

Proper Scoring Rules with Additional Properties

Preliminary Report

Peng Shi^{*}
Duke University,
Durham, NC 27708
ps46@duke.edu

ABSTRACT

Proper scoring rules are reward systems designed to elicit an agent’s best estimate of some probabilistic distribution. They are used extensively in obtaining forecasts such as for the weather, and they are fundamental to the theory of market-scoring rules and prediction markets. In this paper, we present a geometric interpretation to a previously known characterization of proper scoring rules. Moreover, we advance the theory of proper scoring rules with additional properties. One variant we examine are *society-aligned* proper scoring rules: suppose society has a certain utility for each outcome, a society-aligned scoring rule will not incentivize an agent to change the true distribution in a direction that harms society’s expected welfare. We show that all society-aligned proper scoring rules are equivalent to an a one-dimensional scoring rule that asks agents to estimate society’s expected utility.

1. INTRODUCTION

In a world where information is becoming more valuable, one may try to elicit information through some kind of payment scheme. For example, suppose one seeks an estimate of the probability of some event, say whether it rains tomorrow, and one asks an expert. One may promise a certain payment if it rains, and some other payment it does not, and the payments depend on the expert’s report. It turns out by choosing the payment functions carefully, one can incentivize the agent to report her best estimate. Such a payment scheme is called a proper scoring rule.

Beside being used in weather forecasting [5], proper scoring rules are used to assess the quality of predictions in a multitude of other settings [6, 8]. Furthermore, they are the basis of what are called market-scoring rules, which underlies the theory of prediction markets [9]. As predictions markets are becoming more popular and more trusted as sources of in-

^{*}This project is funded under Duke University’s C-SURF undergraduate Research Fellowship.

formation, the theory of proper scoring rules are becoming more important.

In this paper, we advance the theory of proper scoring rules satisfying some additional properties. One motivated property we want to add to proper scoring rules is to make it aligned with society’s interests (Section 3). For example, suppose we use a proper scoring rule to elicit an agent’s beliefs about the probability of a terrorist attack, a commonly used proper scoring rule, such as the quadratic or logarithmic rules, might incentivize her to change the underlying probability by committing a terrorist attack. For such purposes, we want rules that never incentivize agents to act in a way that damages society’s expected welfare. In this paper, we characterize all such society-aligned rules, showing that any such rule is equivalent to a one-dimensional proper scoring rule that asks the agent to estimate society’s expected utility.

Our results are based on a general characterization theorem, parts of which can be attributed to Savage [7], McCarthy [4], Hendrickson and Buehler [2], and Gneiting and Raftery [1]. We prove the characterization theorem and present a geometric interpretation, which greatly aids the derivation or our other results.

This is only a preliminary report, and many possible directions of research are yet to be explored. The immediate next questions and other ideas are listed in Section 6.

2. REVIEW OF GENERAL THEORY

We first present the general theory of proper scoring rules and a characterization theorem, which was discovered at first by Savage [7] but the version shown here is due to Gneiting and Raftery [1]. While proper scoring rules can be defined to elicit any statistical property [3], we will focus on proper scoring rules which elicit the whole distribution, represented by a probability for each possible outcome. None of the content in this section are new results, but some are presented in a new way.

2.1 Review of Convexity Theory

We review general theories of convex real-valued functions in R^n , as well as notions of subgradient vectors and subtangent hyperplanes. In the following, let U denotes an open domain $U \in \mathbb{R}^n$.

Definition 1. A function $G : U \rightarrow \mathbb{R}$ is *convex* if $\forall x, y \in U$, and $\forall \lambda \in (0, 1)$,

$$G(\lambda x + (1 - \lambda)y) \leq \lambda G(x) + (1 - \lambda)G(y)$$

G is called *strictly convex* if the inequality is strict.

One nice property of convex functions is that the supremum of a family of convex functions is still convex.

THEOREM 1. *Given a (possibly infinite) family J of convex functions defined on domain U , define the function $G : U \rightarrow \mathbb{R}$ as $G(x) = \sup_{G' \in J} G'(x)$, then G is convex.*

PROOF.

$$\begin{aligned} & G(\lambda x + (1 - \lambda)y) \\ &= \sup_{G' \in J} G'(\lambda x + (1 - \lambda)y) \\ &\leq \sup_{G' \in J} (\lambda G'(x) + (1 - \lambda)G'(y)) \\ &\leq \sup_{G' \in J} \lambda G'(x) + \sup_{G' \in J} (1 - \lambda)G'(y) \\ &= \lambda G(x) + (1 - \lambda)G(y) \end{aligned}$$

□

By Definition 1, the domain U of a convex function is necessarily a convex set (meaning that if $x, y \in U$, the line segment from x to y is also in U .) On open domains, convex functions are continuous everywhere.

THEOREM 2. *A convex function G defined on an open domain U is everywhere continuous.*

Definition 2. Given convex function $G : U \rightarrow \mathbb{R}$, a *subgradient* is a vector function $G^* : U \rightarrow \mathbb{R}^n$ such that $\forall x, y \in U$,

$$G(y) \geq G(x) + G^*(x) \cdot (y - x)$$

Where \cdot denotes the regular dot product. G^* is called a *strict subgradient* if the above inequality is strict $\forall y \neq x$.

It is well known that subgradient vectors exist for every convex function defined on open domains. If G is differentiable at $x \in U$, then the subgradient at x has to be the gradient: $G^*(x) = (\Delta G)(x)$. Otherwise there might be many choices of $G^*(x)$.

A less conventional but more intuitive term is the *subtangent hyperplane*, which are closely tied to subgradients. Intuitively, these are hyperplanes touching but lying entirely below the convex function.

Definition 3. Given convex function $G : U \rightarrow \mathbb{R}$, a *subtangent hyperplane* at $x \in U$ is a linear function $H_x : U \rightarrow \mathbb{R}$ such that

$$\begin{aligned} H_x(x) &= G(x) \\ H_x(y) &\leq G(y) \quad \forall y \in U \end{aligned}$$

H_x is called a *strictly subtangent hyperplane* if the above inequality is strict $\forall y \neq x$.

Given subgradient vectors $G^*(x)$, it is straightforward to check that we can obtain a subtangent hyperplane in the following way.

LEMMA 1. *Given convex function G and (strict) subgradient vector $G^*(x)$ at point $x \in U$. Define $H_x(y) = G(x) + G^*(x) \cdot (y - x)$, then $H_x(y)$ is a (strictly) subtangent hyperplane of G at x .*

2.2 Definition of Proper Scoring Rules

Let Ω be a set of n possible disjoint outcomes; label $\Omega = \{1, 2, \dots, n\}$. Let $\mathbb{P} = \{P \in \mathbb{R}^n : 0 < P_i < 1, \sum_{i=1}^n P_i = 1\}$ be the set of possible probabilistic distributions over the outcomes, in which every probability must be non-zero. (This condition helps us handle peculiar cases with discontinuities at the boundary, and makes the ensuing math much nicer. After building our theory on distributions with full support, we can handle the edge cases by taking limits.) Define the standard basis $B(i) \in \mathbb{P}$, $B_i(i) = 1$, $B_j(i) = 0 \forall j \neq i$.

Definition 4. A *scoring rule* is a function $S : \mathbb{P} \times \Omega \rightarrow \mathbb{R}$. Given a scoring rule, we can define the expected payment function $\tilde{S} : \mathbb{P} \times \mathbb{P} \rightarrow \mathbb{R}$ as

$$\tilde{S}(P, Q) = \sum_{i=1}^n S(P, i)Q_i$$

A scoring rule can be thought of as a reward system in which we ask an agent to predict the distribution P of outcomes in Ω , and we reward the agent $S(P, i)$ if outcome i occurs. Since the outcomes are disjoint the agent's expected utility assuming the true distribution is $Q \in \mathbb{P}$ is $\sum_{i=1}^n S(P, i)Q_i = \tilde{S}(P, Q)$.

We want to use the scoring rule to incentivize the agent to report her best estimate of the true distribution Q .

Definition 5. A scoring rule $S : \mathbb{P} \times \Omega \rightarrow \mathbb{R}$ is *proper* if $\forall Q \in \mathbb{P}$,

$$\tilde{S}(Q, Q) \geq \tilde{S}(P, Q) \quad \forall P \in \mathbb{P}$$

It is *strictly proper* if equality occurs iff $Q = P$.

2.3 General Characterization of Proper Scoring Rules

Essentially, every proper scoring rule is equivalent to a family of subtangent hyperplanes to some convex function. This general, geometric characterization is due to Gneiting and Raftery [1].

THEOREM 3 (GNEITING & RAFETERY). *Given a convex function $G : \mathbb{P} \rightarrow \mathbb{R}$ and a family of subtangent hyperplanes $H_P : \mathbb{P} \rightarrow \mathbb{R}$, $P \in \mathbb{P}$, then setting $S(P, i) = H_P(b(i))$ defines a proper scoring rule. Equivalently expressed in terms of subgradient vectors, if $H_P(Q) = G(P) + G^*(P) \cdot (Q - P)$,*

$$S(P, i) = G(P) - G^*(P) \cdot P_i + G_i^*(Q)$$

Conversely, every proper scoring rule can be written in the above form. The above result is true if proper is replaced by strictly proper and subtangent by strictly subtangent.

PROOF. Suppose that H_P is a subtangent hyperplane (recall Definition 3), $\forall Q \in \mathbb{P}$,

$$\tilde{S}(P, Q) = \sum_{i=1}^n H_P(B(i))Q_i = H_P(Q)$$

(Since H_P is a linear function constrained a $n - 1$ dimensional space, it is determined by a weighted average of the n intercepts we specify.) Therefore, because H_P lies below $G(Q)$,

$$\tilde{S}(P, Q) = H_P(Q) \leq G(Q) = H_Q(P) = \tilde{S}(Q, Q)$$

Hence, S is a proper scoring rule. If H_P is strictly sub-tangent, then the above inequality is strict $\forall Q \neq P$ and S is strictly proper.

Conversely, suppose S is a proper scoring rule, define $G(Q) = \tilde{S}(Q, Q)$. Note that $G(Q) = \sup_{P \in \mathbb{P}} \tilde{S}(P, Q)$. For fixed P , $\tilde{S}(P, Q) = \sum_i S(P, i)Q_i$ is a linear and therefore convex function of $Q \in \mathbb{P}$. By Theorem 1, G is convex. Now define $H_P = \tilde{S}(P, Q)$. H_P is a linear function in Q , and $H_P(P) = G(P)$. Moreover, $\forall Q \in \mathbb{P}$,

$$H_P(Q) = \tilde{S}(P, Q) \leq \tilde{S}(Q, Q) = G(Q)$$

Hence, H_P is a subtangent hyperplane to G . (It is strictly sub-tangent if S is strict.) \square

Geometric Interpretation: Given the agent's report P and the true distribution Q , the agent's expected reward is $\sum_i S(P, i)Q_i$, which is equal to the sub-tangent hyperplane $H_P(Q)$ with intercepts at $B(i)$ equal to $S(P, i)$. Since $H_P(P) = G(P)$ the agent can attain expected reward of $G(P)$ if the true distribution $Q = P$. But if the true distribution is $Q \neq P$, the agent gets reward $H_P(Q) \leq G(Q)$ by reporting P , but she could have gotten $G(Q)$ if she had reported the true distribution Q .

For each proper scoring rule S , there exists unique convex function $G(P)$ which represents the maximum expected reward the agent can obtain by reporting P . Conversely, any convex function corresponds to some proper scoring rule¹

2.4 Connections to the Financial Asset Interpretation

For one-dimensional proper scoring rules (two outcomes 0 and 1), one well-known and popular interpretation deals with buying shares of an asset which pays \$1 if the event happens, and 0 if the event does not happen. We are allowed to buy shares for various prices, and the starting price is 0. From price level p_1 to p_2 , we are allowed to buy $\int_{p_1}^{p_2} N(q) dq$ shares for some predefined density function N . By participating in this market, we get the lump-sum reward A . Hence, our total reward if the event does not happen, $S(p, 0) = A - \int_0^p qN(q) dq$ and our total reward if the event does happen is $S(p, 1) = A + \int_0^p (1-q)N(q) dq$. See Figure 1 for an illustration.

¹If G is not differentiable, then many families of sub-tangent hyperplanes can be specified, each of which corresponds to a proper scoring rule.

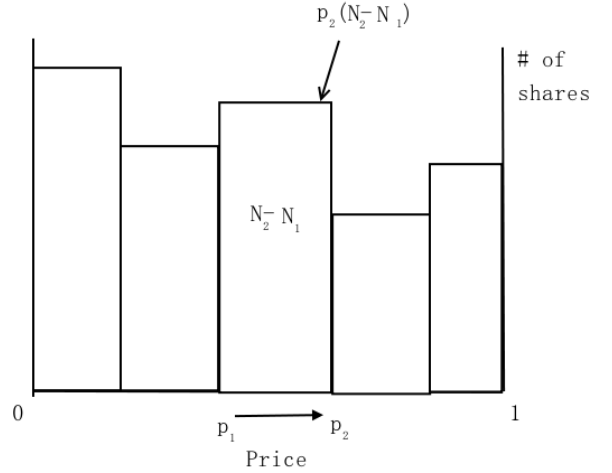


Figure 1: The financial asset interpretation of a one-dimensional proper scoring rule. Suppose that from price p_1 to p_2 , we are allowed to buy $N_2 - N_1$ securities. As we move from probability p_1 to p_2 , the number of shares purchased changes from N_1 to N_2 . $S(p, 0)$ (A - (how much we have to pay)) goes down by $p_2(N_2 - N_1)$, and $S(p, 1)$ (reward if 1 occurs - (how much we have to pay)) goes up by $N_2 - N_1 - p_2(N_2 - N_1)$.

It is well-known that every one-dimensional proper scoring rule can be expressed in this form. Intuitively, this also explains why proper scoring rules are important for prediction markets, as the interpretation resembles such a market.

By Theorem 3, we can also interpret a one-dimensional proper scoring rule as the intercepts of sub-tangents to some convex function G , with the convex function being the expected reward of reporting correctly at each point. See Figure 2 for an illustration.

These two interpretations turn out to be intricately intertwined. If we interpret $N(p)$ as the slope of the convex function G , then buying shares at price p in the financial asset model corresponds to pivoting by point p in the convex function interpretation. By comparing Figures 1 and 2, one can easily check that the induced payment functions $S(p, 0)$ and $S(p, 1)$ are identical in these interpretations.

While the financial asset interpretation can prove very useful, it is unclear how to generalize it to more than two outcomes. The convex function interpretation, on the other hand, generalizes directly to multiple dimensions.

3. SOCIETY-ALIGNED PROPER SCORING RULES

3.1 Definition

Sometimes we want proper scoring rules to be aligned with society's interests. For example, suppose we use a proper scoring rule to elicit an agent's beliefs about the probability of a terrorist attack, which can be viewed as placing a bet. Suppose the agent bets heavily for a terrorist attack, she would lose money if such an attack does not happen.

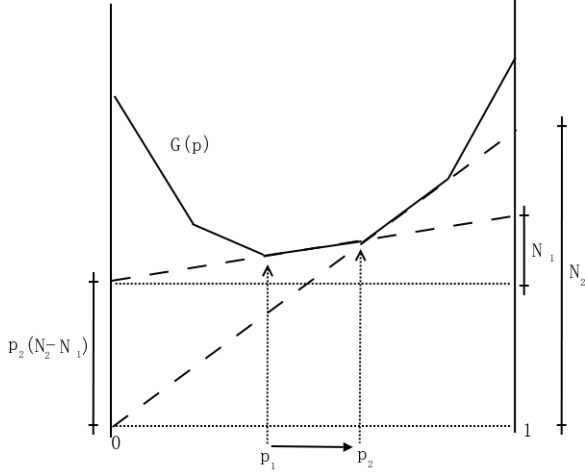


Figure 2: The convex function interpretation of a one-dimensional proper scoring rule. The slope before p_2 is N_1 and the slope after p_2 is N_2 . As we move from probability p_1 to p_2 , we pivot around point p_2 . $S(p, 0)$ (the intercept of the subtangent at 0) goes down by $p_2(N_2 - N_1)$, and $S(p, 1)$ (the intercept at 1) goes up by $N_2 - N_1 - p_2(N_2 - N_1)$.

Hence, she might be incentivized by our bet to change the underlying probability by committing a terrorist attack. We want the proper scoring rule to prevent this type of behavior, by ensuring that no agent can make money by changing the true distribution in a direction that hurts society.

More formally, let $T \in \mathbb{R}^n$ be a vector of society's utility on each of the n outcomes. Then given the true distribution $Q \in \mathbb{P}$, society's expected utility is $Q \cdot T$. We define a scoring rule that is aligned with society's interests.

Definition 6. A proper scoring rule S is *aligned* with vector T if $\forall Q_1, Q_2 \in \mathbb{P}$, if $(Q_2 - Q_1) \cdot T > 0$, then

$$\tilde{S}(Q_2, Q_2) \geq \tilde{S}(P, Q_1) \quad \forall P \in \mathbb{P}$$

We call S *strictly aligned* if the inequality is strict.

We call T *uniform* if T is orthogonal to the hyperplane of possible distributions \mathbb{P} . *i.e.* $T_1 = T_2 = \dots = T_n$. Note that the above definition does not say anything when T is uniform—society is indifferent toward all outcomes.

3.2 Characterization of Society-Aligned Proper Scoring Rules

It turns out that Theorem 3 allows us to formulate a simple characterization of all society-aligned proper scoring rules.

Definition 7. A convex function $G : U \rightarrow \mathbb{R}$ is *monotonically increasing with respect to direction* $T \in \mathbb{R}^n$ if $\forall x_1, x_2 \in U$, $(x_2 - x_1) \cdot T > 0$ implies

$$G(x_2) \geq G(x_1)$$

It is *strictly* monotonically increasing if the above inequality is strict.

The above definition, along with the truthful reporting property of proper scoring rules and Theorem 3, implies the following.

COROLLARY 1. A proper scoring rule S is (strictly) aligned with vector T iff its maximum payment function $G(Q) = \tilde{S}(Q, Q)$ is (strictly) monotonically increasing with respect to T .

When the outcome space is one-dimensional (two outcomes), this classifies all society-aligned proper scoring rules.

COROLLARY 2. A one-dimensional (strictly) proper scoring rule $S : (0, 1) \times \Omega \rightarrow \mathbb{R}$ is aligned with direction $+1$ iff its maximum payment function $G(q) = (1 - q)S(q, 0) + qS(q, 1)$ is (strictly) convex and monotonically increasing. Conversely, any (strictly) convex and monotonically increasing function $G(q)$, along with a family of subtangent lines $H_q : (0, 1) \rightarrow \mathbb{R}$ defines a (strictly) proper scoring rule aligned with $+1$, with $S(q, 0) = H_q(0)$ and $S(q, 1) = H_q(1)$.

The above statements are true when aligned is replaced by strictly aligned, and when monotonicity is replaced by strict monotonicity.

When the outcome space is multi-dimensional, we need the following lemma.

LEMMA 2. Let $U \subset \mathbb{R}^n$ be a convex, open domain. If convex function $G : U \rightarrow \mathbb{R}$ is (strictly) monotonically increasing with respect to direction T , (T is not normal to U), then $G(x) = g(x \cdot T)$ for some single variable (strictly) monotonically increasing convex function g .

PROOF. For each $x \in U$, define the set $U_x = \{y \mid y \in U, (x - y) \cdot T > 0\}$. U_x is non-empty because if $N \subset U$ is an open neighborhood of x , then because T is not normal to U , $\exists y \in N$ such that $(x - y) \cdot T \neq 0$, in which case either y or $-y$ is in U_x . Moreover, U_x is open because it is the intersection of open sets U and $\{y \mid (x - y) \cdot T > 0\}$.

By monotonicity of G , $G(x) \geq \sup_{y \in U_x} G(y)$. But x lies in the closure \bar{U}_x , so by continuity of G , $G(x) \leq \sup_{y \in U_x} G(y)$. This means that $\forall x \in U$, $G(x) = \sup_{y \in U_x} G(y)$.

Note now that $x_1 \cdot T = x_2 \cdot T$, $U_{x_1} \equiv U_{x_2}$, which implies that $G(x_1) = G(x_2)$. Hence, $G(x) = g(x \cdot T)$ for some single variable function g . Moreover, g must be convex because G is convex, and g must be monotonically increasing because G is monotonically increasing w.r.t. T . The above proof holds if monotonicity is replaced by strict monotonicity. \square

This implies the following perhaps surprising characterization of society-aligned proper scoring rules.

THEOREM 4. Given non-uniform societal welfare vector $T \in \mathbb{R}^n$, every proper scoring rule S aligned with T is equivalent to a one-dimensional proper scoring rule aligned with

+1 that asks the agent to estimate

$$f(Q) = \frac{E[T] - T_{min}}{T_{max} - T_{min}}$$

Where $E[T] = \sum_i T_i Q_i$ is society's expected welfare, $T_{min} = \min T_i$ and $T_{max} = \max T_i$ are society's worst and best welfare over the outcomes. (Two scoring rules are equivalent if the expected payments are the same regardless of the agent's belief $P \in \mathbb{P}$ or the true distribution $Q \in \mathbb{P}$.) The above statement holds when proper is replaced by strictly proper.

PROOF. By Corollary 1, S being aligned with T is equivalent to $G(Q) = \tilde{S}(Q, Q)$ being monotonically increasing with respect to direction T . Since T is non-uniform, T is not orthogonal to the domain \mathbb{P} of G , so Lemma 2 implies that $G(Q) = g'(Q \cdot T)$ for some monotonically increasing single variable convex function g' . Now, $Q \cdot T$ is exactly society's expected welfare $E[T]$, and we can perform the linear transformation

$$g(x) = g'(x(T_{max} - T_{min}) + T_{min})$$

to yield monotonically increasing convex function $g : (0, 1) \rightarrow \mathbb{R}$. $\forall P \in \mathbb{P}$, define $f(P) = \frac{P \cdot T - T_{min}}{T_{max} - T_{min}}$, then $G(P) = g(f(P))$. g can be viewed the maximum payment function corresponding to some one-dimensional proper scoring S' rule aligned with +1 that asks the agent to estimate $f(Q)$. We show that $\forall P, Q$, the expected payment under S equals that under S' .

Suppose that $G^*(P)$ is a subtangent for G , and let $g^*(f(P))$ be the corresponding subtangent for g . We have $G(P) = g(f(P))$, and

$$G_i^*(P) = \frac{T_i}{T_{max} - T_{min}} g^*(f(P))$$

Using the characterization theorem (Theorem 3), we have that the expected payment $\tilde{S}(P, Q)$ is

$$\begin{aligned} G(P) + \sum_i G_i^*(P)(Q_i - P_i) \\ = g(f(P)) + \sum_i \frac{T_i}{T_{max} - T_{min}} g^*(f(P))(Q_i - P_i) \\ = g(f(P)) + (f(Q) - f(P))g^*(f(P)) \end{aligned}$$

The last expression is exactly the payment of g if the user reports $f(P)$ and the observed scaled social welfare is $f(Q)$. Hence, S and S' are equivalent. \square

4. PROPER SCORING RULES WITH MINIMAL INCENTIVES

In this section, we examine proper scoring rules with a minimal level of incentivization in certain regions. Intuitively, these are the neighborhoods of the distribution for which we care about the most, and we want to provide the greatest incentives for the agent to be as accurate as possible in these regions.

Definition 8. A proper scoring rule satisfies (δ, k) incentivization at point $Q \in \mathbb{P}$, if for any $P \in \mathbb{P}$, $\|P - Q\| \geq \delta$, the expected payment

$$\tilde{S}(Q, Q) \geq \tilde{S}(P, Q) + k$$

LEMMA 3. Let G be the maximum payment function of a proper scoring rule S , and let H_P be the family of subtangent hyperplanes associated with S . S satisfies (δ, k) incentivization at point $Q \in \mathbb{P}$ iff for any $P \in \mathbb{P}$, $\|P - Q\| \geq \delta$,

$$G(Q) \geq H_P(Q) + k$$

EXAMPLE 1. The quadratic scoring rule defined with by maximum payment function $G(Q) = A\|Q - Q_0\|^2$ for some constants $A > 0$ and $Q_0 \in \mathbb{R}^n$ satisfies $(\delta, A\delta^2)$ incentivization $\forall Q \in \mathbb{P}$.

THEOREM 5. Suppose we want a proper scoring rule S with (δ, k) incentivization everywhere, and the maximal expected payment for every true distribution Q is non-negative. The rule which accomplishes with this and minimizes the maximum expected payment is the quadratic scoring rule defined by the maximum payment function $G(Q) = A\|Q - Q_0\|^2$ with $Q_0 = (1/n, 1/n, \dots, 1/n)$ and $A = k/\delta^2$.

Much work in this area remains to be done.

5. OTHER MISCELLANEOUS RESULTS

We state some other miscellaneous results. So far, all of them follow almost immediately from results proven before.

THEOREM 6 (ADDING PROPER SCORING RULES). Given a proper scoring rule S_1 associated with maximal payment function G_1 , and proper scoring rule S_2 associated with G_2 , the sum $S = S_1 + S_2$ is a proper scoring rule associated with maximal payment function $G = G_1 + G_2$.

COROLLARY 3. Let S be a proper scoring rule for n outcomes whose maximal payment function G is twice differentiable. S can be expressed as the sum of proper scoring rules each eliciting the relative probabilities of some two outcomes iff the cross partial derivatives of G , $\frac{\partial^2 G}{\partial i \partial j} = 0 \forall i \neq j$.

6. NEXT STEPS

- Read literature on proper scoring rules and find out why scientists like to use certain types of rules. *i.e.* the logarithmic rule.
- Suppose one wishes to modify a proper scoring rule S to make it society-aligned, and for every underlying distribution P , one is only allowed to add to the maximal expected payment at that point, and one wishes to minimize the maximum addition of expected payments. To do this for one-dimensional rules and direction +1, we simply change the maximum payment function to $G'(p) = \max(G(p), G(0))$. For multi-dimensional rules, the minimum change is unclear.
- Study the multiagent setting. Examine how to turn a society-aligned proper scoring rule into a market scoring rule, and eventually perhaps into a type of prediction market. One possible model involves trusted users who are allowed to move the probabilities against the direction that benefits society, and normal users who can only move probabilities in society's direction.

7. ACKNOWLEDGMENT

Many thanks to my C-SURF advisor Vincent Conitzer and my graduate student mentor Mingyu Guo for helpful comments and discussion. Many of the ideas in this paper are due to them.

8. REFERENCES

- [1] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, March 2007.
- [2] A. Hendrickson and R. Buehler. Proper scores for probability forecasters. *The Annals of Mathematical Statistics*, 42(6):1916–1921, 1971.
- [3] Nicolas Lambert, David M. Pennock, and Yoav Shoham. Eliciting properties of probability distributions: the highlights. *SIGecom Exch.*, 7(3):1–5, 2008.
- [4] J McCarthy. Measures of the value of information. In *Proceedings of the National Academy of Sciences*, volume 42, pages 654–655, 1956.
- [5] A. Murphy and R. Winkler. Probability forecasting in meteorology. *Journal of the American Statistical Association*, 79(387):489–500, 1984.
- [6] R. Nelson and D. Bessler. Subjective probabilities and scoring rules: Experimental evidence. *American Journal of Agricultural Economics*, 71(2):363–369, 1989.
- [7] Leonard J. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.
- [8] D. Spiegelhalter. Probabilistic prediction in patient management and clinical trials. *Stat Med*, 5(5):421–433, 1986.
- [9] J. Wolfers and E. Zitzewitz. Prediction markets. *The Journal of Economic Perspectives*, 18(2):107–126, 2004.