SPEED-OPTIMAL INDUCTION AND DYNAMIC COHERENCE

MICHAEL NIELSEN AND ERIC WOFSEY

ABSTRACT. A standard way to challenge convergence-based accounts of inductive success is to claim that they are too weak to constrain inductive inferences in the short run. We respond to such a challenge by answering some questions raised by Juhl (1994). When it comes to predicting limiting relative frequencies in the framework of Reichenbach, we show that speed-optimal convergence—a long-run success condition—induces dynamic coherence in the short run.

Keywords. Dynamic coherence; induction; limiting relative frequencies; long run and short run; martingales; Reichenbach; speed-optimality

1. INTRODUCTION

Reichenbach held that the inductive methods used in the sciences are essentially rules for estimating probabilities. Probabilities, in turn, received a frequency interpretation, and this led Reichenbach to regard the discovery of limiting relative frequencies as a primary aim of scientific inquiry.[1]

Reichenbach advocated a particularly simple inductive method for predicting frequencies, the *straight rule*. Using the straight rule, he attempted a "pragmatic vindication" of induction in response to Humean skepticism.[2] We must concede to Hume, Reichenbach thought, that we cannot be certain that nature is regular. But if it is regular, then the straight rule will reveal this to us in the limit of inquiry. In particular, if the relative frequency of an outcome in a repeated experiment approaches a stable limit, then the straight rule's conjectures about the outcome's frequency necessarily approach the same limit.

The chief problem with Reichenbach's account of inductive success—convergence to the correct limiting relative frequencies, when the limits exist—is that it places no constraints whatsoever on the kinds of inductive inferences that can be made in the short-term.[3] Arbitrary conjectures in response to a finite amount of data can always be extended in a way that secures convergence in the long run.

In view of this limitation of Reichenbach's account, it is natural to ask whether stronger criteria of inductive success are able to induce substantive short-term constraints. This line of thought is pursued by Juhl (1994), who introduces the notion of *speed-optimal convergence*. Juhl shows that the straight rule is speed-optimal in his sense and that there are inductive methods that, although convergent, are not speed-optimal. Considered as a criterion of inductive success, then, speed-optimality places more constraints on inductive methodology than the Reichenbachian account, which requires mere convergence.

Importantly, however, Juhl's analysis leaves open one of the questions that motivates it. Namely, does requiring speed-optimal convergence of one's inductive method induce significant short-term constraints? Or can arbitrary short-term behavior always be extended in

---

Date: June 28, 2019.

[1] Reichenbach (1938, 1949). Also see van Fraassen (2000).

[2] Salmon (1991).

[3] Salmon (1966). For a more contemporary take on some of the limitations of Reichenbach's account, see Huttegger (2017a, 3.1).

a speed-optimal way? The primary aim of this paper is to answer these questions. Speed-optimal convergence *does* give rise to short-term constraints; not all short-term behavior can be extended without loss of speed-optimality. Rather surprisingly, the short-term constraint that can be derived from the requirement of speed-optimal convergence is one of great independent interest in the philosophy of induction and learning: *dynamic coherence*. The importance of dynamic coherence in the Reichenbachian frequency prediction framework, which our results reveal, suggests some intriguing connections with probabilistic frameworks for learning, in which dynamic coherence also plays an important role. We discuss these connections in the paper's penultimate section by drawing on Brian Skyrms's work on probabilistic learning. In short, in both frameworks, there appear to be deep connections between convergence, speed-optimality, martingale-like structures, and dynamic coherence. Along the way to proving the results that connect speed-optimality and dynamic coherence, we will also answer another open question of Juhl's by providing a complete characterization of the speed-optimal inductive methods.[4]

The paper is organized as follows. We begin, in the next section, by presenting the standard mathematical framework for induction on relative frequencies. In section 3, we discuss speed-optimality and present our characterization result. Section 4 contains our main result: short-term inductive behavior can be extended in a speed-optimal way if and only if it is dynamically coherent. In section 5, we discuss the possible connections to probabilistic learning and martingales, and we also pose some open questions. Section 6 concludes. Proofs are in the Appendix.

2. Mathematical Preliminaries

Let $\mathcal{C}$ be the collection of all binary sequences. Sequences in $\mathcal{C}$ will be denoted by variants of $\sigma = (\sigma_1, \sigma_2, ...)$.

Let $\mathcal{S} = \bigcup_n \{0,1\}^n$ be the collection of all finite length binary sequences, which we will call *strings*. Generic elements of $\mathcal{S}$ will be denoted by variants of $s = (s_1, ..., s_n)$. Let $|s|$ denote the length of the string s.

If $\sigma \in \mathcal{C}$ and $n \in \mathbb{N}$, let $\sigma^n = (\sigma_1, ..., \sigma_n) \in \mathcal{S}$ denote the initial segment of σ of length n. Similarly, if $s \in \mathcal{S}$ and $n \leq |s|$, then $s^n = (s_1, ..., s_n)$ is the initial segment of s of length n. If $s \in \mathcal{S}$ and $a \in \{0,1\}$, let $sa = (s_1, ..., s_{|s|}, a)$. A sequence σ (resp. string s') *extends* a string s if $\sigma^{|s|} = s$ (resp. $s'^{|s|} = s$).

Let $\mathcal{C}_{\text{conv}}$ denote the collection of sequences in $\mathcal{C}$ such that the limiting relative frequency of 1s exists. That is, if $\sigma \in \mathcal{C}_{\text{conv}}$, then

$$\ell(\sigma) := \lim_{n \to \infty} \frac{\sigma_1 + ... + \sigma_n}{n}$$

exists.

An *inductive method* ϕ is a function of $\mathcal{S}$ into $[0,1]$. The value $\phi(s)$ is interpreted as a conjecture about the limiting relative frequency of 1s based on an observation of s. Note that inductive methods are just arbitrary functions from strings into the unit interval. In particular, they are not assumed to have any probabilistic structure.

---

[4]We should note at the outset that although formal learning theory has made many advances in the years since Juhl's paper, including in the study of fast and efficient inductive methods (Kelly, 1996; Schulte, 1999a,b), the results in that literature do not, to the best of our knowledge, provide immediate answers to the questions raised above.

An inductive method is called *convergent* if

$$\lim_{n \to \infty} \phi(\sigma^n) = \ell(\sigma)$$

for all $\sigma \in \mathcal{C}_{\text{conv}}$. The conjectures of convergent methods approach actual limiting relative frequencies whenever the limits exist.

The *straight rule sr* is the inductive method defined by

$$sr(s) = \frac{s_1 + ... + s_{|s|}}{|s|}$$

for all $s \in \mathcal{S}$. The straight rule always conjectures the observed relative frequency of 1s. It is immediate from the definitions of the relevant terms that the straight rule is convergent. Conjectures equal to observed frequencies are guaranteed to converge to actual limiting relative frequencies whenever the limits exist. It is in this sense that the straight rule is supposed to vindicate induction: we cannot be certain that regularities will emerge as observations unfold—for all we know, the relative frequency of 1s may oscillate forever—but if there is regularity, then the straight rule is sure to identify it in the limit.

The claim that convergence does not constrain short-term behavior can be demonstrated as follows. Let $s_1, ..., s_n$ be an arbitrary, finite collection of strings, and let $r_1, ..., r_n$ be an arbitrary collection of real numbers in $[0, 1]$. Then, there is a convergent inductive method ϕ such that $\phi(s_i) = r_i$ for all $i \in \{1, ..., n\}$. For example, let ϕ agree with sr on all strings besides, perhaps, $s_1, ..., s_n$. Since conjectures on $s_1, ..., s_n$ are irrelevant to ϕ's behavior in the limit, convergence is consistent with arbitrarily divergent predictions in the short-term.

3. Speed-Optimality

In response to Salmon's (1966) criticisms of Reichenbach, Juhl (1994) raises the question whether we can derive short-term constraints by requiring more of our inductive methods than mere convergence. If, following Reichenbach, a primary aim of scientific inquiry is to learn frequencies, then it seems reasonable to favor those inductive methods that converge to the correct frequencies as quickly as possible. A natural way to make this idea precise is as follows.

For all $\sigma \in \mathcal{C}_{\text{conv}}$, $\epsilon > 0$, and convergent inductive methods ϕ, ϕ', we write $\phi \succ_{\sigma, \epsilon} \phi'$ if and only if there exists $m \in \mathbb{N}$ such that

$$|\phi(\sigma^n) - \ell(\sigma)| \leq \epsilon < |\phi'(\sigma^m) - \ell(\sigma)| \text{ whenever } n \geq m.$$

In other words, $\phi \succ_{\sigma, \epsilon} \phi'$ holds if and only if ϕ is within, and forever remains within, ϵ of $\ell(\sigma)$ strictly before ϕ'. When $\phi \succ_{\sigma, \epsilon} \phi'$ holds, we say that ϕ *beats* ϕ' *on* σ, ϵ.

We say that an inductive method ϕ is *faster* than another inductive method ϕ' if and only if ϕ beats ϕ' on some σ, ϵ and ϕ' does not beat ϕ on any σ, ϵ. More formally:

$$\exists \sigma \in \mathcal{C}_{\text{conv}}, \exists \epsilon > 0 : \ \phi \succ_{\sigma, \epsilon} \phi' \ \text{and} \ \forall \sigma \in \mathcal{C}_{\text{conv}}, \forall \epsilon > 0 : \neg[\phi' \succ_{\sigma, \epsilon} \phi].$$

Note that the *faster than* relation is not complete. If ϕ beats ϕ' on some σ, ϵ and ϕ' beats ϕ on some other σ', ϵ', then neither ϕ nor ϕ' is faster than the other. So there is no fastest inductive method. We will say that a convergent inductive method ϕ is *speed-optimal* if and only if there does not exist a convergent inductive method ϕ' that is faster than ϕ.

We now summarize the known facts about speed-optimal convergence and provide an additional example of the property. An inductive method ϕ is called *monotonic* if and only if

$$\phi(s0) \leq \phi(s) \leq \phi(s1)$$

for all $s \in \mathcal{S}$. In other words, monotonic methods do not decrease (resp. increase) their conjectures about the frequency of 1s in response to observing an additional 1 (resp. 0).

Facts (*Juhl, 1994*). If an inductive method is convergent and monotonic, then it is speed-optimal. Hence, the straight rule is speed-optimal. There exist non-monotonic, speed-optimal inductive methods. There exist convergent inductive methods that are not speed-optimal.

Example 1. Say that an inductive method ϕ is *Laplacian* if there exist parameters $\alpha_1, \alpha_2 \geq 0$ such that for all $s \in \mathcal{S}$

$$\phi(s) = \frac{\sum_i s_i + \alpha_1}{|s| + \alpha_1 + \alpha_2}.^5$$

The straight rule is Laplacian with parameters $\alpha_1 = \alpha_2 = 0$. Intuitively, Laplacian methods are biased straight rules, with the biases encoded by the parameters α_1 and α_2. It is a straightforward exercise to show that Laplacian methods are convergent and monotonic, and therefore speed-optimal.

In view of the partial results recorded in the Facts above, Juhl asks, "Exactly which [convergent inductive methods] are speed-optimal?" (862). In the remainder of this section, we provide an answer to this question. In the next section, we extend our answer to show that speed-optimal convergence induces an interesting short-run constraint.

The formal definition of our characterizing condition for speed-optimality is somewhat technical, but the idea behind it is simple and easy to explain: the conjectures of speed-optimal methods are *rigid* in the sense that they cannot be changed without sacrificing speed. Intuitively, our characterization result shows that if ϕ's conjecture at s can be changed without the resulting method being any slower than ϕ, then ϕ cannot have been speed-optimal in the first place; and, conversely, if any change to ϕ's conjecture at s results in a slower inductive method, then ϕ is speed-optimal.

We'll now introduce the formal definition of rigidity. If we are given an inductive method ϕ, a string $s \in \mathcal{S}$, and a sequence $\sigma \in \mathcal{C}_{\text{conv}}$ extending s, we write

$$\epsilon_{\phi,s,\sigma} = \sup_{n \geq |s|} |\phi(\sigma^n) - \ell(\sigma)|,$$

which is the largest distance between $\phi(\sigma^n)$ and $\ell(\sigma)$ after time $|s|$. We also define

$$I_{\phi,s,\sigma} = [\ell(\sigma) - \epsilon_{\phi,s,\sigma}, \ell(\sigma) + \epsilon_{\phi,s,\sigma}] \cap [0,1].$$

This is the smallest closed interval, centered at $\ell(\sigma)$, that contains $\phi(\sigma^n)$ for all $n \geq |s|$. Finally, we define

$$I_{\phi,s} = \bigcap_{\sigma \in \mathcal{C}_{\text{conv}}} I_{\phi,s,\sigma}.$$

One important thing to note is that, by the definitions just given, $\phi(s) \in I_{\phi,s}$ for all ϕ and s. So, $I_{\phi,s}$ is always nonempty.

Let us take a moment to discuss how we think about $I_{\phi,s}$. Intuitively, the closed interval $I_{\phi,s}$ represents ways that ϕ's conjecture at s can be changed without sacrificing speed. To see this, suppose that $x \in I_{\phi,s}$ and define a new method ϕ' from ϕ by $\phi'(s) = x$ and $\phi'(t) = \phi(t)$ for all $t \neq s$. The method ϕ' is the result of setting ϕ's conjecture at s to x and leaving ϕ's other conjectures unchanged. By unpacking the definitions above, we can see that ϕ is not faster than ϕ'. In particular, ϕ does not beat ϕ' on any σ, ϵ. To spell this out in a bit more detail, consider any sequence $\sigma \in \mathcal{C}_{\text{conv}}$ that extends s (ϕ cannot beat ϕ' on σ, ϵ if σ is *not*

---

[5]This formula generalizes Laplace's rule of succession, which is the case $\alpha_1 = \alpha_2 = 1$, and was developed independently by Johnson (1924; 1932) and Carnap (1950; 1952). See Huttegger (2017a) for more details.

an extension of s because ϕ and ϕ' make the same conjectures about all initial segments of such a σ, by construction). Now, $\phi'(s)$ is in the interval $I_{\phi,s,\sigma}$ by definition. The radius of the interval $I_{\phi,s,\sigma}$ is (by definition) the *smallest* ϵ such that $\phi(\sigma^n)$ is within ϵ of $\ell(\sigma)$ at all times n after $|s|$. Since $\phi'(s)$ lies within this interval, we do *not* have

$$|\phi(\sigma^n) - \ell(\sigma)| \leq \epsilon < |\phi'(s) - \ell(\sigma)|, \ \forall n \geq |s|$$

for any ϵ. So, since s is the only string on which the conjectures of ϕ and ϕ' differ, ϕ is not faster than ϕ'. Thus, the numbers $x \neq \phi(s)$ in the interval $I_{\phi,s}$ are *speed-preserving alternatives* to $\phi(s)$ in the sense that changing ϕ's conjecture at s to x does not result in a slower inductive method.

Rigid inductive methods do not have speed-preserving alternatives. Formally, we say that an inductive method ϕ is *rigid* if $I_{\phi,s} = \{\phi(s)\}$ for all $s \in \mathcal{S}$. By way of illustrating the concept, consider the following simple example of a non-rigid inductive method.

Example 2. Let ψ make the same conjectures as the straight rule sr, with the exception that $\psi(00) = 0.5$. Now consider $s = 0$. First, we have $\psi(0) = sr(0) = 0$ and $\psi(01) = sr(01) = 0.5$. Next, every $\sigma \in \mathcal{C}_{\text{conv}}$ that extends s is such that $\sigma^2 \in \{00, 01\}$, and therefore $[0, 0.5] \subseteq I_{\psi,s,\sigma}$ for every $\sigma \in \mathcal{C}_{\text{conv}}$ extending s. It follows that $[0, 0.5] \subseteq I_{\psi,s}$, so ψ is not rigid.

The inductive method ψ is also not speed-optimal. To see this, let ξ make the same conjectures as ψ, with the exception that $\xi(0) = 0.5$. Then ξ is faster than ψ. Indeed, ξ beats ψ on σ, ϵ if σ is the sequence that repeats 01 forever and $\epsilon < 0.5$. On the other hand, ψ never beats ξ, since if $\psi(0) = 0$ is ever a better conjecture than $\xi(0) = 0.5$, this cannot yield faster convergence for ψ since its next conjecture $\psi(00)$ or $\psi(01)$ will be 0.5.

Our first result is that rigidity characterizes speed-optimality.

Theorem 1. *A convergent inductive method is speed-optimal if and only if it is rigid.*

Theorem 1 implies that ψ in Example 2 is not speed-optimal, and the proof of the theorem provides a general method for constructing methods that are faster than non-rigid methods (one such faster inductive method that can be obtained from the proof is ξ). We will explore rigidity further in the next section, but for now we turn to the main question posed in the introduction.

4. Dynamic Coherence

The question that motivated us at the outset is whether requiring one's inductive method to be speed-optimally convergent places any substantive constraints on the method's predictions in the short-term. Juhl (1994) asks precisely this as an open question at the end of his paper:

> Is *any* short-term behavior compatible with speed-optimality?...If a negative answer to this question can be proved, then we will have established the existence of short-run norms on estimation rules. If non-trivial short-term norms can be shown to be induced by the requirement of speed-optimality, then the chief intuitive objection to Reichenbach's attempts to 'vindicate induction' would be answered (862).

The aim of this section is to show that, indeed, a negative answer to the question in the Juhl quote can be established. Speed-optimality does induce a non-trivial and, we will argue, particularly interesting short-run inductive constraint.

Let us begin by formalizing the problem. Call a function f from a finite subset A of $\mathcal{S}$ into $[0, 1]$ a *partial inductive method*. Partial inductive methods represent prediction behavior in

the short run. Since a partial inductive method f is defined on a finite set, there is some long-run time horizon n such that f is undefined for all strings of length more than n. An inductive method $\phi : \mathcal{S} \to [0,1]$ is an *extension* of $f : A \to [0,1]$ if $\phi(s) = f(s)$ for all $s \in A$, and we say that ϕ *extends* f. Now, the question whether any short-term behavior is compatible with speed-optimality becomes: *Is every partial inductive method extended by some speed-optimal, convergent inductive method?* If the answer to this question is negative, and it can be shown that exactly those partial inductive methods with property P admit speed-optimal extensions, then we say that *speed-optimality induces the short-term constraint P*. We will now start working towards a result along these lines.

Our short-run constraint turns out to be closely related to a central idea in the philosophy of induction, and in theorizing about rational learning more generally. The idea is that the conjectures that a rational inductive method makes at a given time are constrained in a particular way by the conjectures that it might make in the future. It is irrational, the idea goes, to conjecture x now while at the same time expecting to conjecture $y \neq x$ in the future *no matter what new data one observes between now and then*.

This general idea has been formalized in a number of ways across several disciplines. In the philosophy of probability and formal epistemology, the *principle of reflection* captures the idea. It says that a rational agent's conditional probability for A, given that her probability for A will be x after learning some new evidence, must be equal to x.[6] In the theory of finitely additive probability, there is a great deal of research on the concept of *non-conglomerability*, which occurs when an unconditional, finitely additive probability value does not reside in the interval spanned by its conditional probability values, given the members of a countably infinite partition.[7] In decision theory, Savage's *sure-thing principle* says that if option 1 is preferred to option 2 conditional on every member of some partition, then option 1 ought to be preferred to option 2 unconditionally.[8] Similar principles of *dynamic consistency* appear in the economics literature.[9] In statistics, results due to Lane and Sudderth (1984; 1985) show that probability estimates are *dynamically coherent* (avoid Dutch book) just in case they are contained within the closed, convex hull of possible future estimates.

The condition that we articulate is similar to all of these, and we borrow some terminology accordingly. Let $\mathcal{S}_n$ denote the collection of strings of length at most n. We say that a partial inductive method $f : \mathcal{S}_n \to [0,1]$ is *dynamically coherent* (or sometimes simply *coherent*) if for all $s \in \mathcal{S}_{n-1}$

$$f(s0) \leq f(s) \leq f(s1) \text{ or } f(s1) \leq f(s) \leq f(s0).$$

For example, dynamic coherence rules out the possibility of conjecturing 0.5 now and 0.6 after the next observation no matter what is observed. Put another way, the conjectures of dynamically coherent methods are always contained within the interval spanned by the conjectures that might be made after observing more data. We note that monotonic partial inductive methods are coherent.

We now have the following preliminary result.

---

[6] van Fraassen (1984, 1999); van Fraassen and Halpern (2016); Huttegger (2013, 2014).

[7] de Finetti (1972); Dubins (1975); Schervish et al. (1984); Kadane et al. (1996). A related phenomenon in the theory of imprecise probability is *dilation* (Seidenfeld and Wasserman, 1993; Pedersen and Wheeler, 2014, 2015).

[8] Savage (1972). Gaifman (2013) discusses connections between some of the phenomena mentioned above. Also see Gaifman and Vasudevan (2012).

[9] Epstein and Le Breton (1993); Epstein and Schneider (2003).

Lemma 1. *Let $n \in \mathbb{N}$ and $f : \mathcal{S}_n \to [0,1]$. Then there exists a speed-optimal convergent inductive method that extends f if and only if f is dynamically coherent.*

Lemma 1 provides a partial answer to the question that we raised above. There are partial inductive methods that do not have speed-optimal extensions. In particular, any method that fails to be dynamically coherent cannot be extended in a speed-optimal way. We see that speed-optimality, then, induces dynamic coherence in the short-term *for all partial inductive methods with domains of the form $\mathcal{S}_n$.*

Removing this last proviso, so that the conclusion of Lemma 1 applies to *all* partial inductive methods, requires a modest generalization of the definition of dynamic coherence. Given $s, t \in \mathcal{S}$, we write $s \sqsubseteq t$ (resp. $s \sqsubset t$) if t is a (resp. strict) extension of s. Suppose $A \subseteq \mathcal{S}_n$. If $s \in \mathcal{S}_n$, we say that A *covers* s if for each t of length n which extends s, there exists $u \in A$ such that $s \sqsubset u \sqsubseteq t$. In other words, A covers s if a partial inductive method with domain A is guaranteed to always make another conjecture after reaching s. If A covers s, we write $c_A(s)$ for the set of $t \in A$ such that $s \sqsubset t$ and there does not exist any $u \in A$ such that $s \sqsubset u \sqsubset t$. That is, $c_A(s)$ is the set of possible "next times" a partial inductive method with domain A will make a conjecture after reaching s.

Finally, we say that a partial inductive method $f : A \to [0,1]$ is *dynamically coherent* if

$$\min_{t \in c_A(s)} f(t) \leq f(s) \leq \max_{t \in c_A(s)} f(t)$$

for all $s \in A$ such that A covers s. In other words, f is dynamically coherent if each (non-final) conjecture that it makes is contained in the interval spanned by the set of possible next conjectures. If $A = \mathcal{S}_n$, then A covers every $s \in \mathcal{S}_{n-1}$ and $c_A(s) = \{s0, s1\}$, so this is indeed a generalization of the previous definition of dynamic coherence. We are now able to state our main result.

Theorem 2. *Let A be a finite subset of $\mathcal{S}$ and $f : A \to [0,1]$. Then f extends to a speed-optimal, convergent inductive method if and only if f is dynamically coherent.*

This result provides a completely general answer to the question whether arbitrary short-term behavior is compatible with speed-optimality: the partial inductive methods that have speed-optimal extensions are exactly the dynamically coherent ones. In other words, speed-optimality induces dynamic coherence in the short-term. By strengthening Reichenbach's convergence criterion so that speed-optimality is required, one can avoid the objection that long-run requirements do not constrain short-term behavior. Dynamic coherence is necessary (and sufficient) in the short-term if the long-run goal of speed-optimal convergence is to be achieved.

Before concluding this section, we address a question that arises naturally in view of the preceding results. Is there anything more precise to be said about the relation between rigidity and dynamic coherence? There is. Roughly, we will show that rigidity on a larger domain of binary sequences than $\mathcal{C}_{\text{conv}}$ is *equivalent* to dynamic coherence. In other words, the two concepts are equivalent given the right domain of definition.

To show this let us say that an inductive method ϕ is *dynamically coherent* if the restriction of ϕ to $\mathcal{S}_n$ is dynamically coherent for all n. That is, ϕ is dynamically coherent if for all $s \in \mathcal{S}$

$$\phi(s0) \leq \phi(s) \leq \phi(s1) \text{ or } \phi(s1) \leq \phi(s) \leq \phi(s0).$$

Since each instance of dynamic coherence involves only finitely many values of ϕ, Theorem 2 implies that if ϕ is a speed-optimal convergent inductive method, then ϕ is dynamically coherent.

Example 3. Return to Example 2. There, ψ is not dynamically coherent because $\psi(0) = 0$ while $\psi(00) = \psi(01) = 0.5$. It follows from Theorem 2 that ψ is not speed-optimal, which we also showed above by pointing out that ξ is faster than ψ.

Next, let $\mathcal{C}_\phi \subseteq \mathcal{C}$ be the set of binary sequences σ such that $(\phi(\sigma^n))_{n \in \mathbb{N}}$ converges to a limit. Note that $\mathcal{C}_\phi \supseteq \mathcal{C}_{\mathrm{conv}}$ if ϕ is convergent. For all $s \in \mathcal{S}$, let

$$I_{\phi,s}^* = \bigcap_{\sigma \in \mathcal{C}_\phi} I_{\phi,s,\sigma},$$

and say that ϕ is $rigid^*$ if $I_{\phi,s}^* = \{\phi(s)\}$ for all s. Note that if ϕ is convergent, then $I_{\phi,s}^* \subseteq I_{\phi,s}$ since $\mathcal{C}_\phi \supseteq \mathcal{C}_{\mathrm{conv}}$.

Proposition 1. *A convergent inductive method is dynamically coherent if and only if it is rigid**.*

Since a convergent method is rigid* if it is rigid, Proposition 1 shows that we can view rigidity (or equivalently, speed-optimality) as nothing more than a slight strengthening of dynamic coherence. We do not currently know whether there are convergent inductive methods that are rigid* but not rigid. This leaves open the possibility that the two rigidity concepts are equivalent for convergent methods. In other words, it is an open question whether every convergent, dynamically coherent inductive method is speed-optimal. We will discuss this open question more in the next section.

5. Discussion

In addition to discussing the problems that our analysis has left open, we would like to conclude by drawing some connections between our results and probabilistic learning. This will raise some interesting possibilities for future research.

One of the most distinguished proponents of dynamic coherence in settings where agents' degrees of beliefs are represented by probability measures is Brian Skyrms.[10] A key insight of Skyrms's work in this area is that dynamically coherent degrees of belief form *martingales*. Martingales, in turn, have especially nice convergence properties. Using this fact, Skyrms has shown that dynamic coherence implies (almost surely) convergent degrees of belief—and this holds quite generally, without the assumption that beliefs change by Bayesian conditionalization, for instance. An important philosophical consequence of this result, emphasized by Skyrms, is that dynamic coherence rules out a particularly strong kind of inductive skepticism (Skyrms, 2014).[11] If one's beliefs are dynamically coherent, and so convergent, then one must expect that one's own beliefs will exhibit regularities in the long-run. In the presence of coherence, absolute skepticism—the view that nature exhibits no regularities whatsoever—is untenable.

One question that Skyrms does not explicitly answer is whether dynamic coherence is necessary for convergent degrees of belief in the probabilistic setting. In fact, it is not. This follows from a small body of mathematical literature produced in the 1970s that, so far as we know, has never been mentioned in philosophical work on convergence and coherence.[12] This literature shows that degrees of belief are convergent in Skyrms's sense just in case they are *martingales in the limit*, a property strictly weaker than being a martingale. So convergent

---

[10]Skyrms (1987, 1990, 1996, 2006). Also see Huttegger (2013, 2014, 2015, 2017b).

[11]Also see Diaconis and Skyrms (2017, ch. 10).

[12]In particular, this follows from results in Blake (1978). Also see Blake (1970); Mucci (1973, 1976); Edgar and Sucheston (1976, 1977).

degrees of belief need not be dynamically coherent. In view of the fact that convergence for degrees of belief is a strictly weaker property than dynamic coherence, the following question arises naturally: Is there a compelling notion of convergence for degrees of belief that strengthens Skyrms's notion and *does* imply dynamic coherence in the probabilistic sense? More specifically, is there a notion of speed-optimal convergence in the probabilistic setting that is sufficient for dynamic coherence? To the best of our knowledge, these questions are wide open.

These gaps in the probabilistic setting are, in a sense, dual to the ones that we have left open in our paper. To show this, we begin by remarking that there is a connection to be made with martingales in our framework as well. In our case, the relevant notion of martingale comes not from probability theory but the theory of algorithmic randomness. An inductive method ϕ is called a *martingale* if

$$\phi(s) = \frac{\phi(s0) + \phi(s1)}{2}$$

for all $s \in \mathcal{S}$. This notion of martingale was introduced by Jean Ville (1936; 1939) and plays an important role in contemporary studies of random binary sequences.[13]

It is clear from the definition that any inductive method that is a martingale is dynamically coherent in the sense of the previous section. As we also indicated in the previous section, our analysis has left open the question whether convergent martingales are necessarily speed-optimal. More generally, an important question for future research in our framework is: Are convergent, dynamically coherent inductive methods speed-optimal? Or, equivalently: Are convergent, rigid* methods rigid?

In the probabilistic setting, then, there is the question whether a notion of speed-optimal convergence is *sufficient* for coherence. And in the frequency prediction setting there is the question whether speed-optimal convergence is *necessary* for coherence. These questions, while independently interesting, are especially intriguing when considered together. We hope that future research will not only answer the open questions raised here but also shed light on unifying connections between induction in the probabilistic setting of Skyrms and induction in the frequency prediction setting of Reichenbach.[14]

6. CONCLUSION

In this paper, we have shown that a well-known objection to Reichenbach's criterion of inductive success—mere convergence—can be resisted by appealing to Juhl's notion of speed-optimality. Inductive methods that are speed-optimally convergent cannot behave arbitrarily in the short run, unlike inductive methods that are merely convergent. In order to secure speed-optimality, dynamic coherence is necessary. Dynamic coherence, in turn, is a substantive short-term constraint on inductive inference of considerable philosophical interest. Of special interest, to us, are the possible connections between this paper's results and known

---

[13]Nies (2009). Also see Shafer and Vovk (2005) and Bienvenu et al. (2009).

[14]An anonymous referee has suggested another promising avenue for future research. An alternative criterion to speed-optimality, which also strengthens Reichenbach's criterion of mere convergence, is the minimization of mind changes. This criterion has already proved itself to be a powerful tool in formal learning theory. See, for example, Kelly (1996) and Schulte (2018). It would be interesting to investigate the relation between minimizing mind changes and dynamic coherence. For example, is dynamic coherence necessary in order to minimize mind changes, as it is in order to secure speed-optimal convergence? We thank the referee for suggesting this question.

results concerning dynamic coherence for probabilistic learning. In the previous section, we suggested several open questions that we hope future research will be able to settle.

APPENDIX

PROOF OF THEOREM 1

First, suppose ϕ is a rigid convergent inductive method and ϕ' is any other convergent inductive method. Then $\phi(s) \neq \phi'(s)$ for some s. Since ϕ is rigid, there exists some $\sigma \in \mathcal{C}_{\text{conv}}$ extending s such that $\phi'(s) \notin I_{\phi,s,\sigma}$. It follows that $\epsilon_{\phi',s,\sigma} > \epsilon_{\phi,s,\sigma}$ and so ϕ' cannot be faster than ϕ.

Conversely, suppose ϕ is a convergent inductive method which is not rigid. Then for some s, $I_{\phi,s} = [a,b]$ is a nondegenerate interval.

Now let $t \in \mathcal{S}$ be a minimal finite extension of s such that $I_{\phi,s} \nsubseteq I_{\phi,t}$ (such a t exists since ϕ is convergent). Let u be t with its last bit removed; then u is also an extension of s.

Define

$$x = \begin{cases} \phi(t) & \phi(t) \in I_{\phi,s} \\ a & \phi(t) < a \\ b & \phi(t) > b. \end{cases}$$

Now define $\phi'(u) = x$ and $\phi'(v) = \phi(v)$ for all $v \neq u$. Note that since $x \in I_{\phi,s}$ and $I_{\phi,s} \subseteq I_{\phi,u}$ by minimality of t, $\neg[\phi \succ_{\sigma,\epsilon} \phi']$ for all $\sigma \in \mathcal{C}_{\text{conv}}$ and all $\epsilon > 0$. On the other hand, since $I_{\phi,s} \nsubseteq I_{\phi,t}$, there is some $\sigma \in \mathcal{C}_{\text{conv}}$ extending t such that $I_{\phi,s} \nsubseteq I_{\phi,t,\sigma}$. Note that $\phi(t) \in I_{\phi,t,\sigma}$. So, if $\phi(t) \in I_{\phi,s}$, then $\phi'(u) = \phi(t) \in I_{\phi,t,\sigma} = I_{\phi',t,\sigma}$ and so $I_{\phi',u,\sigma} = I_{\phi,t,\sigma} \nsupseteq I_{\phi,s}$. If $\phi(t) < a$, then $b \notin I_{\phi,t,\sigma}$ (otherwise $I_{\phi,t,\sigma}$ would contain all of $I_{\phi,s}$) and it follows that $b \notin I_{\phi',u,\sigma}$ since we defined $\phi'(u) = a$. Similarly, if $\phi(t) > b$, then $a \notin I_{\phi',u,\sigma}$.

So in all cases, we have $I_{\phi',u,\sigma} \nsupseteq I_{\phi,s}$, and in particular $I_{\phi',u,\sigma} \neq I_{\phi,u,\sigma}$ since $I_{\phi,u,\sigma} \supseteq I_{\phi,u} \supseteq I_{\phi,s}$. It follows that $I_{\phi',u,\sigma} \subset I_{\phi,u,\sigma}$, and therefore $\epsilon_{\phi',u,\sigma} < \epsilon_{\phi,u,\sigma}$. So, $\phi' \succ_{\sigma,\epsilon} \phi$ for $\epsilon = \epsilon_{\phi',u,\sigma}$. Thus ϕ' is faster than ϕ, and ϕ is not speed-optimal. $\square$

PROOF OF LEMMA 1

First, suppose that $\phi : \mathcal{S} \to [0,1]$ is a convergent inductive method extending f and that f is not dynamically coherent. Let $s \in \mathcal{S}_{n-1}$ witness that f is incoherent. We assume that $f(s0) < f(s)$ and $f(s1) < f(s)$, as the other case is similar. Let $a = \max(f(s0), f(s1))$. Then $[a, f(s)] = [a, \phi(s)] \subseteq I_{\phi,s}$ since any $\sigma \in \mathcal{C}_{\text{conv}}$ extending s must have $\sigma^{|s|+1} \in \{s0, s1\}$. Hence, ϕ is not rigid, and by Theorem 1, not speed-optimal.

Now suppose that f is dynamically coherent. Define $\phi : \mathcal{S} \to [0,1]$ by $\phi(s) = f(s)$ if $|s| \leq n$ and

$$\phi(s) = \frac{\sum_{i>n} s_i + f(s^n)}{|s| - n + 1}$$

if $|s| > n$. Then ϕ is convergent and extends f.

To prove that ϕ is rigid and hence speed-optimal, let $s \in \mathcal{S}$. If $|s| \geq n$, let $\sigma = s0000\ldots$ be the sequence obtained by extending s with all 0s. Then $\sigma \in \mathcal{C}_{\text{conv}}$ with $\ell(\sigma) = 0$ and the values $\phi(\sigma^m)$ are monotone decreasing for $m \geq |s|$. It follows that $I_{\phi,s,\sigma} = [0, \phi(s)]$. Similarly, if we take $\sigma' = s1111\ldots$, then $I_{\phi,s,\sigma'} = [\phi(s), 1]$. Thus $I_{\phi,s} \subseteq I_{\phi,s,\sigma} \cap I_{\phi,s,\sigma'} = \{\phi(s)\}$ and so $I_{\phi,s} = \{\phi(s)\}$.

In the case $|s| < n$, we use a similar argument but with different sequences. Since ϕ extends f and f is dynamically coherent, we can choose $a_1, a_2 \ldots, a_{n-|s|}$ such that

$$\phi(s) \geq \phi(sa_1) \geq \phi(sa_1a_2) \geq \cdots \geq \phi(sa_1a_2 \ldots a_{n-|s|}).$$

Taking $\sigma = sa_1a_2 \ldots a_{n-|s|}0000\ldots$, then we have as before that the values $\phi(\sigma^m)$ are monotone decreasing for $m \geq |s|$ and so $I_{\phi,s,\sigma} = [0, \phi(s)]$. Similarly, we can choose $b_1, b_2 \ldots, b_{n-|s|}$

such that
$$\phi(s) \le \phi(sb_1) \le \phi(sb_1b_2) \le \cdots \le \phi(sb_1b_2 \ldots b_{n-|s|})$$
and then $\sigma' = sb_1b_2 \ldots b_{n-|s|}1111\ldots$ satisfies $I_{\phi,s,\sigma'} = [\phi(s), 1]$. Thus again, we have $I_{\phi,s} = \{\phi(s)\}$.   □

Proof of Theorem 2

The proof requires a preliminary lemma.

In the proof of the lemma, it will be convenient to say f *is coherent at* s, by which we mean that $f(s)$ does not witness a counterexample to dynamic coherence, as defined in the main text.

Lemma 2. *Let* $A \subset \mathcal{S}_n$ *and let* $f : A \to [0,1]$ *be dynamically coherent. Then there exists* $s \in \mathcal{S}_n \setminus A$ *and an extension* $g : A \cup \{s\} \to [0,1]$ *of* f *which is dynamically coherent.*

Proof. Let $s \in \mathcal{S}_n \setminus A$ be minimal with respect to extension. If s is not the empty string, let $t \in \mathcal{S}_n$ be such that (without loss of generality) $s = t0$. By minimality of s, we must have $t \in A$. Note that $A \cup \{s\}$ will not cover any elements of $\mathcal{S}_n$ not covered by A, except possibly t. Moreover, if A covers u and $u \ne t$, then $c_{A \cup \{s\}}(u) = c_A(u)$. So to check that an extension $g : A \cup \{s\} \to [0,1]$ of f is coherent, we need only check coherence at s and at t.

To define g and prove it is coherent, we consider several cases.

First, suppose s is not covered by A. In that case, we define $g(s) = f(t)$, or we define $g(s)$ arbitrarily if s is the empty string. In this case $A \cup \{s\}$ still will not cover s. If $A \cup \{s\}$ covers t, note that $s \in c_{A \cup \{s\}}(t)$ and so since $g(s) = g(t)$, s witnesses the coherence of g at t.

Now suppose s is covered by A. Let I be the closed interval spanned by $f(c_A(s))$. As long as we define $g(s)$ to be some element of I, then g will be coherent at s. So if s is the empty string, we just define $g(s)$ to be any element of I.

If s is not the empty string, first suppose that t is not covered by A. Since $s = t0$ is covered by A but t is not covered by A, there must be some u of length n extending $t1$ such that there is no $v \in A$ with $t < v \le u$. But then this u witnesses that t is still not covered by $A \cup \{s\}$, so we may define $g(s)$ to be any element of I.

Finally, suppose t is covered by A. Since f is coherent at t, there exist $u, v \in c_A(t)$ such that $f(u) \le f(t) \le f(v)$. If u and v both extend $t1$, then $u, v \in c_{A \cup \{s\}}(t)$ as well, so we can define $g(s)$ to be any element of I. If u extends $t0$ and v extends $t1$, then $u \in c_A(s)$ and so $f(u) \in I$, so we can define $g(s) = f(u)$. Then g is coherent at t because $s, v \in c_{A \cup \{s\}}(t)$ and $g(s) \le g(t) \le g(v)$. Similarly if u extends $t1$ and v extends $t0$, we can define $g(s) = f(v)$. Finally, if u and v both extend $t0$, then $u, v \in c_A(s)$ and so $f(t) \in I$ since it is between $f(u)$ and $f(v)$. So, we may define $g(s) = f(t)$, and then g is coherent at t since $s \in c_{A \cup \{s\}}(t)$.   □

Proof of Theorem 2. Let $A \subseteq \mathcal{S}_n$. First, suppose $f : A \to [0,1]$ is dynamically coherent. If $A \ne \mathcal{S}_n$, then by Lemma 2, we can extend f to one more element of $\mathcal{S}_n$ while preserving its coherence. Iterating this, we may extend f to a partial inductive method $g : \mathcal{S}_n \to [0,1]$ which is coherent. By Lemma 1, we can then extend g to a speed-optimal convergent inductive method $\phi : \mathcal{S} \to [0,1]$.

Conversely, suppose f is not dynamically coherent, and let $s \in A$ witness the incoherence of f. Then A covers s and either $\min_{t \in c_A(s)} f(t) > f(s)$ or $\max_{t \in c_A(s)} f(t) < f(s)$. We write $a = \max_{t \in c_A(s)} f(t)$ and assume that $a < f(s)$ as the other case is similar.

Now suppose that ϕ is any convergent inductive method extending f. For any $\sigma \in \mathcal{C}_{\text{conv}}$ extending s, we have $\sigma^m \in c_A(s)$ for some $m > |s|$, since A covers s. We thus have $\phi(\sigma^m) =$

$f(\sigma^m) \leq a$ for some $m > |s|$. It follows that $[a, f(s)] = [a, \phi(s)] \subseteq I_{\phi,s}$. Hence ϕ is not rigid, and by Theorem 1, not speed-optimal. $\qquad\square$

Proof of Proposition 1

If a convergent inductive method $\phi : \mathcal{S} \to [0,1]$ is not dynamically coherent, then its restriction to some $\mathcal{S}_n$ is not dynamically coherent. The proof of Lemma 1, shows that ϕ is not rigid, but actually the same argument (applied to all $\sigma \in \mathcal{C}_\phi$ and not just all $\sigma \in \mathcal{C}_{\text{conv}}$) shows that ϕ is not rigid*.

Conversely, suppose a convergent inductive method $\phi : \mathcal{S} \to [0,1]$ is dynamically coherent. Let $s \in \mathcal{S}$; we wish to show $I^*_{\phi,s} = \{\phi(s)\}$. Since ϕ is dynamically coherent, there is some $a_1 \in \{0,1\}$ such that $\phi(sa_1) \leq \phi(s)$. There is similarly $a_2 \in \{0,1\}$ such that $\phi(sa_1a_2) \leq \phi(sa_1)$. Continuing by induction, we obtain a sequence σ extending s such that $\phi(\sigma^m) \leq \phi(\sigma^n)$ for all $m, n \geq |s|$ such that $m \geq n$.

Since the values $\phi(\sigma^m)$ form an eventually monotone sequence, they converge to some limit, so $\sigma \in \mathcal{C}_\phi$. Moreover, since $\phi(\sigma^m)$ is decreasing for $m \geq |s|$, the right endpoint of $I_{\phi,s,\sigma}$ is $\phi(s)$.

We may similarly construct a sequence σ' such that $\phi(\sigma'^m)$ is increasing for $m \geq |s|$ and so the left endpoint of $I_{\phi,s,\sigma'}$ is $\phi(s)$. Thus $I^*_{\phi,s} \subseteq I_{\phi,s,\sigma} \cap I_{\phi,s,\sigma'} = \{\phi(s)\}$ and so $I^*_{\phi,s} = \{\phi(s)\}$, as desired. $\qquad\square$

References

Bienvenu, L., G. Shafer, and A. Shen (2009). On the history of martingales in the study of randomness. *Electronic Journal for History of Probability and Statistics 5*(1), 1–40.

Blake, L. (1970). A generalization of martingales and two consequent convergence theorems. *Pacific Journal of Mathematics 35*(2), 279–283.

Blake, L. H. (1978). Every amart is a martingale in the limit. *Journal of the London Mathematical Society 2*(2), 381–384.

Carnap, R. (1950). *Logical Foundations of Probability.* University of Chicago Press.

Carnap, R. (1952). *The Continuum of Inductive Methods.* University of Chicago Press.

de Finetti, B. (1972). *Probability, induction, and statistics.* John Wiley & Sons.

Diaconis, P. and B. Skyrms (2017). *Ten Great Ideas about Chance.* Princeton University Press.

Dubins, L. E. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability 3*(1), 89–99.

Edgar, G. and L. Sucheston (1977). Martingales in the limit and amarts. *Proceedings of the American Mathematical Society 67*(2), 315–320.

Edgar, G. A. and L. Sucheston (1976). Amarts: A class of asymptotic martingales a. discrete parameter. *Journal of Multivariate Analysis 6*(2), 193–221.

Epstein, L. G. and M. Le Breton (1993). Dynamically consistent beliefs must be Bayesian. *Journal of Economic theory 61*(1), 1–22.

Epstein, L. G. and M. Schneider (2003). Recursive multiple-priors. *Journal of Economic Theory 113*(1), 1–31.

Gaifman, H. (2013). The sure thing principle, dilations, and objective probabilities. *Journal of Applied Logic 11*(4), 373–385.

Gaifman, H. and A. Vasudevan (2012). Deceptive updating and minimal information methods. *Synthese 187*(1), 147–178.

Huttegger, S. M. (2013). In defense of reflection. *Philosophy of Science 80*(3), 413–433.

Huttegger, S. M. (2014). Learning experiences and the value of knowledge. *Philosophical Studies 171*(2), 279–288.

Huttegger, S. M. (2015). Merging of opinions and probability kinematics. *The Review of Symbolic Logic 8*(04), 611–648.

Huttegger, S. M. (2017a). Analogical predictive probabilities. *Mind.*

Huttegger, S. M. (2017b). *The Probabilistic Foundations of Rational Learning.* Cambridge University Press.

Johnson, W. E. (1924). *Logic, Part III: The Logical Foundations of Science.* Cambridge University Press.

Johnson, W. E. (1932). Probability: The deductive and inductive problems. *Mind 41*(164), 409–423.

Juhl, C. F. (1994). The speed-optimality of Reichenbach's straight rule of induction. *The British Journal for the Philosophy of Science 45*(3), 857–863.

Kadane, J. B., M. J. Schervish, and T. Seidenfeld (1996). Reasoning to a foregone conclusion. *Journal of the American Statistical Association 91*(435), 1228–1235.

Kelly, K. T. (1996). *The Logic of Reliable Inquiry.* Oxford University Press.

Lane, D. A. and W. D. Sudderth (1984). Coherent predictive inference. *Sankhyā: The Indian Journal of Statistics, Series A: The Indian Journal of Statistics, Series A 46*(2), 166–185.

Lane, D. A. and W. D. Sudderth (1985). Coherent predictions are strategic. *The Annals of Statistics 13*(3), 1244–1248.

Mucci, A. G. (1973). Limits for martingale-like sequences. *Pacific Journal of Mathematics 48*(1), 197–202.

Mucci, A. G. (1976). Another martingale convergence theorem. *Pacific Journal of Mathematics 64*(2), 539–541.

Nies, A. (2009). *Computability and randomness.* Oxford University Press.

Pedersen, A. P. and G. Wheeler (2014). Demystifying dilation. *Erkenntnis 79*(6), 1305–1342.

Pedersen, A. P. and G. Wheeler (2015). Dilation, disintegrations, and delayed decisions. In *Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications*, pp. 227–236.

Reichenbach, H. (1938). *Experience and Prediction: An Analysis of the Foundations and the Structure of Knowledge.* University of Chicago press.

Reichenbach, H. (1949). *The Theory of Probability.* University of California Press.

Salmon, W. C. (1966). *The Foundations of Scientific Inference.* University of Pittsburgh Press.

Salmon, W. C. (1991). Hans reichenbach's vindication of induction. *Erkenntnis 35*(1-3), 99–122.

Savage, L. (1972). *The Foundations of Statistics.* New York: John Wiley & Sons.

Schervish, M. J., T. Seidenfeld, and J. B. Kadane (1984). The extent of non-conglomerability of finitely additive probabilities. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 66*(2), 205–226.

Schulte, O. (1999a). The logic of reliable and efficient inquiry. *Journal of Philosophical Logic 28*(4), 399–438.

Schulte, O. (1999b). Means-ends epistemology. *The British Journal for the Philosophy of Science 50*(1), 1–31.

Schulte, O. (2018). Formal learning theory. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy.*

Seidenfeld, T. and L. Wasserman (1993). Dilation for sets of probabilities. *The Annals of Statistics 21*(3), 1139–1154.

Shafer, G. and V. Vovk (2005). *Probability and finance: it's only a game!*, Volume 491. John Wiley & Sons.

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

Skyrms, B. (1990). *The dynamics of rational deliberation.* Harvard University Press.

Skyrms, B. (1996). The structure of radical probabilism. *Erkenntnis 45*(2-3), 285–297.

Skyrms, B. (2006). Diachronic coherence and radical probabilism. *Philosophy of Science 73*(5), 959–968.

Skyrms, B. (2014). Grades of inductive skepticism. *Philosophy of Science 81*(3), 303–312.

van Fraassen, B. C. (1984). Belief and the will. *The Journal of Philosophy 81*(5), 235–256.

van Fraassen, B. C. (1999). Conditionalization, a new argument for. *Topoi 18*(2), 93–96.

van Fraassen, B. C. (2000). The false hopes of traditional epistemology. *Philosophical and Phenomenological Research LX*(2), 253–280.

van Fraassen, B. C. and J. Y. Halpern (2016). Updating probability: Tracking statistics as criterion. *The British Journal for the Philosophy of Science 68*(3), 725–743.

Ville, J. (1936). Sur la notion de collectif. *Comptes rendus des Séances de l'Académie des Sciences 203*, 26–27.

Ville, J. (1939). *Etude critique de la notion de collectif.* Gauthier-Villars Paris.