Why be (approximately) coherent?

Glauber De Bona and Julia Staffel

Introduction

Bayesian epistemology proposes norms on degrees of belief that are supposed to constitute rational ideals. The most widely endorsed norm is probabilism, which requires ideally rational agents to have degrees of belief that can be represented by a probability function. Unfortunately, probabilistic coherence is unattainable for human thinkers, because fully complying with the norm is too difficult for us. In response, Bayesians suggest that for limited thinkers, probabilistic coherence is an ideal to be approximated. We are supposedly better off the more closely our credences approximate the ideal. However, it is rarely discussed exactly in what sense credences are better if they approximate coherence more closely. In this paper, we first clarify the way in which approximating coherence needs to be beneficial in order for probabilism to constitute an ideal in the intended sense. In Section 2, we present existing results from the literature that support the idea that probabilism is an ideal that should be approximated: On some measures of incoherence, being less incoherent reduces vulnerability to Dutch books. Furthermore, given certain other incoherence measures, some ways of being less incoherent have guaranteed benefits for the accuracy of one's credences. The problem is that these known results rely on different ways of measuring closeness to coherence. Hence, the existing results don't settle whether, for a given incoherent credence function, there always are alternative, less incoherent credence functions that are both better from an accuracy perspective and from

1

a Dutch book perspective. In Section 3, we show that is is in fact always possible, for any given credence function, to identify a series of alternative credence functions that are less incoherent, and guarantee simultaneous improvements in accuracy and DB vulnerability. This is good news for the Bayesian project, because it strengthens the claim that Bayesian ideal norms apply to non-ideal agents as regulative ideals.[1]

1 Bayesian Ideal Norms

Bayesians maintain that norms of ideal rationality apply to non-ideal thinkers in the sense that a non-ideal epistemic state is better the more closely it approximates this ideal. For this line of argument to be plausible, Bayesians need to specify what it means to approximate a rational ideal like coherence, and why doing so is beneficial. Earman makes this point quite forcefully in *Bayes or Bust?*: "The response that Bayesian norms should be regarded as ideals towards which we should strive even if we always fall short is idle puffery unless it is specified how we can take steps to bring us closer to these goals." (Earman 1992)

Reflection on ideal scenarios in other domains reveals that is not always true that being closer to an ideal is actually better. Compare the ideal of being fluent in Portuguese to the ideal of being the first choice for your dream job. Being the first choice for your dream job is clearly an ideal situation. You would benefit from it by having the opportunity to pursue a stimulating career, to earn money by doing something you love, etc. However, there are natural ways of conceiving of approximations to this ideal that don't get you an increasing share of these benefits. If you are the second choice for the position, you get no more of a portion of the benefit associated with the ideal state than if you are the fifth choice for the position. The benefits don't obtain in any of the non-

---

[1]Our result demonstrates that it is possible to jointly improve accuracy and reduce DB vulnerability by approximating coherence, but it is not intended as advice for non-ideal agents about how to improve, or as a reasoning strategy that non-ideal reasoners can implement.

ideal situations, even if some of them are intuitively closer to the ideal state than others. By contrast, suppose that your ideal is to be fluent in Portuguese. Benefits of fluency in Portuguese include being able to talk to other Portuguese speakers, understanding Portuguese literature and films, etc. Approximating the ideal of being fluent clearly gives you an increasing portion of the benefits associated with being completely fluent. The closer you are to being fluent, the better you can converse with other Portuguese speakers, and understand films and literature.

Bayesians clearly think of coherence as being analogous to the language rather than the job example. Yet, this view needs to be defended. There is a large literature on why probabilistic coherence is a rational ideal, in which it is argued that being probabilistically coherent is beneficial. The most prominent results show that having probabilistically coherent degrees of belief benefits both the accuracy of one's attitudes and their ability to guide actions. The accuracy-dominance argument for probabilism shows that, given a suitable measure of inaccuracy, if an agent has incoherent credences, there is an alternative coherent credence function that is closer to the truth no matter what the truth is. If an agent has coherent credences, her attitudes are not accuracy-dominated in this way. The Dutch book argument for probabilism shows that incoherent, but not coherent, degrees of belief justify betting in ways that lead to guaranteed losses.

Hence, coherent credences are better suited than incoherent credences to play two characteristic roles of belief states, namely accurately representing the world and guiding action. If coherence is supposed to be an ideal in the sense in which being fluent in Portuguese is an ideal, then approximating coherence must deliver an increasing portion of the benefits that fully coherent credences guarantee.

2 Reducing Dutch Book Vulnerability and Improving Accuracy

Some research has already been done on whether approximating coherence leads to decreased Dutch book-vulnerability, and also on whether it leads to improved accuracy. To state the results, we define a credence function as a function $c : F \to [0, 1]$ that assigns a real number from the interval $[0, 1]$ to each proposition in a finite set of propositions $F \subseteq \mathcal{L}$, where $\mathcal{L}$ is a propositional language defined in the usual way based on a set of atomic statements $\{A_i\}$ and the standard Boolean connectives $\neg, \vee, \wedge$. A *coherent* credence function is a credence function whose credence assignments agree with some probability function on $\mathcal{L}$, whereas an *incoherent* credence function lacks this property. [2]

A general method for measuring the degree to which an incoherent credence function approximates coherence is by representing credence functions as vectors, and measuring the distance between the vector corresponding to an incoherent credence function and a vector corresponding to some closest coherent credence function. For any credence function $c : \{\varphi_1, \ldots, \varphi_m\} \to [0, 1]$ such that $c(\varphi_i) = q_i$ for $1 \leq i \leq m$, its corresponding vector is $q = \langle q_1, \ldots, q_m \rangle$.

In principle, any divergence can be used to formulate such an incoherence measure, where a divergence is defined as any function $d : [0, 1]^n \times [0, 1]^n \to [0, \infty)$, over vectors $X, Y \in [0, 1]^n$, where $d(X, Y) = 0$ iff $X = Y$.

We can then define an incoherence measure based on some divergence d as follows:

$$\mathcal{I}_d(c) = \min\{d(c, c')|c' \text{ is coherent}\}$$

Here, c and c' are credence functions that are defined over the same set of propositions, and their divergence is measured based on their corresponding vectors.

---

[2] Here and in what follows, we are considering only unconditional, point-valued credence functions. The results by SSK (2002, 2003, 2000) and De Bona and Finger (2015) also extend to imprecise and conditional credences, but since the accuracy-results by De Bona and Staffel (2017) apply to precise, unconditional credences, the results presented here are limited to those cases as well.

A subset of divergences are metrics, which are the most familiar type of distance measures. They are characterized by being symmetrical and satisfying the triangle inequality. The p-norm distance, which is a metric, between two vectors $X = \langle x_1, x_2, \ldots, x_n \rangle$ and $Y = \langle y_1, y_2, \ldots, y_n \rangle$ is defined as the p-norm of their difference, with $p \in [1, \infty)$:

$$d_p(X, Y) = \|X - Y\|_p = (\sum_{i=1}^{n} |x_i - y_i|^p)^{1/p}$$

For example, we can set $p = 1$, which gives us the absolute distance measure, or $p = 2$, which gives us Euclidean distance, and so on. Taking the limit $p \to \infty$, we have $d_\infty(X, Y) = \max_i |x_i - p_i|$, which gives us Chebyshev distance.

It has been found that incoherence measures based on two particular distance measures have natural interpretations in terms of Dutch book vulnerability (De Bona and Finger 2015). A Dutch book is a collection of bets that leads to a guaranteed loss. For an agent whose credence function c is defined over F, we can represent a bet on or against a proposition $\varphi_i \in F$ based on the agent's credence in φ_i as follows:

$$(Ind(\varphi_i) - c(\varphi_i))\alpha$$

The indicator function $Ind(\varphi_i)$ assigns a value of 1 if φ_i is true and 0 if false. The coefficient α (the bet stake) determines the size and direction of the bet. If $\alpha > 0$, then the agent bets on the truth of φ_i; if $\alpha < 0$, the agent bets on its falsity. An agent is said to be Dutch-bookable if a clever bookie can set up a betting arrangement based on her credences that lead to a guaranteed loss no matter what the world is like. More formally, if we can choose coefficients $\alpha_1, ..., \alpha_n$ (where α_i can be 0), and there is a constant $\epsilon > 0$ so that the sum of the payoffs of the bets based on the agent's credence function c on F is lower than $-\epsilon$ for every possible state of the world $s \in \mathcal{S}$, then the agent is vulnerable to a Dutch book. The guaranteed loss G that results from a set of bets based on credences in c with coefficients $\alpha_1, ..., \alpha_n$ is then:

$$G(c) = -min\{0, sup_{s \in \mathcal{S}} \sum_{i=1}^{n} \alpha_i (Ind_{\varphi_i}(s) - c(\varphi_i))\}$$

Since the coefficients are not constrained, betting losses from Dutch books can be arbitrarily large. To be useable for measuring incoherence, the Dutch book loss from a set of bets must be normalized. Two normalizations that have been proposed by Schervish, Seidenfeld, and Kadane (SSK) are of particular interest for our purposes (2002, 2003, 2000). Their *neutral/max* measure determines the incoherence of a credence function by choosing coefficients $\alpha_1, ..., \alpha_n$ that maximize the following quantity: $\frac{G(c)}{max\{|\alpha_1|, ..., |\alpha_n|\}}$. Their *neutral/sum* measure instead requires maximizing this quantity: $\frac{G(c)}{\sum_{i=1}^{n} |\alpha_i|}$. In more intuitive terms, the neutral/max measure gathers and adds up incoherence anywhere in the agent's credence function, whereas the neutral/sum measure detects the subset of the agent's credences that contains the most dramatic incoherence, and measures the Dutch book vulnerability of the credences in this set.[3]

Each of these Dutch book measures correspond to an incoherence measure based on a p-norm distance (De Bona and Finger 2015). The *neutral/max* measure is equivalent to determining the incoherence of a credence function as its absolute distance ($p = 1$) to some closest coherent credence function. The *neutral/sum* measure is equivalent to determining the incoherence of a credence function as its Chebyshev distance ($p = \infty$) to some closest coherent credence function. Hence, the closer a credence function is to coherence according to the absolute (Chebyshev) distance measure, the lower its normalized Dutch book loss according to the *neutral/max* (*neutral/sum*) measure. This result gives a partial vindication to the idea that approximating coherence is valuable: if closeness to coherence is measured with the appropriate p-norm distance mea-

---

[3]SSK (2002) not only introduce Dutch book measures of incoherence based on a variety of possible normalizations, they also show under which conditions the extent of incoherence is a continuous function of the bets and previsions being considered, and under which conditions their incoherence measures preserve dominance relations between incoherent credence functions. SSK characterize dominance as follows: Suppose c and c' are incoherent assignments of previsions to the same set of random variables. Furthermore, assume that γ is a gambling strategy that determines for each prevision how much is gambled on the associated bet. c dominates c' with respect to γ just in case the guaranteed loss that can be extracted from c' is greater than the guaranteed loss resulting from c in every possible state of the world. SSK go on to determine which of their Dutch book measures of incoherence preserve dominance. The neutral/max and neutral/sum measures we focus on here both do so, among others.

sures, then approximating coherence has the benefit of reduced Dutch book vulnerability.

Coherence also has a different benefit - coherent credences are not accuracy-dominated, provided a suitable measure of accuracy. Does approximating coherence lead to better accuracy? De Bona and Staffel (2017) show that particular ways of reducing incoherence lead to guaranteed improvements in the accuracy of one's credences. However, as explained below, their results depend on using a set of divergences to measure incoherence that contains neither absolute nor Chebyshev distances.

Suppose inaccuracy is measured in the usual way with an additive inaccuracy measure $\mathfrak{I}$ that is based on a continuous, strictly proper scoring rule, and assume moreover that the inaccuracy measure is convex.[4] Let $\mathfrak{I}(c, c')$ denote the expected inaccuracy of a credence function c relative to some coherent credence function c'. Every such inaccuracy measure $\mathfrak{I}$ defines a divergence

$$d_{\mathfrak{I}}(c, c') = \mathfrak{I}(c, c') - \mathfrak{I}(c', c')$$

In other words, $d_{\mathfrak{I}}(c, c')$ measures the expected gain in inaccuracy, according to a coherent c', due to adopting c instead of c'. Suppose then that inaccuracy is measured with some inaccuracy measure $\mathfrak{I}$ as defined above, and incoherence is measured with $\mathcal{I}_{d_{\mathfrak{I}}}$ as the divergence from some closest coherent credence function, where the divergence[5] is generated by its corresponding inaccuracy measure as just described. Under those circumstances, the following holds (De Bona and Staffel 2017):

Proposition 1. *If c is an incoherent credence function on F, and c^* is some $d_{\mathfrak{I}}$-*

---

[4]For a characterization of inaccuracy measures, see (Predd, Seiringer, Lieb, Osherson, Poor, and Kulkarni 2009). Their argument for probabilism doesn't require convexity. However, De Bona and Staffel's result does. For an argument that it is desirable to use convex inaccuracy measures, see (Joyce 2009).The importance of the continuity requirement is illustrated by SSK (2009).

[5]As divergences are assumed to be finite, the technical results presented here in principle exclude the logarithmic scoring rule and the associated KL-divergence $d_{KL}(\cdot, \cdot)$. However, all results still hold for this pair if we focus on credence functions c with $\mathcal{I}_{d_{KL}}(c) < \infty$, which is the case when $c(\varphi) \in (0, 1)$ for all $\varphi \in F$, for instance.

closest coherent credence function to c, then c^* strongly $\mathfrak{I}$-accuracy-dominates
c. Moreover, for any $\lambda \in (0,1)$, let c_λ be a credence function on F such that
$c_\lambda(\varphi) = (1-\lambda)c(\varphi) + \lambda c^*(\varphi)$ for all $\varphi \in F$; then c_λ strongly $\mathfrak{I}$-accuracy-
dominates c for any $\lambda \in (0,1)$.

In other words, De Bona and Staffel show that for any incoherent credence
function c, the credence functions that lie on the direct path between c and
some closest coherent credence function to c are more accurate in every possible
world, given a suitable inaccuracy measure $\mathfrak{I}$.

These credence functions that are more accurate than c are also less incoher-
ent than c, provided we use a suitable divergence in our measure of incoherence.
This is captured by Proposition 2 (proved in the appendix):

Proposition 2. *Let $\mathfrak{I}$ be a convex, additive inaccuracy measure that is based on*
a continuous, strictly proper scoring rule, c, an incoherent credence function on
F, and c^*, some $d_{\mathfrak{I}}$-closest coherent credence function to c. For any $\lambda \in (0,1]$,
let c_λ be a credence function on F such that $c_\lambda(\varphi) = (1-\lambda)c(\varphi) + \lambda c^*(\varphi)$ for
all $\varphi \in F$; then $\mathcal{I}_{d_J}(c_\lambda) < \mathcal{I}_{d_J}(c)$ for any $\lambda \in (0,1]$.

Hence, we know two interesting facts: (i) For any incoherent credence func-
tion c, any alternative credence functions that are closer to coherence according
to the absolute (or Chebyshev) distance measure are also less vulnerable to
Dutch books according to the *neutral/max* (*neutral/sum*) measure. (ii) For any
incoherent credence function c, any alternative credence functions that lie on
the direct path between c and some closest coherent credence function as mea-
sured by a suitable divergence $d_{\mathfrak{I}}$ are less incoherent than c according to the
incoherence measure $\mathcal{I}_{d_{\mathfrak{I}}}$, and also more accurate than c in every possible world
according to the associated inaccuracy measure $\mathfrak{I}$. The problem is that the two
results rely on different incoherence measures, so we can't be sure that, for a
given incoherent credence function c, there are always alternative, less incoher-
ent credence functions that are both less vulnerable to Dutch books and more
accurate in every possible worlds.

We're thus interested in investigating the following question *Q1*: For a given inaccuracy measure $\mathfrak{I}$, is there a way of approximating coherence that guarantees both a gain in $\mathfrak{I}$-accuracy in every possible world, *and* a reduction in Dutch book vulnerability? On the simplest understanding, a "way of approximating coherence" refers to a reduction of incoherence according to a *way* of measuring it; i.e. there is one way of approximating coherence for each divergence we can use to measure incoherence. As De Bona and Staffel (2017) have shown, if "a way of approximating coherence" is understood as a reduction in d-incoherence for some divergence d, the answer to *Q1* is "no": there are no combinations of continuous inaccuracy measure $\mathfrak{I}$ and divergence d, such that a reduction of d-incoherence is guaranteed to lead to improved $\mathfrak{I}$-accuracy in every possible world. However, we've learned above that if we understand a "way of approximating coherence" more generally - in a sense that includes moving on the direct path to some $d_{\mathfrak{I}}$-closest coherent credence function - then there is a way of approximating coherence that is guaranteed to improve accuracy. Hence, we will focus on the more specific question *Q2*: For a given inaccuracy measure $\mathfrak{I}$ that is additive, convex, and based on a continuous, strictly proper scoring rule, does moving towards a $d_{\mathfrak{I}}$-closest coherent credence function decrease Dutch book vulnerability?

3 We Can Have It All

Fortunately for Bayesian, we can answer *Q2* affirmatively, which implies a positive answer to the more general question *Q1*, assuming we're focusing on convex, additive inaccuracy measures that are based on continuous, strictly proper scoring rules. We can show that, for any incoherent credence function c, any credence function c_λ is less $\mathfrak{I}$-inaccurate than c and less Dutch book vulnerable than c according to both the neutral/sum and the neutral/max measures, provided that c_λ lies on the direct path towards some closest coherent credence function according to some divergence $d_{\mathfrak{I}}$ that is generated by a suitable inac-

curacy measure $\mathfrak{I}$.

This result is based on the following more general finding (proved in the appendix):

Proposition 3. *Let c be an incoherent credence function on F, and c^* be an arbitrary coherent credence function on F. For any λ within the interval $(0,1]$, let c_λ be a credence function such that $c_\lambda(\varphi) = (1-\lambda)c(\varphi) + \lambda c^*(\varphi)$ for all $\varphi \in F$. Then, for any p-norm distance measure d_p and any $\lambda \in (0,1]$: $\mathcal{I}_{d_p}(c_\lambda) < \mathcal{I}_{d_p}(c)$.*

Proposition 3 helps us answer Q2 as follows: Due to Proposition 1, we know that, for any incoherent credence function c, any credence function c_λ that lies on the direct path between c and some $d_\mathfrak{I}$-closest coherent credence function c^* is more $\mathfrak{I}$-accurate than c in every world. Due to Proposition 2, we also know that c_λ is less incoherent than c, provided that we measure incoherence with $\mathcal{I}_{d_\mathfrak{I}}$. Due to Proposition 3, we know that c_λ is also less incoherent than c if we measure incoherence as either the absolute or the Chebyshev distance to some closest coherent credence function. It then follows due to the results proven in (De Bona and Finger 2015) that c_λ is less vulnerable to Dutch books than c according to both the neutral/max and neutral/sum measures of Dutch book vulnerability. Hence, we know that, for a given incoherent credence function c, we can always find alternative credence functions that are more accurate in every possible world and less vulnerable to Dutch books. These alternative credence functions are less incoherent than c on any p-norm distance measure of incoherence, as well as the on the incoherence measure $\mathcal{I}_{d_\mathfrak{I}}$ that is based the relevant inaccuracy measure $\mathfrak{I}$.

Example 1. *Suppose an agent has an incoherent credence function $c(p) = 0.3$ and $c(\neg p) = 0.4$. Using the Brier score $\mathfrak{B}$ as our inaccuracy measure, the $d_\mathfrak{B}$-closest coherent credence function is $c^*(p) = 0.45$ and $c^*(\neg p) = 0.55$; $d_\mathfrak{B}$ is the squared Euclidean distance. The degree of incoherence of c as measured by the increase of expected $\mathfrak{B}$-inaccuracy relative to c^* is $I_{d_\mathfrak{B}}(c) = d_\mathfrak{B}(c, c^*) = 0.045$. c^* is also a closest coherent credence function to c according to the*

absolute and Chebyshev distance measures. The degrees of incoherence of c according to these measures are $\mathcal{I}_{d_1}(c) = 0.3$ and $\mathcal{I}_{d_\infty}(c) = 0.15$. We can now examine the properties of $c_\lambda(p) = 0.375$ and $c_\lambda(\neg p) = 0.475$. c_λ lies on the direct path between c and c^* ($\lambda = 0.5$), and it is less incoherent than c on all three measures: $\mathcal{I}_{d_1}(c) = 0.15$, $\mathcal{I}_{d_\infty}(c) = 0.075$, $\mathcal{I}_{d_\mathfrak{B}}(c) = 0.0.01125$. As expected, c_λ $\mathfrak{B}$-dominates c. It also has a lower normalized Dutch book loss: $neutral/max(c_\lambda) = 0.15 < neutral/max(c) = 0.3$, and $neutral/sum(c_\lambda) = 0.075 < neutral/max(c) = 0.15$.

4 Conclusion

This result is very good news for the Bayesian project. Bayesians claim that ideal norms of rationality should be approximated by non-ideal reasoners. To argue for this, they must show that there is in fact a way of approximating the ideal that confers an increasing share of the benefits associated with perfect norm compliance. The most commonly cited benefits of probabilistic coherence are that coherent credence functions are not accuracy-dominated, and that they are not vulnerable to Dutch books. We showed that it is indeed the case that for any incoherent credence function, there always are alternative, less incoherent credence functions that are both more accurate in every possible world according to some suitable inaccuracy measure $\mathfrak{I}$, and less vulnerable to Dutch books on either the neutral/max or neutral/sum measure.

Appendix

Proposition 2 . *Let $\mathfrak{I}$ be a convex, additive inaccuracy measure that is based on a continuous, strictly proper scoring rule, c, an incoherent credence function on F, and c^*, some $d_\mathfrak{I}$-closest coherent credence function to c. For any $\lambda \in (0, 1]$, let c_λ be a credence function on F such that $c_\lambda(\varphi) = (1 - \lambda)c(\varphi) + \lambda c^*(\varphi)$ for all $\varphi \in F$; then $\mathcal{I}_{d_J}(c_\lambda) < \mathcal{I}_{d_J}(c)$ for any $\lambda \in (0, 1]$.*

Proof. Since $\mathfrak{I}$ is convex, $d_{\mathfrak{I}}(\cdot, \cdot)$ is convex in its first argument. Thus:

$$d_{\mathfrak{I}}(c_{\lambda}, c^*) \leq (1 - \lambda)d_{\mathfrak{I}}(c, c^*) + \lambda d_{\mathfrak{I}}(c^*, c^*) = (1 - \lambda)d_{\mathfrak{I}}(c, c^*)$$

From $\mathcal{I}_{d_{\mathfrak{I}}}(c_{\lambda}) \leq d_{\mathfrak{I}}(c_{\lambda}, c^*)$, it follows that, for any $\lambda \in (0, 1]$:

$$\mathcal{I}_{d_{\mathfrak{I}}}(c_{\lambda}) \leq (1 - \lambda)d_{\mathfrak{I}}(c, c^*) < d_{\mathfrak{I}}(c, c^*) = \mathcal{I}_{d_{\mathfrak{I}}}(c)$$

$\square$

Proposition 3 . *Let c be an incoherent credence function on F, and c^* be an arbitrary coherent credence function on F. For any λ within the interval $(0, 1]$, let c_{λ} be a credence function such that $c_{\lambda}(\varphi) = (1-\lambda)c(\varphi)+\lambda c^*(\varphi)$ for all $\varphi \in F$. Then, for any p-norm distance measure d_p and any $\lambda \in (0, 1]$: $\mathcal{I}_{d_p}(c_{\lambda}) < \mathcal{I}_{d_p}(c)$.*

Proof. Let q, q^* and q^{λ} be the vectors in $[0, 1]^m$, with $m = |F|$, corresponding to the credence functions c, c^* and c_{λ}, respectively. It follows that $q^{\lambda} = (1 - \lambda)q + \lambda q^*$. For a given positive $p \in \mathbb{N} \cup \{\infty\}$, let $q^p \in [0, 1]^m$ be the vector corresponding to a coherent credence functions c^p such that $\mathcal{I}_{d_p}(c) = d_p(c, c^p)$ – i.e., c^p is a coherent credence function that is closest to c according to d_p. Consider an arbitrary $\lambda \in (0, 1]$. Let $q^{p,\lambda} = (1 - \lambda)q^p + \lambda q^*$. Since the set of vectors corresponding coherent credence functions is convex, $q^{p,\lambda}$ corresponds to coherent credence function, say $c^{p,\lambda}$. Hence, $\mathcal{I}_{d_p}(c^{\lambda}) \leq d_p(c^{\lambda}, c^{p,\lambda})$. Furthermore, we have:

$$
\begin{aligned}
d_p(c^{\lambda}, c^{p,\lambda}) &= \|q^{\lambda} - q^{p,\lambda}\|_p = \|(1 - \lambda)q - (1 - \lambda)q^p\|_p \\
&= (1 - \lambda)\|q - q^p\|_p = (1 - \lambda)d(c, c^p) = (1 - \lambda)\mathcal{I}_{d_p}(c)
\end{aligned}
\tag{1}
$$

Finally, as $\lambda \in (0, 1]$, $\mathcal{I}_{d_p}(c^{\lambda}) \leq d_p(c^{\lambda}, c^{p,\lambda}) < \mathcal{I}_{d_p}(c)$. $\square$

References

De Bona, G. and M. Finger (2015). Measuring inconsistency in probabilistic logic: rationality postulates and Dutch book interpretation. *Artificial Intelligence 227*, 140–164.

De Bona, G. and J. Staffel (2017). Graded incoherence for accuracy-firsters. *Philosophy of Science 84*(2), 189–213.

Earman, J. (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. MIT Press.

Joyce, J. M. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber and C. Schmidt-Petri (Eds.), *Degrees of belief*, Volume 342 of *Synthese Library*, pp. 263–297. Springer.

Predd, J. B., R. Seiringer, E. H. Lieb, D. N. Osherson, H. V. Poor, and S. R. Kulkarni (2009). Probabilistic coherence and proper scoring rules. *IEEE Transactions on Information Theory 55*(10), 4786–4792.

Schervish, M. J., J. B. Kadane, and T. Seidenfeld (2003). Measures of incoherence: How not to gamble if you must. In *Bayesian Statistics 7: Proceedings of the 7th Valencia Conference on Bayesian Statistics*, pp. 385–402.

Schervish, M. J., T. Seidenfeld, and J. B. Kadane (2000). How sets of coherent probabilities may serve as models for degrees of incoherence. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 8*(03), 347–355.

Schervish, M. J., T. Seidenfeld, and J. B. Kadane (2002). Measuring incoherence. *Sankhyā: The Indian Journal of Statistics, Series A 64*, 561–587.

Schervish, M. J., T. Seidenfeld, and J. B. Kadane (2009). Proper scoring rules, dominated forecasts, and coherence. *Decision Analysis 6*(4), 202–221.