HOW TO SERVE TWO EPISTEMIC MASTERS

Leszek Wroński* Zalán Gyenis*

April 17, 2024

Abstract

We extend a result by Gallow concerning the impossibility of following two epistemic masters, so that it covers a larger class of pooling methods. We also investigate a few ways of avoiding the issue, such as using nonconvex pooling methods, employing the notion of imperfect trust or moving to higher-order probability spaces. Along the way we suggest a conceptual issue with the conditions used by Gallow: whenever two experts are considered, whether we can trust one of them is decided by the features of the other!

Keywords: credence, deference.

1 INTRODUCTION

In this article we will extend [5]'s impossibility result concerning complete deference to two experts. Gallow showed that full trust in each expert's opinion when considered separately is incompatible with linear averaging of their opinions when considered together. We present a new proof of this result and show that this behaviour is displayed if we consider any convex pooling method. We will also point to some limitations of Gallow's conditions.

Where Gallow uses "serves", we frequently use "trusts" or "defers to". This is meant to only be a terminological difference. Similarly with "fully" and "perfectly"; the expressions "fully trusts" and "perfectly serves" are synonymous for the purposes of this paper. "Propositions" (understood as sets) and "events" will also be used interchangeably here.

2 CONVEXITY AND EXPERT DEFERENCE

We will employ the approach from [5], where we're interested in a single proposition r (e.g., that it will rain this afternoon). Since we're after an impossibility result, one proposition of interest suffices: if we can't trust two experts' opinions about *one* thing, it's natural to think we *a fortiori* can't do so if *more* things are involved. Regarding the proposition r, then, we seek guidance from two forecasters Al and Bert and we weigh their opinions by α and $\beta = 1 - \alpha$, respectively. If our own degrees of confidence are given by the function P, A is a random variable whose value is Al's forecast of rain, and B is a random variable whose value is Bert's forecast of rain, then it might be natural to stipulate that P should be such that:

*Faculty of Philosophy, Jagiellonian University. Corresponding author's email address: leszek.wronski@uj.edu.pl.

1. $P(A = B) < 1$;
2. for any a, $P(r|A = a) = a$ ("P (perfectly) trusts A");
3. for any b, $P(r|B = b) = b$ ("P (perfectly) trusts B");
4. for any a, b, $P(r|A = a, B = b) = \alpha a + \beta b$.

However, if P is a probability, it cannot satisfy all these requirements at the same time. We will show a new proof of this result, in a setting which is more general in one respect: Gallow's condition 4. above has P use linear averaging, while we will talk about *strictly convex pooling methods*, introduced here after a few introductory remarks.

We diverge from Gallow in explicitly considering the sets of experts 'options', OP_A and OP_B. These are sets of numbers x such that $x \in \mathrm{OP}_Z$ iff the proposition $Z = x$ is in the domain of P. We assume these sets are finite, linearly ordered and labelled so that the lowest index signifies the lowest number. That is:

$$\mathrm{OP}_A = \{a_1, \dots, a_n\} \subset \mathbb{R} \text{ such that } i < j \text{ implies } a_i < a_j;$$

$$\mathrm{OP}_B = \{b_1, \dots, b_m\} \subset \mathbb{R} \text{ such that } i < j \text{ implies } b_i < b_j.$$

As for the finiteness assumption, it allows us the specific way of proving our result. However, it does not allow us to recover the full force of Gallow's argument for the specific case of linear averaging. Still, we do not think this is a big drawback. If the values of P are not chosen in a very deliberate way, there will only be a finite number of propositions of the form $A = x$ to which P assigns a nonzero value, and so in all other cases the relevant conditional probabilities will be undefined.

We will say that A **ranges below** B iff $a_1 < b_1$. We will say that A **ranges above** B iff $a_n > b_n$. In the context of r, A and B, we will say that P **uses a strictly convex pooling method** iff for any a different from b, $P(r|A = a, B = b) \in (\min\{a, b\}, \max\{a, b\})$, and $P(r|A = a, B = a) = a$.

We have assumed the agent is interested in a single r; from a different perspective, they are interested in whether r or $\neg r$, that is, in which member of a certain partition turns out to be true. Since we are dealing with two-element partitions, the class of strictly convex pooling methods contains not only the linear averaging, but also e.g. geometric (logarithmic) pooling [1].

Our main vehicle will be the following simple piece of classical probability, to which we will refer by the name "Partition Lemma". Where X is any proposition, we will say that X is **partitioned by** $\{X_i\}_{i \in \{1,\dots,n\}}$ iff

1. $i \neq j \to X_i \cap X_j = \emptyset$;
2. $\cup\{X_i\}_{i \in \{1,\dots,n\}} = X$;
3. $X_i \neq \emptyset$ for all $1 \leq i \leq n$.

2.0.1 Lemma. Suppose X is partitioned by $\{X_i\}_{i \in \{1,\dots,n\}}$. Then for any event r, $P(r|X)$ is a convex combination of the probabilities $P(r|X_i)$.

Proof. Given the assumptions, we want to show that there exist $\lambda_1, \dots, \lambda_n$ such that $P(r|X) = \sum_{i=1}^{n} \lambda_i P(r|X_i)$, with all the λ_i's nonnegative and summing up to 1.

Note that $P(rX) = P(rX_1 \cup \dots \cup rX_n)$. This, by the law of total probability, is equal to

$$P(X_1)P(rX_1 \cup \dots \cup rX_n|X_1) + \dots + P(X_n)P(rX_1 \cup \dots \cup rX_n|X_n),$$

that is, ultimately, since the X_i's are pairwise disjoint,

$$P(rX) = \sum_{i=1}^{n} P(X_i)P(r|X_i).$$

Dividing both sides by $P(X)$ we get

$$P(r|X) = \sum_{i=1}^{n} P(X_i|X)P(r|X_i).$$

■

We will apply the Partition Lemma to the case of X being equal to $A = a$, for some choice of a, being partitioned by $\{A = a, B = b\}_{b \in \mathrm{OP}_B}$: $P(r|A = a) = \sum_{i=1}^{m} P(B = b_i|A = a)P(r|A = a, B = b_i)$. We thus get immediately the following:

2.0.2 Corollary. Suppose P perfectly trusts A and uses a strictly convex pooling method. Then for any $a \in \mathrm{OP}_A$, $P(r|A = a) \geq \min\{P(r|A = a, B = b) \mid b \in \mathrm{OP}_B\}$.

The Partition Lemma allows us to prove the general impossibility result for perfectly serving two epistemic masters, assuming the probability in question is uses a strictly convex pooling method.

2.0.3 Theorem. Assume two experts A and B are given. If $P(A = B) < 1$ and P uses a strictly convex pooling method, then P does not perfectly trust A or P does not perfectly trust B.

Proof. Recall that by our general assumption the sets OP_A and OP_B are finite. Assuming that P perfectly trusts both A and B, we will now construct infinite increasing sequences of elements of OP_A and OP_B, to achieve a contradiction. These sequences will be indexed by left superscript in what follows.

Since $P(A = B) < 1$, there are i and j such that $a_i \neq b_j$ and $P(A = a_i, B = b_j) > 0$; fix such a pair. Set the initial elements of our sequences: ${}^0b = b_j$ and ${}^0a = a_i$. Since P perfectly trusts A, $P(r|A = a_i) = a_i$. Assume $P(r|A = a_i, B = b_j) < a_i$ (the $>$ direction is analogous). By Lemma 2.0.1 and the fact that P uses a strictly convex pooling method this implies there is a b_k such that $P(r|A = a_i, B = b_k) > a_i$, for $b_j < b_k$ and $a_i < b_k$. Set ${}^1b = b_k$. Since P perfectly trusts B, $P(r|B = b_k) = b_k$, so, as before, since $P(r|A = a_i, B = b_k) < b_k$, there is an a_l such that $P(r|A = a_l, B = b_k) > b_k$, where $a_i < a_l$, $b_j < b_k$, $a_i < b_k$, $a_l > b_k$. Set ${}^1a = a_l$. Since P perfectly trusts A, $P(r|A = a_l) = a_l$, so, as before, since $P(r|A = a_l, B = b_k) < a_l$, there is a b_o such that $P(r|A = a_l, B = b_o) > a_l$, for $b_j < b_k < b_o$, $a_i < a_l$, $a_i < b_k$, $a_l > b_k$. Set ${}^2b = b_o$. Notice that at these point we have ${}^0b < {}^1b < {}^2b$ and and ${}^0a < {}^1a$. The procedure of producing subsequent elements of these sequences does not terminate. This contradicts the assumption that the sets OP_A and OP_B are finite. ■

As already noted, this is more general than Gallow's result in that more pooling methods fall under its scope; it is less general in that the sets of options of the experts under consideration is assumed to be finite.

2.1 CHOOSING WHOM AND HOW TO TRUST

The Partition Lemma has important consequences for which of the two agents you can trust – if you can trust any of them! We will introduce this somewhat surprising issue by means of the following example.

The Two Kindergarten Teachers. Assume that your kindergarten-going kid is somewhat allergic to peanuts. Unfortunately, you have failed to inform the kindergarten teachers of this. Your kid will eat one cookie at the kindergarten today; you have trained him so that he always randomly determines a cookie from a given bowl. You're interested whether he ate a cookie with peanut butter jelly filling. The cookies have two colors, black or white, and two shapes, square or round. Your proposition r is whether there was jelly in the cookie your kid ate. Before meeting the kid himself you ask two teachers, Dan (D) and Eve (E), about the chance of r. Dan always recalls only the colors of the cookies eaten by the kids and based on what he saw in the kitchen today thinks 60% black and 40% white cookies were filled with jelly. Eve only recalls the shapes, and thinks 30% of the square cookies and 70% of the round cookies were filled with jelly. Thus, in the context of inquiring about r, $\mathrm{OP}_D = \{.4, .6\}$ and $\mathrm{OP}_E = \{.3, .7\}$. It turns out that if your credence uses a strictly convex pooling method, you not only cannot trust both Dan and Eve, but if you wish to trust any of the two, it has to be Dan. You simply cannot trust Eve.

The reason for this is because the next corollary follows straight from Corollary 2.0.2:

2.1.1 Corollary. Assume P uses a strictly convex pooling method. Then if A ranges below or above B, P does not perfectly trust A.

Proof. Assume A ranges below B (the other case is analogous). Then $a_1 < b_1$. Since P uses a strictly convex pooling method, $\min\{P(r|A = a_1, B = b) \mid b \in \mathrm{OP}_B\} > a_1$. However, by Corollary 2.0.2, if – as we assume for *reductio* – P perfectly trusts A, then $P(r|A = a_1) = a_1 \geq \min\{P(r|A = a_1, B = b) \mid b \in \mathrm{OP}_B\}$. Therefore $a_1 > a_1$. Contradiction. ■

One understandable reaction to this would be to point to how the Partition Lemma helps us appreciate the somewhat trivial fact that, for any a, the proposition $A = a$ is actually a disjunction of m conjunctions $A = a, B = b_1$, $A = a, B = b_2$, …, and $A = a, B = b_m$. When considering whether P perfectly trusts A, we do not simply investigate the credence in r given that A says the chance of r is a: on every occasion this happens, B also gives their suggestion. We could, then, add a "silent" choice to the set of options of each agent, signified e.g. by the bullet $\bullet$: this would let us consider situations where one of the agents does not put forward any suggestion. It has to be noted, however, that with the sets of agents' options thus extended, but on the reasonable assumption that $P(r|A = a, B = \bullet) = a$ (with the corresponding symmetric condition in play when it is A who stays silent), every reasoning considered so far remains valid: and so Theorem 2.0.3 and Corollary 2.1.1 still hold. We cannot perfectly serve two epistemic masters even if one of them sometimes keeps their mouth shut. If one of the masters ranges below or above the other, we cannot trust them even if the other stays silent on occasion.

Before exploring the unfortunate consequences of this, let us note more constructively that if one of the two agents ranges both below and above the other one, that other one can be perfectly trusted when P uses a strictly convex pooling method to pool the experts' opinions (for our purposes such a method will be treated as a function c of two arguments):

2.1.2 Fact. Fix a proposition r and two agents A and B. If A ranges both below and above B, then for a strictly convex pooling method c there exists a probability function P satisfying the conditions

(i) $P(A = B) < 1$,
(ii) for any $b \in \mathrm{OP}_B$, $P(r \mid B = b) = b$,
(iii) for any $a \in \mathrm{OP}_A$, $b \in \mathrm{OP}_B$, $P(r \mid A = a, B = b) = c(a, b)$.

Proof. Our job will be done once we have set the values of all probabilities of the form $P(r, A = a_i, B = b_j)$ and $P(\neg r, A = a_i, B = b_j)$; that is, probabilities of the probability space's atoms. We need, in other words, $2 \cdot n \cdot m$ nonnegative numbers x_{1ij} and x_{0ij} which sum up to 1, chosen so that the conditions (i)-(iii) are satisfied. Let us write $c_{ij} = c(a_i, b_j)$. Conditions (ii) and (iii) translate to the following equations:

$$\begin{aligned} (1-b_j)\sum_{i=1}^{n} x_{1ij} &= b_j \sum_{i=1}^{n} x_{0ij} && \forall j; \\ (1-c_{ij})x_{1ij} &= c_{ij}x_{0ij} && \forall i, j. \end{aligned}$$

Reordering the equations yields

$$\sum_{i=1}^{n} \left(\frac{c_{ij}}{1-c_{ij}} - \frac{b_j}{1-b_j} \right) x_{0ij} = 0 \qquad \forall j. \tag{1}$$

We need to find non-negative x_{0ij}'s that satisfy (1). Since A ranges both below and above B, $a_1 < b_j < a_n$ for each j. By strict convexity of the pooling method,

$$\begin{aligned} c_{1j} &< b_j \text{ as } c_{1j} \in (a_1, b_j) \\ c_{nj} &> b_j \text{ as } c_{nj} \in (b_j, a_n). \end{aligned}$$

Thus, there are both positive and negative coefficients in the sum in (1). This allows for picking positive x_{0ij}'s satisfying (1) (and, in particular, though we omit the details here, for doing so in such a way that for some $i \neq j$ $x_{0ij} > 0$, ensuring the condition (i)). The condition that the x_{1ij}'s and x_{0ij}'s sum up to 1 is expressed by the equality

$$\sum_i \sum_j \frac{1}{1-c_{ij}} x_{0ij} = 1. \tag{2}$$

The matrix corresponding to the linear equations in (1) and (2) has $m+1$ linearly independent rows (the rows corresponding to (1) have both negative and positive members, while the row corresponding to (2) is strictly positive, ensuring linear independence), and there are $n \cdot m$ variables. Therefore, the system has a continuum of positive solutions. ■

Let us take stock. We cannot serve two epistemic masters if we're using a strictly convex pooling method. Sometimes we can benefit from combining the experts' opinions while also perfectly trusting one of them: but whether such perfect trust is possible depends on the other expert's set of options. For example, if $\text{OP}_A = \{.3, .7\}$, $\text{OP}_B = \{.4, .8\}$, and we wish to use a strictly convex pooling method, due to Corollary 2.1.1 we cannot perfectly trust any of the two experts, since A ranges below B, and B ranges above A. However, that very same A can be trusted if the other agent is a C with $\text{OP}_C = \{.2, .9\}$, thanks to Fact 2.1.2. So whether an expert can be trusted or not depends on the properties of the *other* expert.[1] We find this troublesome and we would lean to the conclusion that this shows that conditions such as 2. and 3. do not properly capture the notion of perfect trust. However, the root of the problem seems to be simple classical probabilistic phenomenon captured in the Partition Lemma, something entirely unavoidable if we do not use an alternative modelling framework. We will now consider two options salient for anyone not willing to lightly abandon classical probability: imperfect trust and a non-convex pooling method.

[1] Tangentially: if the finite sets OP_A and OP_B are identical (and contain at least 3 options), it is always possible to define a P satisfying 1., 4., and one of the two "trust" conditions. We leave this as an exercise for the Reader.

2.2 IMPERFECT TRUST

Given the expert's opinion, a range of credences might be *permissible*, depending on how strictly the agent wishes to follow the expert. We will say that C **serves** A **up to** ε iff, for any a,[2] $C(r|A = a) \in [a - \varepsilon, a + \varepsilon]$.

Let's try to construct an example which satisfies 1., 3., 4., and is "close" to satisfying 2.: that is, the agent serves Bert perfectly, but serves Al up to some nonzero ε. Al's opinion should weigh less, then: let's put the weight α at 0.4. Consider the following:

	r	$\neg r$
$A = .2,\ B = .4$	$0.32 \cdot x$	$0.68 \cdot x$
$A = .2,\ B = .6$	$0.44 \cdot y$	$0.56 \cdot y$
$A = .8,\ B = .4$	$0.56 \cdot w$	$0.44 \cdot w$
$A = .8,\ B = .6$	$0.68 \cdot z$	$0.32 \cdot z$

Notice that if we can find four nonzero numbers x, y, w, z so that everything in our table sums up nicely to 1, we will have a P which satisfies:

- $P(A = B) < 1$, since the two masters just cannot agree (condition 1.);
- $P(r|A = .2, B = .4) = .32$, $P(r|A = .2, B = .6) = .44$, $P(r|A = .8, B = .4) = .56$, $P(r|A = .8, B = .6) = .68$, that is: condition 4. with weights $\alpha = .4$, $\beta = .6$.

Our goal is to find a solution where we also have $P(r|B = .4) = .4$ and $P(r|B = .6) = .6$, that is, condition 3. Then how much $P(r|A = .2)$ deviates from .2 and $P(r|A = .8)$ from .8 will determine the ε up to which P serves A.

It turns out we can achieve $\varepsilon = 2/15$ with "easy" numbers; let us set then as our target $P(r|A = .2) = 1/3$ and $P(r|A = .8) = 2/3$. On these assumptions a modicum of arithmetic indeed finds a solution at $x = 8/15$, $y = 1/15$, $w = 4/15$ and $z = 2/15$. So we can thus serve one master perfectly and at the same time the other up to $2/15$.

We have worked this out in detail to present an example of how one can in general approach the task of generating a P which trusts one of the experts perfectly and the other imperfectly (and assuming the pooling is performed *via* linear averaging). Smaller values of ε are achievable for experts who are more similar to each other, and for experts regarding which P specifies a low (but positive) credence in their disagreement. For example, here is a case where $\mathrm{OP}_B \subset \mathrm{OP}_A$, and the probability of disagreement is relatively low (0.4):

	r	$\neg r$
$A = .2,\ B = .2$	.096	.384
$A = .2,\ B = .4$	0	0
$A = .3,\ B = .2$	.005	.015
$A = .3,\ B = .4$	.007	.013
$A = .4,\ B = .2$	0	0
$A = .4,\ B = .4$	.192	.288

For uniform weights, this satisfies conditions 1., 2., and 4. Thus P perfectly trusts A. Note that $P(r|B = .2) = .202$ and $P(r|B = .4) = .402$; so P serves B up to 0.002. Here, the probability that A and B disagree is relatively low, and we can trust both almost-perfectly.

[2] Assume in what follows that in such contexts the quantification is over the set of options of the relevant agent.

2.3 A NON-CONVEX POOLING METHOD

Our Theorem 2.0.3 shows that we cannot serve two epistemic masters if we use a strictly convex pooling method. However, this possibility is open if a pooling method of a different sort is used. We will now present an example using a non-convex pooling rule introduced in [3], called "upco". The name derives from "updating on credences of others"; however, we will find use for it here even if we are not (at least not explicitly) concerned with updating. For the specific case of a fixed r, A and B, we will say that P **uses upco** iff

$$P(r|A = a, B = b) = \frac{a \cdot b}{a \cdot b + (1 - a) \cdot (1 - b)}.$$

It turns out that if our P uses upco, we can trust both Dan and Eve from our kindergarten example. Consider the following probability space:

	r	$\neg r$
$D = .4,\ E = .3$	$2/9 \cdot x$	$7/9 \cdot x$
$D = .4,\ E = .7$	$14/23 \cdot y$	$9/23 \cdot y$
$D = .6,\ E = .3$	$9/23 \cdot w$	$14/23 \cdot w$
$D = .6,\ E = .7$	$7/9 \cdot z$	$2/9 \cdot z$

It is a mundane matter to check that setting $w = y = 0.23$ and $x = z = 0.27$ generates a probability P which perfectly trusts both D and E, at the same time satisfying $P(r|D = .4, E = .3) = 2/9$, $P(r|D = .4, E = .7) = 14/23$, $P(r|D = .6, E = .3) = 9/23$, and $P(r|D = .6, E = .7) = 7/9$, that is, using upco.

However, if trusting epistemic masters is our general goal, we will not be perfectly happy after adopting upco as our pooling method. Suppose $\mathrm{OP}_A = \{.7, .8\}$ and $\mathrm{OP}_B = \{.6, .9\}$. We know from Fact 2.1.2 that for arbitrary nonzero weights α, $\beta = 1 - \alpha$ there is a probability P satisfying 1., 2., and 4., that is, that we can perfectly trust A if we combine the experts' opinions using linear averaging. Notice, however, that if P uses upco instead, then both $P(r|A = .8, B = .6)$ and $P(r|A = .8, B = .9)$ are above .8; this, in light of Corollary 2.0.2, is incompatible with $P(r|A = .8) = .8$. Therefore, if P uses upco, then P cannot perfectly trust A. There are, then, cases when upco, in contrast with strictly convex pooling methods, allows us to serve two epistemic masters; but there are also cases where upco prohibits us from serving a master whom we could trust perfectly if linear averaging was used.

3 GOING HIGHER-ORDER

Given the role the Partition Lemma seems to play in the apparent unwieldiness of the classical trust conditions, it might be tempting to consider some approach to expert deference which would go beyond classical probability. The hope could be that if related phenomena also arose there, perhaps they would be somewhat further 'below the surface' and would not block us from trusting experts we should be able to trust. One approach *prima facie* suitable for such a job is that of Higher Order Probability Spaces [HOPs, 4]. Introduced by Gaifman in 1986 and revised in 1988, they, despite the clear formal success of the main construction, have not achieved popularity in formal epistemology; however, they do come up from time to time in the literature.[3] Intuitively, the theory deals with probability spaces where *that the probability P of the event X belongs to the segment* $[x, y] \subseteq [0, 1]$ is also an event in the domain of P.

[3] Sometimes this is not immediately apparent. We suspect, for example, that there is a close kinship between [2]'s probability frames and [4]'s kernels of HOPs. Arriving at a precise formulation of this relationship is a task for another day.

Let $\mathcal{I}$ be the set of closed intervals with endpoints in $[0,1]$. A quadruple $(W, \mathcal{F}, P, pr)$ is called a **higher order probability space** if $(W, \mathcal{F}, P)$ is a probability space, and pr is a mapping $pr : \mathcal{F} \times \mathcal{I} \to \mathcal{F}$ so that certain conditions (I)-(V) are satisfied which essentially stipulate that the function pr behaves in harmony with measure-theoretic probability,[4] while it is usually intended that the structure satisfy an additional axiom (VI), which gives pr the intended interpretation: $pr(A, \Delta)$ is the event "that the expert probability of A lies in the interval Δ" [4, p. 192]. The axiom (VI) is supposed to capture the agent's perfect trust in the expert:

(VI) If $E \in \mathcal{F}$ is a finite intersection of the form

$$E = pr(A_1, \Delta_1) \cap pr(A_2, \Delta_2) \cap \ldots \cap pr(A_n, \Delta_n) \tag{6}$$

then for all $A \in \mathcal{F}$ for which $P(pr(A, \Delta) \cap E) \neq 0$ we have

$$P(A \mid pr(A, \Delta) \cap E) \in \Delta. \tag{7}$$

That is, "the mere knowledge of the [expert] assignment will make the agent adopt it as subjective probability" (*ibid.*). Closed intervals may, of course, be single points: for $0 \leq x \leq 1$, $pr(A, x)$ denotes $pr(A, [x, x])$. Axiom (VI) applied to a single point interval and $E = W$ is essentially Miller's Principle:[5]

$$P\left(A \mid pr(A, x)\right) = x \tag{8}$$

provided $P(pr(A, x)) \neq 0$. This bears an obvious similarity to the previously discussed trust conditions such as 2. above.

Suppose $(W, \mathcal{F}, P, pr)$ is a finite HOP. There is a unique mapping p which associates with every $x \in W$ a probability p_x over $\mathcal{F}$ such that

$$pr(A, \Delta) = \{x : \ p_x(A) \in \Delta\} \tag{9}$$

holds. Vice versa, if $\{x : \ p_x(A) \in \Delta\}$ is in $\mathcal{F}$ for all $A \in \mathcal{F}$ and Δ, then $pr(A, \Delta)$ defined by equation (9) gives back a HOP satisfying axioms (I)-(V). The mapping p is called the *kernel* of the HOP and is considered to be a $|W| \times |W|$ matrix ([4, Theorem 1]). A HOP satisfies axiom (VI)[6] if and only if its kernel is idempotent and P is a left-eigenvector of the kernel [4, Theorem 2]. Roughly, such kernels are square matrices whose rows are probabilistic vectors, featuring along the up-left-to-down-right diagonal square blocks with identical rows—see the examples on the following pages.

---

[4]Here are the conditions, for completeness:

(I) $pr(A, \emptyset) = \emptyset$ for all $A \in \mathcal{F}$;

(II) $pr(A, [0,1]) = X$ for all $A \in \mathcal{F}$;

(III) if Δ_1, Δ_2 and $\Delta_1 \cup \Delta_2$ are all intervals from $\mathcal{I}$, then for all $A \in \mathcal{F}$

$$pr(A, \Delta_1 \cup \Delta_2) = pr(A, \Delta_1) \cup pr(A, \Delta_2); \tag{3}$$

(IV) If Δ_i $(i = 1, 2, \ldots n \ldots)$ is a countable set of intervals from $\mathcal{I}$, then

$$\cap_i pr(A, \Delta_i) = pr(A, \cap_i \Delta_i); \tag{4}$$

(V) If A_i $(i = 1, 2, \ldots n \ldots)$ is a countable set of pairwise disjoint elements from $\mathcal{F}$, and $[\alpha_i, \beta_i]$ is a countable set of intervals from $\mathcal{I}$, and $[\alpha, \beta]$ is the interval with $\alpha = \sup_i \alpha_i$ and $\beta = \sup_i \beta_i$, then

$$\cup_i pr(A_i, [\alpha_i, \beta_i]) \subseteq pr(\cup A_i, [\alpha, \beta]). \tag{5}$$

[5]The original paper by Miller, where he argues that this principle is actually untenable, is [7]; be sure to read the review by Jeffrey, [6].

[6]Gaifman's paper is ambiguous whether axiom (VI) is in fact constitutive of HOPs; we take it that it isn't.

Since our topic is deference to two experts, we should note that nothing in principle prohibits adding a second expert probability function to a HOP. We propose that a quintuple $(W, \mathcal{F}, P, pr_1, pr_2)$ be called a **bare double higher order probability space** (a bare double HOP) exactly when both $(W, \mathcal{F}, P, pr_1)$ and $(W, \mathcal{F}, P, pr_2)$ are HOPs satisfying the axiom (VI). (It's "bare" since no condition linking the two expert functions is stipulated at this point.) We could then investigate whether conditional probabilities of the form $P(A|pr_1(A, \Delta_1) \cap pr_2(A, \Delta_2))$, understood as conditional credences given two opinions by the different experts, display some principled or otherwise interesting behaviour. We will now present three examples intended to suggest that the concept of a bare double HOP is perhaps not sufficiently fruitful as is, and we should search for some meaningful condition linking the two expert functions to be added to its definition.

Example 1: lack of unanimity preservation. Consider the propability space $W = \{w_1, \dots, w_4\}$, $\mathcal{F} = \mathcal{P}(W)$ and

$$P = \begin{bmatrix} 3/8 & 1/8 & 1/8 & 3/8 \end{bmatrix}.$$

Let two kernels p^1 and p^2 be defined as follows:

$$p^1 = \begin{bmatrix} .6 & .2 & .2 & 0 \\ .6 & .2 & .2 & 0 \\ .6 & .2 & .2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad p^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & .2 & .2 & .6 \\ 0 & .2 & .2 & .6 \\ 0 & .2 & .2 & .6 \end{bmatrix}.$$

These two kernels give rise to functions pr_1 and pr_2 via (9). Easy calculations show that P is a mixture of both kernels (a left-eigenvector) with coefficients respectively

$$[5/24, 5/24, 5/24, 3/8], \quad \text{and} \quad [3/8, 5/24, 5/24, 5/24].$$

As the kernels are idempotent, it follows that $(W, \mathcal{F}, P, pr_1)$, and $(W, \mathcal{F}, P, pr_2)$ are both HOPs satisfying axiom (VI); that is, $(W, \mathcal{F}, P, pr_1, pr_2)$ is a bare double HOP. Let now $A = \{w_1, w_2, w_4\}$. Then

$$pr_1(A, .8) = \{w_1, w_2, w_3\}, \text{ and } pr_2(A, .8) = \{w_2, w_3, w_4\},$$

and

$$P(A \mid pr_1(A, .8) \cap pr_2(A, .8)) = P(\{w_1, w_2, w_4\} \mid \{w_2, w_3\}) = \frac{1/8}{1/8 + 1/8} = .5$$

Thus, despite the unanimous opinions of the two experts giving their credence in A as .8, the agent sets their credence to .5; in other words, bare double HOPs fail to satisfy the condition of preserving unanimity [1].

Example 2: when one expert is ignored. Define two kernels in this way:

$$p^1 = \begin{bmatrix} .5 & .5 & & & & \\ .5 & .5 & & & & \\ & & .5 & .5 & & \\ & & .5 & .5 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix}, \qquad \text{and } p^2 = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1/6 & 1/6 & 4/6 & \\ & & 1/6 & 1/6 & 4/6 & \\ & & 1/6 & 1/6 & 4/6 & \\ & & & & & 1 \end{bmatrix}.$$

The corresponding probability vector is

$$P = \begin{bmatrix} 1/6 & 1/6 & 3/36 & 3/36 & 12/36 & 1/6 \end{bmatrix},$$

being a mixture of the kernels with respective coefficients

$$[1/6, 1/6, 3/36, 3/36, 12/36, 1/6], \qquad \text{and} \qquad [1/6, 1/6, 1/6, 1/6, 1/6, 1/6].$$

Let $A = \{w_1, w_3, w_5\}$. Then

$$\begin{aligned} pr_1(A, 1/2) &= \{w_1, w_2, w_3, w_4\} \\ pr_2(A, 5/6) &= \{w_3, w_4, w_5\}, \end{aligned}$$

and therefore

$$\begin{aligned} P(A \mid pr_1(A, 1/2) \cap pr_2(A, 5/6)) &= P(A \mid \{w_3, w_4\}) \\ &= \frac{P(\{w_3\})}{P(\{w_3, w_4\})} = \frac{3/36}{3/36 + 3/36} = 1/2. \end{aligned}$$

In this case, therefore, the opinion of the second expert is ignored: the one of the first expert trumps it completely.

Example 3: moving to the extreme. Define two kernels as

$$p^1 = \begin{bmatrix} .5 & .5 & & & & \\ .5 & .5 & & & & \\ & & .5 & .5 & & \\ & & .5 & .5 & & \\ & & & & .5 & .5 \\ & & & & .5 & .5 \end{bmatrix}, \qquad \text{and } p^2 = \begin{bmatrix} 1/3 & 1/3 & 1/3 & & & \\ 1/3 & 1/3 & 1/3 & & & \\ 1/3 & 1/3 & 1/3 & & & \\ & & & 1/3 & 1/3 & 1/3 \\ & & & 1/3 & 1/3 & 1/3 \\ & & & 1/3 & 1/3 & 1/3 \end{bmatrix},$$

corresponding to the uniform probability

$$P = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix}.$$

Let $A = \{w_3\}$. Then

$$\begin{aligned} pr_1(A, 1/2) &= \{w_3, w_4\} \\ pr_2(A, 1/3) &= \{w_1, w_2, w_3\}, \end{aligned}$$

and therefore

$$P(A \mid pr_1(A, 1/2) \cap pr_2(A, 1/3)) = P(\{w_3\} \mid \{w_3\}) = 1.$$

In this case, the conditional credence in A given two rather moderate estimates by the experts is, surprisingly, 1.

These examples suggest that the concept of a bare double HOP is, well, *too* bare to be fruitfully used when modelling simultaneous deference to two experts. This motivates the following research question:

Question. What condition linking the two expert probabilities should be added to the definition of a bare double HOP to make it philosophically fruitful?

This question is difficult if only for the fact that constructing even a single example of a HOP where P would use the same pooling method for all propositions is a nontrivial matter.

4 CONCLUSIONS

In this paper we have, in the context of two experts with finite sets of options, extended Gallow's impossibility result to cover all strictly convex pooling methods (Theorem 2.0.3). We have then shown that the nonconvex upco method sometimes allows the possibility of serving two epistemic masters. However, this cannot be seen as the go-to solution to the problem, since there exist pairs of agents such that one is perfectly trustable if linear averaging is used, but is not if upco is employed. Along the way we have also noted a sufficient condition for any of two given experts not to be perfectly trustable if a strictly convex pooling method is used (Corollary 2.1.1), and a sufficient condition for the existence of a probability function trusting one of the two experts if such a method is employed (Fact 2.1.2). We have also briefly considered limited trust: perhaps one could be fine with serving two masters *almost* perfectly.

We find it troubling that when Gallow's conditions are employed and two experts are concerned, whether one of them can be trusted depends on the properties of the other one. This phenomenon is rooted in the simple classical probabilistic fact captured in our Partition Lemma (Lemma 2.0.1). It is related to convexity, but its implications are not limited to situations where convex pooling methods are used: we have noted a case where it leads to the nonconvex upco method prohibiting us from trusting an agent which is easily trusted if a strictly convex method is employed. This suggests that some nonclassical approach might be fruitful, where the proposition "expert A says φ" would not be assumed from the start to literally be a union of disjoint propositions of the form "expert A says φ and expert B says ψ" for all the possible values of ψ. In section 3 we have considered extending the framework of Higher-Order Probability Spaces to cover the cases of two experts, but the examples we present suggest the need of some additional axiom connecting the functions representing the suggestions of the two experts.

Apart from this particular issue, other questions seem worth considering in the future, even if we stay in the realm of classical probability. If our goal is to find ways of serving two epistemic masters, are there different nonconvex pooling methods aside from upco that at least sometimes do the trick? If yes, do they differ in terms of pairs of perfectly trustable experts, and if so, is there some interesting meaning to that difference? If one pooling method allows us to trust more experts than another, is the former necessarily better? And is it true that in principle any expert should be trustable alongside any other expert? We leave tackling these questions for further research.

FUNDING

The research was supported by the grant 2019/34/E/HS1/00044, 'Epistemic inaccuracy and foundational issues in formal epistemology', of the National Science Centre (Poland).

REFERENCES

[1] Dietrich, F. and List, C. (2016). Probabilistic Opinion Pooling. In Hajek, A. and Hitchcock, C., editors, *Oxford Handbook of Philosophy and Probability*. Oxford University Press.

[2] Dorst, K. (2020). Evidence: A Guide for the Uncertain. *Philosophy and Phenomenological Research*, 100(3):586–632.

[3] Easwaran, K., Fenton-Glynn, L., and Hitchcock, C. (2016). Updating on the Credences of Others: Disagreement, Agreement, and Synergy. *Philosophers' Imprint*, 16(11).

[4] Gaifman, H. (1988). A theory of Higher Order Probabilities. In Skyrms, B. and Harper, W. L., editors, *Causation, Chance and Credence: Proceedings of the Irvine Conference on Probability and Causation Volume 1*, pages 191–219. Springer Netherlands, Dordrecht. Section: A Theory of Higher Order Probabilities.

[5] Gallow, J. D. (2018). No one can serve two epistemic masters. *Philosophical Studies*, 175(10):2389–2398.

[6] Jeffrey, R. (1970). Review. *The Journal of Symbolic Logic*, 35(1):124–127.

[7] Miller, D. (1966). A Paradox of Information. *The British Journal for the Philosophy of Science*, 17(1):59–61.