

There is a Bayesian updating norm that we do have value-based reason to conform to. Earlier I noted that plan conditionalization, which is relevantly similar to the orthodox account, is justified by the norm to minimize expected inaccuracy. Given this, one might reason that credences that conform to plan conditionalization are more valuable than credences that conform to diachronic probabilism. After all, the norm to minimize expected inaccuracy picks out credences that are uniquely best, whereas the norm not to be accuracy-dominated does not necessarily pick out credences that do better than other non-dominated credences. Given that the norm that underwrites diachronic probabilism is even weaker than the norm not to be accuracy-dominated, one might reason that conforming to plan conditionalization yields a more valuable set of credences than conforming to diachronic probabilism.

However, while conforming to plan conditionalization may yield a more valuable set of credences than conforming to diachronic probabilism, plan conditionalization governs a much more restricted range of epistemic situations. Not only does plan conditionalization fail to apply to agents with uncertain evidence, but, as Pettigrew [2016, p. 188] and others have noted, and as we have already seen in the introduction, it only applies to agents who know in advance what evidence partition they will update over. Given that plan conditionalization imposes constraints on a much more restricted set of belief revisions, it's far from clear that we have stronger reason to conform to this norm than we have to conform to diachronic probabilism.

In the end, then, diachronic probabilism may turn out to be the updating norm that strikes just the right balance of being strong enough to impose constraints on a wide range of epistemic situations, while also recommending credences that promote some form of epistemic utility. While these remarks are too brief to establish this once and for all, I hope they at least suggest that we cannot rule out diachronic probabilism as a serious contender for the Bayesian norm of diachronic coherence that we should adopt. More generally, I hope that they show that we cannot rule out the retrospective approach to Bayesian updating.

References

David A Blackwell and Meyer A Girshick. *Theory of games and statistical decisions*. Courier Corporation, 1979.

R. A. Briggs and Richard Pettigrew. An accuracydominance argument for conditionalization. *Noûs*, forthcoming.

Persi Diaconis and Sandy L Zabell. Updating subjective probability. *Journal of the American Statistical Association*, 77(380):822–830, 1982.

John Earman. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, ma: MIT Press, 1992.

Kenny Easwaran. Expected accuracy supports conditionalization—and conglomerability and reflection. *Philosophy of Science*, 80(1):119–142, 2013.

Branden Fitelson and Kenny Easwaran. Accuracy, coherence, and evidence. *Oxford Studies in Epistemology*, 5:61–96, 2015.

I.J. Good. Letter to the editor. *The American Statistician*, 1971.

Hilary Greaves and David Wallace. Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind*, 115(459):607–632, 2006.

Gilbert Harman. *Change in view: Principles of reasoning*. The MIT Press, 1986.

Brian Hedden. Time-slice rationality. *Mind*, 124(494):449–491, 2015a.

Brian Hedden. *Reasons without persons: rationality, identity, and time*. OUP Oxford, 2015b.

Colin Howson and Peter Urbach. *Scientific reasoning: the Bayesian approach*. Open Court Publishing, 2006.

RC Jeffrey. The logic of decision. Technical report, MCGRAW-HILL BOOK CO INC NEW YORK, 1965.

Richard Jeffrey. *Probability and the Art of Judgment*. Cambridge University Press, 1992.

Richard C Jeffrey. Dracula meets wolfman: Acceptance vs. partial belief. In *Induction, acceptance and rational belief*, pages 157–185. Springer, 1970.

James Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In Franz Huber and Christoph Schmidt-Petri, editors, *Degrees of Belief*, pages 263–297. Synthese, 2009.

Niko Kolodny. Ix—how does coherence matter? In *Proceedings of the Aristotelian Society (Hardback)*, volume 107, pages 229–263. Wiley Online Library, 2007.

Hannes Leitgeb and Richard Pettigrew. An objective justification of bayesianism i: Measuring inaccuracy. *Philosophy of Science*, 77(2):201–235, 2010a.

Hannes Leitgeb and Richard Pettigrew. An objective justification of bayesianism ii: The consequences of minimizing inaccuracy. *Philosophy of Science*, 77(2):236–272, 2010b.

Christopher JG Meacham. Unravelling the tangled web: Continuity, internalism, uniqueness and self-locating belief. *Oxford studies in epistemology*, 3:86, 2010.

Christopher JG Meacham. Understanding conditionalization. *Canadian Journal of Philosophy*, 45(5-6):767–797, 2015.

Christopher JG Meacham. Ur-priors, conditionalization, and ur-prior conditionalization. *Ergo, an Open Access Journal of Philosophy*, 3, 2016.

Richard Pettigrew. *Accuracy and the Laws of Credence*. Oxford University Press, 2016.

Miriam Schoenfield. Conditionalization does not maximize expected accuracy. *Mind*, 126(504):1155–1187, 2017.

Brian Skyrms. Three ways to give a probability assignment a memory. 1983.

Brian Skyrms. Dynamic coherence and probability kinematics. *Philosophy of Science*, 54(1):1–20, 1987.

Brian Skyrms. Diachronic coherence and radical probabilism. *Philosophy of science*, 73 (5):959–968, 2006.

Michael G Titelbaum. *Quitting certainties: A Bayesian framework modeling degrees of belief*. Oxford University Press, 2013.

Jonathan Weisberg. Updating, undermining, and independence. *The British Journal for the Philosophy of Science*, 66(1):121–159, 2014.

Roger White. Evidential symmetry and mushy credence. In T. Szabo Gendler and J. Hawthorne, editors, *Oxford Studies in Epistemology*, pages 161–186. Oxford University Press, 2009.