General Truthfulness Characterizations Via Convex Analysis

Rafael M. Frongillo
UC Berkeley
raf@cs.berkeley.edu

Ian A. Kash
MSR Cambridge
iankash@microsoft.com

ABSTRACT

We present a model of truthful elicitation which generalizes and extends both mechanisms and scoring rules. Our main result is a characterization theorem using the tools of convex analysis, of which characterizations of mechanisms and scoring rules represent special cases. Moreover, we demonstrate that a variety of existing results in the mechanism design literature are often simpler and more direct when first phrased in terms of convexity. We then generalize our main theorem to settings where agents report some alternate representation of their private information rather than reporting it directly, which gives new results about both scoring rules and mechanism design.

1. INTRODUCTION

In this paper, we examine a general model of information elicitation where a single agent is endowed with some type t that is private information and is asked to reveal it. After doing so, he receives a score that depends on both his report t' and his true type t. For reasons that will become clear, we represent this as a function $S(t')(t)$ that maps his reported type to a function that maps types to real numbers, with his score being this function applied to his true type (equivalently his reported type selects from a parameterized family of functions with the result applied to his true type). We allow S to be quite general, with the main requirement being that $S(t')(\cdot)$ is an affine[1] function of the true type t, and seek to understand when it is optimal for the agent to truthfully report his type. Given this restriction, it is immediately clear why convexity plays a central role – when an agent's type is t, the score for telling the truth is $S(t)(t) = \sup_{t'} S(t')(t)$, which is convex function of t as the pointwise supremum of affine functions.

One special case of our model is mechanism design, where $S(t')$ can be thought of as the allocation and payment given a report of t', which combine to determine the utility of the agent as a function of his type[2]. In this context, $S(t)(t)$ is the consumer surplus function, and Myerson's [18] well-known characterization states that, in single-parameter settings, a mechanism is truthful if and only if the consumer surplus function is convex and its derivative (or subgradient at points where it is not differentiable) is the allocation rule. More generally, this remains true in higher dimensions (see [1]). Note that here the restriction that $S(t')$ is affine is without loss of generality, since we may consider types as functions mapping an outcome to the agent's utility for that outcome, and the evaluation of a type on an outcome (or a distribution over outcomes) is affine in that type.

Another special case is a scoring rule, where an agent is asked to predict the distribution of a random variable and given a score based on the observed realization of that variable. In this setting, types are distributions over outcomes, and $S(t')(t)$ is the agent's expected score for a report that the distribution is t' when he believes the distribution is t. As an expectation, this score is linear in the agent's type.[3] Gneiting and Raftery [13] unified and generalized existing results in the scoring rules literature by characterizing proper scoring rules in terms of convex functions (the negative of the *Bayes risk* from decision theory) and their subgradients.

The similarities between these two models were noted by Fiat et al. [11], who gave a construction to convert mechanisms into scoring rules and vice versa. In this paper, we prove a general characterization, of which these characterizations of scoring rules and mechanisms are special cases. Our proof is essentially a combination of Gneiting and Raftery's scoring rule construction [13] with a technique from Archer and Kleinberg [1] for handling mechanisms with non-convex type spaces. Our characterization not only shows how mechanisms and scoring rules relate to each other, but also provides an understanding of how results about mechanisms relate to results about scoring rules and vice versa. In particular, many results in each literature are essentially results in convex analysis, and by phrasing them as such it is immediately clear how they apply in the other domain.

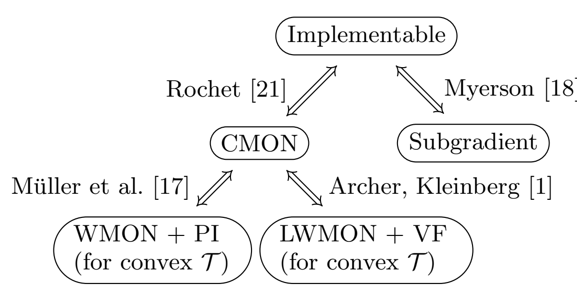

Figure 1: Proof structure of existing mechanism design literature

[1]A mapping between two vector spaces is affine if it consists of a linear transformation followed by a translation.

[2]It suffices to consider a single agent because notions of truthfulness such as dominant strategies and Bayes-Nash are phrased in terms of holding the behavior of other agents constant. See [10, 1] for additional discussion.

[3]Since distributions lie on an affine space, any affine function can be implemented as well.

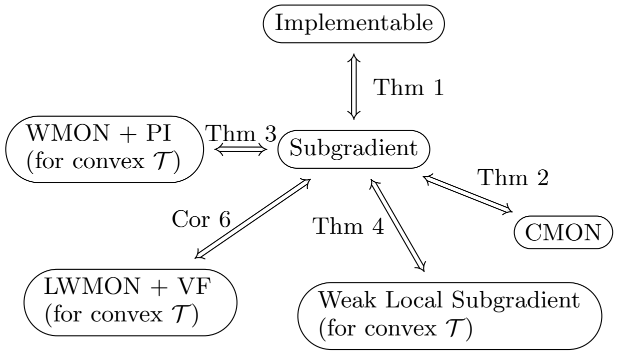

Figure 2: Our proof structure

After proving our main result, we examine a number of characterizations from the mechanism design literature that generally focus on when there exist payments that make a given allocation rule truthful. Figure 1 illustrates these characterizations and how they were proved. As it shows, several of them rely on showing equivalence to a condition known as *cyclic monotonicity*. Instead, we translate these results into convex analysis terms and prove them by showing equivalence to the property of being a family of subgradients of a convex function (see Figure 2). This has two main benefits. First, since cyclic monotonicity is a difficult condition to work with, we are able to greatly simplify the proofs of these results. Second, our proofs generally proceed by explicitly constructing the convex function, which gives a natural characterization of the payments rather than just showing that they exist. This approach also illuminates how a result by Carroll [7] about truthful mechanisms is essentially a similar characterization (see Theorem 4).

Next, we turn analyzing situations where, rather than asking for an agent's type, we wish to elicit a simpler representation of it. In the scoring rules context, this has been studied as the elicitation of properties of a distribution such as the mean or median [12, 16]. In mechanism design, this is implicit in settings such as matching where a ranking over potential matches is elicited rather than the agent's utility for them. We generalize our main result to this setting, which allows us to prove several new results. Perhaps the most notable is that if the set of distributions has positive measure in its convex hull then then there is a unique value of the property almost everywhere.

In addition to our more significant results, we get a number of smaller new results along the way essentially for free. These include:

- the fact that in mechanisms which select among a finite set of (allocation,payment) pairs, the set of sets of types that select each outcome forms not just a set of polyhedra but a power diagram;
- a relaxation of an outcome compactness assumption needed by Archer and Kleinberg [1] for allocation rules on non-convex type spaces;
- a characterization of proper scoring rules for non-convex sets of distributions; and
- a characterization that a scoring rule is proper if and only if it is locally proper

Notation..

We define $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ to be the extended real numbers. Given a set of measures M on X, a function $f : X \to \overline{\mathbb{R}}$ is M-quasi-integrable if $\int_X f(x)d\mu(x) \in \overline{\mathbb{R}}$ for all $\mu \in M$. Let $\Delta(X)$ be the set of all probability measures on X. . We denote by $\mathsf{Aff}(X \to Y)$ and $\mathsf{Lin}(X \to Y)$ the set of functions from X to Y which are affine and linear, respectively. We write $\mathsf{Conv}(X)$ to denote the convex hull of vector space X, the set of all (finite) convex combinations of elements X.

2. MODEL AND MAIN RESULT

We consider a very general model with an agent who has a given type $t \in \mathcal{T}$ and reports some possibly distinct type $t' \in \mathcal{T}$, at which point the agent is rewarded according to some score S which is affine in the true type t. This reward we call an affine score. We wish to characterize all *truthful* affine scores, those which incentivize the agent to report her true type t.

DEFINITION 1. *Any function $S : \mathcal{T} \to \mathcal{A}$, where $\mathcal{T} \subseteq \mathcal{V}$ for some vector space $\mathcal{V}$ and $\mathcal{A} \subseteq \mathsf{Aff}(\mathcal{T} \to \overline{\mathbb{R}})$, is a* affine score. *We say S is* truthful *if for all $t, t' \in T$,*

$$S(t')(t) \leq S(t)(t). \tag{1}$$

If this inequality is strict for all t, t', then S is strictly truthful.

Before stating our characterization result, we note two important applications of, and indeed motivations for, this framework: mechanism design, and proper scoring rules.

DEFINITION 2. *Given outcome space $\mathcal{O}$ and a type space $\mathcal{T} \subseteq (\mathcal{O} \to \mathbb{R})$, consisting of functions mapping outcomes to reals, a* mechanism *is a pair (f, p) where $f : \mathcal{T} \to \mathcal{O}$ is an* allocation rule *and $p : \mathcal{T} \to \mathbb{R}$ is a* payment. *The utility of the agent with type t and report t' to the mechanism is $U(t', t) = t(f(t')) - p(t')$; we say the mechanism (f, p) is* truthful *if $U(t', t) \leq U(t, t)$ for all $t, t' \in \mathcal{T}$.*

DEFINITION 3. *Given outcome space $\mathcal{O}$ and set of probability measures $\mathcal{P} \subseteq \Delta(\mathcal{O})$, a* scoring rule *is a function $S : \mathcal{P} \times \mathcal{O} \to \overline{\mathbb{R}}$. We say S is* proper *if for all $p, q \in \mathcal{P}$,*

$$\mathop{\mathbb{E}}_{o \sim p}[S(q, o)] \leq \mathop{\mathbb{E}}_{o \sim p}[S(p, o)]. \tag{2}$$

If the inequality in (2) *is strict, then S is* strictly proper.

These two models fit comfortable within our framework. For the mechanism, viewing $\mathcal{T}$ as a vector space of functions, the term $t(f(t'))$ is linear in t. Hence, $U(\cdot, t')$ is affine. Thus, the mechanism is simply the affine score with $\mathcal{A} = \{t \mapsto t(o) + c \,|\, o \in \mathcal{O}, c \in \mathbb{R}\}$ defined by $S(t')(t) := U(t, t')$. Similarly, a scoring rule S is an affine score $S : \mathcal{T} \to \mathcal{A}$ with $\mathcal{T} = \mathcal{P}$, and $\mathcal{A} \subseteq \{p \mapsto \int_{\mathcal{O}} f(o)dp(o) \,|\, f \text{ is } \mathcal{P}\text{-quasi-integrable}\}$.

Our characterization relies heavily on convex analysis, a central concept of which is the subgradient of a function.

DEFINITION 4. *Given some function $G : \mathcal{T} \to \mathbb{R}$, a function $dG_t \in \mathsf{Lin}(\mathcal{T} \to \overline{\mathbb{R}})$ is a* subgradient *to G at t if for all $t' \in \mathcal{T}$,*

$$G(t') \geq G(t) + dG_t(t' - t). \tag{3}$$

We denote by ∂G_t the set of subgradients to G at t, and $\partial G = \cup_{t \in \mathcal{T}} \partial G_t$.

For mechanism design, it is typical to assume that utilities are always real valued. However, the log scoring rule (one of the most popular scoring rules) has the property that if an agent reports that an outcome will occur with probability 0 and then that outcome does occur, the agent receives a score of $-\infty$. Essentially solely to accommodate this, we allow affine scores and subgradients to take on values from the extended reals. It is standard (see [13]) to restrict consideration to the *regular* case, where essentially only things like the log score are permitted to be infinite. In particular, an affine score S is regular if for all $t \in \mathcal{T}$ $S(t)(t) \in \mathbb{R}$ and for $t' \neq t$ $S(t')(t) \in \mathbb{R} \cup \{-\infty\}$. Similarly, a parameterized family of linear functions (e.g. a family of subgradients) $\{dG_t \in \mathsf{Lin}(\mathcal{T} \to \overline{\mathbb{R}})\}_{t\in\mathcal{T}}$ is regular if for all $t \in \mathcal{T}$ $dG_t(t) \in \mathbb{R}$ and for $t' \neq t$ $dG_{t'}(t) \in \mathbb{R}\cup\{-\infty\}$. For the remaineder of the paper we assume all affine scores and parameterized families of linear functions are regular. Note that certain results in Section 3 require a stronger assumption that the relevant parameterized families are in fact real valued rather than simply regular. A reader not interested in the details of how our framework incorporates the log scoring rule can assume that all affine scores and families are real valued throughout the paper with little loss.

We now state, and prove, our characterization theorem. Note that the proof draws techniques and insights from Gneiting and Raftery [13] and Archer and Kleinberg [1], but in such a way as to simplify the argument considerably.

THEOREM 1. *Let an affine score $S : \mathcal{T} \to \mathcal{A}$ be given. S is truthful if and only if there exists some convex $G : \mathsf{Conv}(\mathcal{T}) \to \overline{\mathbb{R}}$ with $G(\mathcal{T}) \subseteq \mathbb{R}$, and some selection of subgradients $\{dG_t\}_{t\in\mathcal{T}}$, such that*

$$S(t')(t) = G(t') + dG_{t'}(t - t'). \tag{4}$$

PROOF. It is trivial from the subgradient inequality (3) that the proposed form is in fact truthful. For the converse, we are given some truthful $S : \mathcal{T} \to \mathcal{A}$. Note first that for any $\hat{t} \in \mathsf{Conv}(\mathcal{T})$ we may write $\hat{t}$ as a finite convex combination $\hat{t} = \sum_{i=1}^{m} \alpha_i t_i$ where $t_i \in \mathcal{T}$. Now, as the range of S is affine, we may naturally extend $S(t)$ to all of $\mathsf{Conv}(\mathcal{T})$ by defining

$$S(t)(\hat{t}) = \sum_{i=1}^{m} \alpha_i S(t)(t_i). \tag{5}$$

One easily checks that this definition coincides with the given S on $\mathcal{T}$.

Now we let $G(\hat{t}) := \sup_{t\in\mathcal{T}} S(t)(\hat{t})$, which is convex as the pointwise supremum of convex (in our case affine) functions. Since S is truthful, we in particular have $G(t) = S(t)(t) \in \mathbb{R}$ for all $t \in \mathcal{T}$ by our regularity assumption. Also by truthfulness, we have for all $t, t' \in \mathcal{T}$,

$$\begin{aligned} G(t) = S(t)(t) \geq S(t')(t) &= S(t')(t') + S_\ell(t')(t - t') \\ &= G(t') + S_\ell(t')(t - t'), \end{aligned}$$

where $S_\ell(t')$ is the linear part of $S(t')$. Hence, $S_\ell(t')$ satisfies (3) for G at t' and so S is of the form (4). ☐

3. CONVEXITY IN MECHANISM DESIGN

Interpreted in the general mechanism design framework given by Definition 2, Theorem 1 says that a mechanism (f, p) is truthful if and only if the consumer surplus function $t \mapsto U(t, t)$ that it implicitly defines on the convex hull of $\mathcal{T}$ is a convex function which has a subgradient consistent with f on $\mathcal{T}$. This is a known characterization for the case of convex $\mathcal{T}$ (as well as non-convex $\mathcal{T}$ that satisfy an assumption known as outcome compactness) [1], but in practice the consumer surplus function is not always the most natural representation of a mechanism. In this section, we examine two other approaches to characterizing truthful mechanisms that have been explored in the literature and show that they have insightful interpretations in convex analysis. This interpretation has two benefits. First, by focusing on the essential convex analysis questions we are able to greatly simplify many of the proofs. Second, our proofs are constructive; in many cases we explicitly construct a consumer surplus function G, which when the mechanism is being represented by its allocation rule gives the necessary payments rather than simply providing a proof that payments exist.

3.1 Subgradient characterizations

From an algorithmic perspective, it may be more natural to focus on the design of the allocation rule f rather than the specific payments. There is a large literature that focuses on when there exists a choice of payments p to make f into a truthful mechanism (e.g. [24, 2]). Since such payments exist if and only if there is a convex function for which f is a subgradient at points in $\mathcal{T}$, this is essentially a very natural convex analysis question: when is a function f a subgradient of a convex function? Unsurprisingly, the central result in this literature is closely connected to convex analysis.

DEFINITION 5. *A family $\{dG_t \in \mathsf{Lin}(\mathcal{T} \to \mathbb{R})\}_{t\in\mathcal{T}}$ satisfies* cyclic monotonicity (CMON) *if for all finite sets $\{t_0, \ldots, t_k\} \subseteq \mathcal{T}$,*

$$\sum_{i=0}^{k} dG_{t_i}(t_{i+1} - t_i) \leq 0, \tag{6}$$

where indices are taken modulo $k+1$. We refer to the weaker condition that (6) hold for all pairs $\{t_0, t_1\}$ as weak monotonicity (WMON).

A well known characterization from convex analysis is that a function f defined on a convex set is a subgradient of a convex function on that set iff it satisfies CMON [22]. Rochet's [21] proof that a such payments exists on a possibly non-convex $\mathcal{T}$ iff f satisfies CMON is effectively a proof of the following generalization of this theorem.

THEOREM 2. *A family $\{dG_t \in \mathsf{Lin}(\mathcal{T} \to \mathbb{R})\}_{t\in\mathcal{T}}$ satisfies CMON if and only if there exists a convex $G : \mathsf{Conv}(\mathcal{T}) \to \mathbb{R}$ such that dG_t is a subgradient of G at t for all $t \in \mathcal{T}$.*

Rochet notes that his proof is adapted from the one given in Rockafellar's text [22] of the weaker theorem where $\mathcal{T}$ is restricted to be convex. We adapt Rochet's proof to highlight how its core is a construction of G.

PROOF. Given such a G, by (3) we have $dG_{t_i}(t_{i+1} - t_i) \leq G(t_{i+1}) - G(t_i)$. Summing gives (6). Given such a family $\{dG_t\}_{t\in\mathcal{T}}$, fix some $t_0 \in \mathcal{T}$ and define

$$G(t) = \sup_{\substack{\{t_1,\ldots,t_{k+1}\}\subseteq\mathcal{T},\\ t_{k+1}=t}} \sum_{i=0}^{k} dG_{t_i}(t_{i+1} - t_i), \tag{7}$$

where $\{t_1, \ldots, t_k\}$ denotes any finite sequence (k is not fixed).

By CMON, for $t \in \mathcal{T}$ this sum is upper bounded by $-dG_t(t_0 - t)$. Thus, the supremum is finite on $\mathcal{T}$. G is a pointwise supremum of convex functions, so is convex. By convexity, G is also finite on $\mathsf{Conv}(\mathcal{T})$.

For any $t \in \mathcal{T}$ and $t' \in \mathsf{Conv}(\mathcal{T})$,

$$\begin{aligned} G(t) + dG_t(t'-t) &= dG_t(t'-t) + \sup_{\substack{\{t_1,\ldots,t_{k+1}\}\subseteq\mathcal{T},\\ t_{k+1}=t}} \sum_{i=0}^{k} dG_{t_i}(t_{i+1}-t_i) \\ &= \sup_{\substack{\{t_1,\ldots,t_{k+1}\}\subseteq\mathcal{T},\\ t_k=t,\ t_{k+1}=t'}} \sum_{i=0}^{k} dG_{t_i}(t_{i+1}-t_i) \\ &\leq \sup_{\substack{\{t_1,\ldots,t_{k+1}\}\subseteq\mathcal{T},\\ t_{k+1}=t'}} \sum_{i=0}^{k} dG_{t_i}(t_{i+1}-t_i) \\ &= G(t'), \end{aligned}$$

so dG_t satisfies (3). $\square$

A number of papers have sought simpler and more natural conditions than CMON that are necessary and sufficient in special cases, e.g. [24, 1, 2]. These results are typically proven by showing they are equivalent to CMON. However, it is much more natural to directly construct the relevant G. This also often has the advantage of providing a characterization of the payments that is more intuitive than the supremum in Rochet's construction. As an example, we show one such result has a simple proof using our framework.

As in Myerson's [18] construction for the single-parameter case, we construct a G by integrating over dG_t. In particular, for any two types x and y our construction makes use of the line integral

$$\int_{L_{xy}} dG_t(y-x)dt = \int_0^1 dG_{(1-t)x+ty}(y-x)dt.$$

As Berger et al. [5] and Ashlagi et al. [2] observed, if $\{dG_t\}_{t\in\mathcal{T}}$ satisfies WMON and $\mathcal{T}$ is convex, this (Riemann) integral is well defined because it is the integral of a monotone function. If these line integrals vanish around all triangles (equivalently $\int_{L_{xy}} dG_t(y-x)dt + \int_{L_{yz}} dG_t(z-y)dt = \int_{L_{xz}} dG_t(z-x)dt)$) we say $\{dG_t\}$ satisfies *path independence.*

Theorem 3 (adapted from [17]). *For convex $\mathcal{T}$, a family $\{dG_t \in \mathsf{Lin}(\mathcal{T} \to \mathbb{R})\}_{t\in\mathcal{T}}$ is a subgradient of a convex function if and only if $\{dG_t\}_{t\in\mathcal{T}}$ satisfies WMON and path independence.*

Proof. Given a convex function G and $\{dG_t\}$, $\{dG_t\}$ satisfies CMON and thus WMON. Path independence also follows from convexity (Rockafellar [22] p. 232). Now given a $\{dG_t\}$ that satisfies WMON and path independence, fix a type $t_0 \in \mathcal{T}$ and define $G(t') = \int_{L_{t_0 t'}} dG_t(t'-t_0)dt$ (well defined by WMON as the integral of a monotone function). Given $x, y, z \in \mathcal{T}$ such that $z = \lambda x + (1-\lambda)y$, by path independence and the linearity of dG_z we have

$$\begin{aligned} &\lambda G(x) + (1-\lambda)G(y) \\ &= G(z) + \lambda \int_{L_{zx}} dG_t(x-z)dt + (1-\lambda)\int_{L_{zy}} dG_t(y-z)dt \\ &\geq G(z) + \lambda dG_z(x-z) + (1-\lambda)dG_z(y-z) = G(z), \end{aligned}$$

so G is convex. Similarly, for $x, y \in \mathcal{T}$, dG_t satisfies (3) because

$$dG_x(y-x) \leq \int_{L_{xy}} dG_t(y-x)dt = G(y) - G(x). \quad \square$$

3.2 Local convexity

In many settings, it is natural to specify mechanisms in terms of an algorithm that computes the allocation and payment. As it is often easier to reason about the behavior of algorithms given small changes to their input rather than arbitrary changes, several authors have sought to characterize truthful mechanisms using local conditions [1, 5, 7].

We show in this section how many of these results are in essence a consequence of a more fundamental statement, that convexity is an inherently local property. For example, in the twice differentiable case it can be verified by determining whether the Hessian is positive semidefinite at each point. We start with a local convexity result, and use it to show that an affine score is truthful if and only if it satisfies a very weak local truthfulness property introduced by Carroll [7]. Afterwards we turn to a similar characterization by Archer and Kleinberg [1].

Definition 6. *Given a function G, a family $\{dG_t \in \mathsf{Lin}(\mathcal{T} \to \overline{\mathbb{R}})\}_{t\in\mathcal{T}}$ is a* weak local subgradient with respect to G (G-WLSG) *if for all $t \in \mathcal{T}$ there exists an open neighborhood U_t of t such that for all $t' \in U_t$,*

$$G(t) \geq G(t') + dG_{t'}(t-t') \quad \text{and} \quad G(t') \geq G(t) + dG_t(t'-t). \tag{8}$$

We now show that G-WLSG is a sufficient condition for a family of functions to be a subgradient of G. The proof is heavily inspired by Carroll [7].

Theorem 4. *Let $\mathcal{T}$ be convex. A family $\{dG_t \in \mathsf{Lin}(\mathcal{T} \to \overline{\mathbb{R}}\}_{t\in\mathcal{T}}$ is a subgradient of a given function G if and only if it satisfies G-WLSG.*

Proof (Adapted from [7]). As usual, the forward direction is trivial. For the other, let $t, t' \in \mathcal{T}$ be given; we show that the subgradient inequality for $dG_{t'}$ holds at t. By compactness of $\mathsf{Conv}(\{t, t'\})$, we have a finite set $t_i = \alpha_i t' + (1-\alpha_i)t$, where $0 = \alpha_0 \leq \cdots \leq \alpha_{k+1} = 1$, such that G-WLSG holds between each t_i and t_{i+1}. (The cover $\{U_s \mid s \in \mathsf{Conv}(\{t,t'\})\}$ has a finite subcover. Take t_{2i} from the subcover and $t_{2i+1} \in U_{t_{2i}} \cap U_{t_{2i+2}}$.) By the WLSG condition (8), we have for each i,

$$0 \geq G(t_{i+1}) - G(t_i) + dG_{t_{i+1}}(t_i - t_{i+1}) \tag{9}$$

$$0 \geq G(t_i) - G(t_{i+1}) + dG_{t_i}(t_{i+1} - t_i). \tag{10}$$

Now using the identity $t_{i+1} - t_i = (\alpha_{i+1} - \alpha_i)(t' - t)$ and adding $\alpha_i/(\alpha_{i+1}-\alpha_i)$ times (9) to $\alpha_{i+1}/(\alpha_{i+1}-\alpha_i)$ times (10), we have

$$0 \geq G(t_i) - G(t_{i+1}) + \alpha_i dG_{t_i}(t'-t) - \alpha_{i+1} dG_{t_{i+1}}(t'-t). \tag{11}$$

Summing (11) over $0 \leq i \leq k$ gives

$$0 \geq G(t_0) - G(t_{k+1}) + \alpha_0 dG_{t_0}(t'-t) - \alpha_{k+1} dG_{t_{k+1}}(t'-t),$$

which when recalling our definitions for α_i and t_i yields the result. $\square$

The WLSG condition translates to an analogous notion in terms of truthfulness, *weak local truthfulness.*

DEFINITION 7. *An affine score is* weakly locally truthful *if for all $t \in \mathcal{T}$ there exists some open neighborhood U_t of t, such that truthfulness holds between t and every $t' \in U_t$, and vice versa. That is,*

$$\forall t \in \mathcal{T},\ \forall t' \in U_t,\ \ S(t')(t) \le S(t)(t) \ \ \text{and}\ \ S(t)(t') \le S(t')(t'). \tag{12}$$

COROLLARY 5 (GENERALIZATION OF CARROLL [7]). *An affine score $S : \mathcal{T} \to \mathcal{A}$ for convex $\mathcal{T}$ is truthful if and only if it is weakly locally truthful.*

PROOF. Defining $G(t) := S(t)(t)$, by weak local truthfulness we may write

$$G(t) = S(t)(t) \ge S(t')(t) = G(t') + S_\ell(t')(t - t')$$
$$G(t') = S(t')(t') \ge S(t)(t') = G(t) + S_\ell(t)(t' - t),$$

where $S_\ell(\cdot)$ is the linear part of $S(\cdot)$. This says that $dG_t = S_\ell(t)$ satisfies G-WLSG; the rest follows from Theorem 4 and Theorem 1. □

Finally, in the spirit of Section 3.1, Archer and Kleinberg [1] characterized local conditions under which an allocation rule can be made truthful. A key condition from their paper is *vortex-freeness*, which is a condition they show to be equivalent to local path independence. The other condition, local WMON, means that WMON holds in some neighborhood around each type. Their result follows directly from an analogous characterization of subgradients.

COROLLARY 6. *Let $\mathcal{T}$ be convex. A family $\{dG_t \in \mathsf{Lin}(\mathcal{T} \to \mathbb{R})\}_{t\in\mathcal{T}}$ is a subgradient of a convex function if and only if it satisfies local WMON and is vortex-free.*

PROOF. We prove the reverse direction; suppose $\{dG_t\}_{t\in\mathcal{T}}$ satisfies local WMON and is vortex-free. From Lemma 3.5 of [1] we have that vortex-freeness is equivalent to path independence, so by Theorem 3 for all t there exists some open U_t such that $\{dG_{t'}\}_{t'\in U_t}$ is the subgradient of some convex function $G^{(t)} : U_t \to \mathbb{R}$. We need only show the existence of some G such that $\{dG_t\}_{t\in\mathcal{T}}$ is the subgradient of G on each U_t; the rest follows from Theorem 4.

Fix some $t_0 \in \mathcal{T}$ and define $G(t) = \int_{L_{t_0\,t}} dG_{t'} dt'$, which is well defined by compactness of $\mathsf{Conv}(\{t_0, t\})$ and the fact that a locally increasing real-valued function is increasing. But for each t' and $t \in U_{t'}$ we can also write $G^{(t')}(t) = \int_{L_{t'\,t}} dG_{t''} dt''$ by [22, p. 232], and now by path independence we see that G and $G^{(t')}$ differ by a constant. Hence $\{dG_t\}_{t\in\mathcal{T}}$ must be a subgradient of G on $U_{t'}$ as well, for all $t' \in \mathcal{T}$. □

4. EXTENSION TO PROPERTIES

In many settings, it is difficult, or even impossible, to have agents report an entire type $t \in \mathcal{T}$. For example, when $\mathcal{O} = \mathbb{R}$, then a mechanism with $\mathcal{T} = (\mathcal{O} \to \mathbb{R})$ requires agents to submit an infinite-dimensional type, and even type spaces which are exponential in size can be problematic. It is therefore natural to consider a model of truthful reporting where agents provide some sort of summary information about their type. Such a model has been studied in the scoring rules literature, where one wishes to elicit some statistic, or *property*, of a distribution [12, 16]. We follow this line of research, and extend our affine score framework to accept reports from a different (intuitively, smaller) space than $\mathcal{T}$. Our results shed light on the structure of properties and their affine scores.

4.1 General setting

We wish to generalize the notion of an affine score to accept reports from a space R which is different from $\mathcal{T}$. To even discuss truthfulness in this setting, we need a notion of a truthful report r for a given type t. We encapsulate this notion by a general multivalued function which specifies only and all of the correct values for t.

DEFINITION 8. *Let R be some given report space. A* property *is a multivalued map $\Gamma : \mathcal{T} \rightrightarrows R$ which associates a nonempty set of correct report values to each type. We let $\Gamma_r := \{t \in \mathcal{T} \,|\, r \in \Gamma(t)\}$ denote the set of types t corresponding to report value r.*

One can think of Γ_r as the "level set" of Γ corresponding to value r. This concept will be useful mainly when we consider finite-valued properties in Section 4.2.

We extend the notion of an affine score to this setting, where the report space is R instead of $\mathcal{T}$ itself.

DEFINITION 9. *An affine score $S : R \to \mathcal{A}$* elicits *a property $\Gamma : \mathcal{T} \rightrightarrows R$ if for all t,*

$$\Gamma(t) = \operatorname*{argsup}_{r\in R} S(r)(t). \tag{13}$$

A property $\Gamma : \mathcal{T} \rightrightarrows R$ is elicitable *if there exists some affine score $S : R \to \mathcal{A}$ which elicits Γ.*

As before, we need a notion of regularity here as well – an affine score S is Γ-*regular* if $S(r)(t) < \infty$ always and $S(r)(t) \in \mathbb{R}$ whenever $r \in \Gamma(t)$.

The simplest way to come up with an elicitable property is to induce one from an affine score. For any $S : R \to \mathsf{Aff}(\mathcal{T} \to \mathbb{R})$, the property

$$\Gamma_S : t \to \operatorname*{argsup}_{r\in R} S(r)(t) \tag{14}$$

is trivially elicited by S.

Observe also that any affine score S eliciting Γ gives rise to a truthful affine score in the original sense – in fact, this is a version of the so-called *revelation principle*. For each t let $r_t \in \Gamma(t)$ be a report choice for t; then the affine score $S^{\mathcal{T}}(t')(t) := S(r_{t'})(t)$ is truthful. Moreover, by our choices of $\{r_t\}$, we have

$$G(t) := \sup_{t'\in\mathcal{T}} S^{\mathcal{T}}(t')(t) = \sup_{r\in R} S(r)(t). \tag{15}$$

Of course, in general, $S^{\mathcal{T}}$ will not be strictly truthful, since by definition, any reports t', t'' with $r_{t'} = r_{t''}$ (which can happen when $t', t'' \in \Gamma_r$ for some r) will have $S^{\mathcal{T}}(t') \equiv S^{\mathcal{T}}(t'')$. Thus we may think of a property as *refining* the notion of strictness for a truthful affine score. The connection we will draw in Theorem 8 is that, in light of (15), a property Γ therefore specifies the portions of the domain of $\mathcal{T}$ where G must be "flat".

To get at the connection between properties and "flatness", we start with a technical lemma which shows that having the same subgradient at two different points implies that G is flat in between.

LEMMA 7. *For convex $\mathcal{T}$, let $G : \mathcal{T} \to \mathbb{R}$ be convex, and let $d \in \partial G_t$ for some $t \in \mathcal{T}$. Then for all $t' \in \mathcal{T}$,*

$$d \in \partial G_{t'} \iff G(t) - G(t') = d(t - t').$$

PROOF. First, the forward direction. Applying the subgradient inequality (3) at t' for $dG_t = d$ and at t for $dG_{t'} = d$, we have

$$G(t') \geq G(t) + d(t' - t)$$
$$G(t) \geq G(t') + d(t - t'),$$

from which the result follows. For the converse, assume $G(t) = G(t') + d(t - t')$ and let $t'' \in \mathcal{T}$ be arbitrary. Then since $d \in \partial G_t$ we have

$$\begin{aligned} G(t') + d(t'' - t') &= G(t') + d(t'' - t) + d(t - t') \\ &= G(t) + d(t'' - t) \\ &\leq G(t''), \quad \square \end{aligned}$$

We are now ready to state our characterization in this setting, which in essence says that eliciting a property Γ is equivalent to eliciting subgradients of a convex function G.

THEOREM 8. *Let property $\Gamma : \mathcal{T} \rightrightarrows R$ and Γ-regular affine score $S : R \to \mathcal{A}$ be given. Then S elicits Γ if and only if there exists some convex $G : \mathsf{Conv}(\mathcal{T}) \to \overline{\mathbb{R}}$ with $G(\mathcal{T}) \subseteq \mathbb{R}$ and $\varphi : R \to \partial G$ satisfying $r \in \Gamma(t) \iff \varphi(r) \in \partial G_t$, such that*

$$S(r)(t) \; = \; G(t_r) + \varphi(r)(t - t_r), \tag{16}$$

where $\{t_r\}_{r\in R}$ satisfies $r \in \Gamma(t_r)$ for all r.

PROOF. For the forward direction, assume that affine score S elicits Γ. For each r, we may extend $S(r)$ to all $\hat{t} \in \mathsf{Conv}(\mathcal{T})$ by linearity as in Theorem 1, whence we may define $G(\hat{t}) := \sup_{r\in R} S(r)(\hat{t})$, which is finite for $\hat{t} \in \mathcal{T}$ as S is Γ-regular. We wish to show that the choice $\varphi : r \mapsto S_\ell(r)$ suffices, where S_ℓ denotes the linear part of S.

By standard arguments (cf. Proposition 8.12 of [23]), the subgradients to a convex function $F : X \to \mathbb{R}$ at $x \in X$ are precisely the gradients of its affine supports at x. More formally, let $\mathrm{AM}(F) = \{a \in \mathsf{Aff}(X \to \mathbb{R}) \,|\, \forall x\; F(x) \geq a(x)\}$ be the set of affine functions minorizing F, and let $\mathrm{AS}(F, x) = \{a \in \mathrm{AM}(F) \,|\, a(x) = F(x)\}$ be the set of affine supports of F at x. Then we have $\partial F_x = \{\partial a \,|\, a \in \mathrm{AS}(F, x)\}$.

By definition of G, we have $S(r) \in \mathrm{AM}(G)$. Noting that $\partial S(r)_t = S_\ell(r)$, we have for all $t \in \mathcal{T}$ and $r \in R$,

$$\begin{aligned} r \in \Gamma(t) &\iff S(r)(t) = \sup_{r'\in R} S(r')(t) \\ &\iff S(r)(t) = G(t) \\ &\iff S(r)(t) \in \mathrm{AS}(G, t) \\ &\iff S_\ell(r) \in \partial G_t, \end{aligned}$$

where we use elicitation in the first biconditional. Finally, using Γ-regularity and the definition of t_r, we show the form (16):

$$G(t_r) + S_\ell(r)(t - t_r) = S(r)(t_r) + S_\ell(r)(t - t_r) = S(r)(t).$$

For the converse, let G, φ, and t_r be given, and assume S has the form (16). First note that by definition of $\{t_r\}$, and by assumption on φ, we have $\varphi(r) \in \partial G_{t_r}$ for all r. We claim that G dominates S, meaning for all $r \in R$ and all $t \in \mathcal{T}$, $G(t) \geq S(r)(t)$; this follows from the definition of S and the subgradient inequality (3) applied to $\varphi(r)$ at t.

Now by Lemma 7, we have

$$\begin{aligned} \varphi(r) \in \partial G_t &\iff G(t) - G(t_r) = \varphi(r)(t - t_r) \\ &\iff G(t) = S(r)(t). \end{aligned} \tag{17}$$

By definition of a property, we have for all $t \in \mathcal{T}$ that $\Gamma(t) \neq \emptyset$, so by the above $G(t) = S(r)(t)$ for some r. Combining this with the fact that G dominates S, we have $G(t) = \sup_r S(r)(t)$ for all $t \in \mathcal{T}$. Putting this together with (17) and our assumption that $r \in \Gamma(t) \iff \varphi(r) \in \partial G_t$, we have

$$r \in \Gamma(t) \iff G(t) = S(r)(t) \iff r \in \operatorname*{argsup}_{r'} S(r')(t) \quad \square$$

Returning to the above discussion, by focusing on $S^{\mathcal{T}}$ instead of S, we now see how properties are essentially a refinement of strictness. In fact, one can can formalize this by considering that, up to a remapping of R, $S^{\mathcal{T}}$ is strictly truthful if and only if S elicits $\Gamma : t \mapsto \{t\}$, and $S^{\mathcal{T}}$ is truthful if and only if S elicits some Γ such that $t \in \Gamma(t)$ for all $\mathcal{T}$. Hence, Theorem 8 is actually a generalization of our main result, Theorem 1.

Returning our focus to R, our characterization sheds light on the structure of elicitable properties. For example, we may apply the fact that convex functions are differentiable almost everywhere to yield a very strong result about the structure of elicitable properties.

COROLLARY 9. *If $\Gamma : \mathcal{T} \rightrightarrows R$ is an elicitable property, $\mathcal{T}$ is of positive measure in $\mathsf{Conv}(\mathcal{T})$, and no property value is redundant ($\forall r_1, r_2 \in R, \Gamma_{r_1} \not\subseteq \Gamma_{r_2}$) then $|\Gamma(t)| = 1$ almost everywhere.*

4.2 Finite case

We now turn to the case where R is finite. In this scoring rules literature, this case has been thought of as eliciting answers to multiple-choice questions [15]. There are also natural applications to mechanism design, which we mention below.

Assume throughout that R is finite and that $\mathcal{T}$ is a subset of a vector space $\mathcal{V}$ endowed with an inner product, so that we may write $\langle t, t' \rangle$ and in particular $\|t\|^2 = \langle t, t \rangle$. In this more geometrical setting, we will use the concept of a power diagram from computational geometry.

DEFINITION 10. *Given points $p_i \in \mathcal{V}$, called* sites*, and weights $w(p_i) \in \mathbb{R}$ for $i \in \{1, \ldots, m\}$, a* power diagram *is a collection cells $\mathrm{cell}(p_i) \subseteq \mathcal{T}$ defined by*

$$\mathrm{cell}(p_i) = \{t \in \mathcal{T} \,|\, i = \operatorname*{argmin}_j \{\|p_i - t\|^2 - w(p_j)\}\}. \tag{18}$$

The following result is a generalization of Theorem 4.1 of [15], and is essentially a direct application of Proposition 1 from [3].

THEOREM 10. *A property $\Gamma : \mathcal{T} \rightrightarrows R$ for R finite is elicitable if and only if the level sets $\{\Gamma_r\}_{r\in R}$ form a power diagram. Moreover, every $S : R \to \mathcal{A}$ can be written*

$$S(r)(t) = \|t\|^2 - \|p_r - t\|^2 + w(p_r), \tag{19}$$

for some choice of sites $\{p_r \in \mathcal{V}\}_{r\in R}$ and weights $w(p_r) \in \mathbb{R}$.

PROOF. Given sites $\{p_r\}$ and weights $w(p_r)$, let S take the form (19). Then taking $\Gamma(t) := \operatorname{argsup}_r S(r)(t)$, we see $\Gamma_r = \mathrm{cell}(p_r)$.

Given affine score S eliciting Γ, note that since we are in an inner product space, we may write $S(r)(t) = \langle x_r, t \rangle + c_r$ for $x_r \in \mathcal{V}$ and $c_r \in \mathbb{R}$. Now take $p_r = x_r/2$ and $w(p_r) = \|p_r\|^2 + c_r$; then

$$\|t\|^2 - \|p_r - t\|^2 + w(p_r) = 2\,\langle p_r, t \rangle - \|p_r\|^2 + c_r + \|p_r\|^2 = S(r)(t),$$

which simultaneously proves the form (19) and that $\Gamma_r = \text{cell}(p_r)$. $\square$

We can now combine Theorem 10 and Corollary 5 to obtain the following simple truthfulness check for the finite case.

THEOREM 11. *For finite R, an affine score $S : R \to \mathcal{A}$ elicits Γ if and only if for all r, r' such that $\Gamma_r \cap \Gamma_{r'} \neq \emptyset$, we have $S(r')(t) < S(r)(t)$ for all $t \in \Gamma_r \setminus \Gamma_{r'}$.*

We now apply our results to mechanism design with a finite set of allocations. One example of such a setting is ordinal utilities over a finite set of outcomes. Let $\mathcal{T} = \mathcal{O} \to \mathbb{R}$ and let R be the set of preference orderings over $\mathcal{O}$, which we represent as permutations π of $\mathcal{O}$ (higher preference being first). Then a natural property is

$$\Gamma(t) = \{\pi \in R \mid \forall i < j,\ t(\pi_i) \leq t(\pi_j)\}, \tag{20}$$

the set of rankings consistent with type t. More generally, truncated rankings or any mechanism with a finite message space yields such a property.

Carroll [7] points out that it is known that for all finite sets of allocations, the set of types for which a particular allocation is optimal forms a polyhedron and then shows that for all convex set of types it is sufficient to check incentive compatibility constraints between polyhedra that intersect at a face. Our results in this section show something stronger: the types form not just an arbitrary set of polyhedra, but a power diagram. However, the sufficiency condition we get is slightly weaker, as it requires checking incentive compatibility between types that intersect at a vertex as well as at a face.

5. ADDITIONAL APPLICATIONS

In this section, we demonstrate the power of our framework with several additional applications drawn from the scoring rules literature.

5.1 Two Results About Scoring Rules

The role of convexity in the scoring rules literature dates back to 1971 when Savage showed how to construct a scoring rule from a convex function [25]. The general, modern characterization of scoring rules is due to Gneiting and Raftery [13], who used a more nuanced convex analysis approach to clarify and generalize a number of previous characterizations, including Savage [25] and Schervish [26]. In this section, we show that the Gneiting and Raftery characterization is a simple special case of Theorem 1, and moreover that our results enable several generalizations in the scoring rule setting.

Letting $\mathcal{F}$ be the set of $\mathcal{P}$-quasi-integrable functions $f : \mathcal{O} \to \overline{\mathbb{R}}$, we apply our characterization with $\mathcal{T} = \mathcal{P}$ and $\mathcal{A} = \{p \mapsto \int_{\mathcal{O}} f(o)\, dp(o) \mid f \in \mathcal{F}\}$.

COROLLARY 12. *For an arbitrary set $\mathcal{P} \subseteq \Delta(\mathcal{O})$ of probability measures, a regular[4] scoring rule $S : \mathcal{P} \times \mathcal{O} \to \overline{\mathbb{R}}$ is proper if and only if there exists a convex function $G : \mathsf{Conv}(\mathcal{P}) \to \mathbb{R}$ with functions $G_p \in \mathcal{F}$ such that*

$$S(p, o) = G(p) + G_p(o) - \int_{\mathcal{O}} G_p(o)\, dp(o), \tag{21}$$

where $G_p : q \mapsto \int_{\mathcal{O}} G_p(o)\, dq(o)$ is a subgradient of G for all $p \in \mathcal{P}$.

PROOF. The given form is truthful by the subgradient inequality. Let $S : \mathcal{T} \to \mathcal{A}$ be a given truthful affine score. Since $S(p) \in \mathcal{A}$, we have some $f_p \in \mathcal{F}$ generating $S(p)$. We can therefore use $G_p : q \mapsto \int_{\mathcal{O}} f_p(o)\, dq(o)$ as the subgradients in the proof of Theorem 1, thus giving us the desired form. $\square$

Corollary 12 immediately generalizes their characterization to the case where $\mathcal{P}$ is not convex. While this case is not of central importance to the theory of scoring rules, it is sometimes a natural case to consider. For example, Babaioff et al. [4] examine when proper scoring rules can have the additional property that uninformed experts do not wish to make a report (have a negative expected utility), while informed experts do wish to make one. They show that this is possible in some settings where the space of reports is not convex. Thus our characterization shows that, despite not needing to ensure properness on on reports outside $\mathcal{P}$, essentially the only possible scoring rules are still those that are proper on all of $\Delta(\mathcal{O})$.

While local truthfulness is a natural property to examine in mechanism design, it is not obvious how it is as useful for scoring rules since scoring rule designers tend to operate under significantly fewer constraints than mechanism designers. Nevertheless, to results from Section 3.2 apply, so Corollary 5 shows that local properness is equivalent to global properness for scoring rules on convex $\mathcal{P}$.

COROLLARY 13. *For a convex set $\mathcal{P} \subseteq \Delta(\mathcal{O})$ of probability measures, a scoring rule $S : \mathcal{P} \times \mathcal{O} \to \overline{\mathbb{R}}$ is proper if and only if it is (weakly) locally proper.*

5.2 Decision Rules

Theorem 1 also generalizes Gneiting and Raftery's [13] characterization to settings beyond eliciting a single distribution. For example, a line of work has considered a setting where a decision maker needs to select from a finite set $\mathcal{D}$ of decisions and so desires to elicit the distribution over outcomes conditional on selecting each alternative [20, 9, 8]. Since only one decision will be made and so only one conditional distribution can be sampled, simply applying a standard proper scoring rule generally does not result in truthful behavior. Applying Theorem 1 to this setting characterizes what expected scores must be, from which many of the results in these papers follow. As it is not our main focus, we refrain from introducing the model necessary to explicitly state a characterization result similar to Corollary 12.

6. DISCUSSION

We have presented a model of truthful elicitation which generalizes and extends both mechanisms and scoring rules. On the mechanism design side, we have seen how our framework provides simpler and more constructive proofs of a number of known results, some of which (as we saw in Section 5.1) lead to new results about scoring rules. We then generalized our model to allow reports of a succinct representation of an agent's private information rather than that information itself, a topic that has been studied in the scoring rules literature about eliciting properties of distributions. This led us to new results in both mechanism design and scoring rules.

[4]This is the same concept as with affine scores: scores cannot be ∞ and only incorrect reports can yield $-\infty$.

Our analysis makes use of the fact that $S(t')(t)$ is affine in t to ensure that $G(t) = \sup_{t'} S(t')(t)$ is a convex function. However, this property continues to hold if $S(t')(t)$ is instead a convex function of t. Thus, a natural direction for future work is to investigate characterizations of convex scores. While mechanisms can always be represented as affine functions by taking the types to be functions from allocations to $\mathbb{R}$, it may be more natural to treat the type as a parameter of a (convex) utility function. While many such utility functions are affine (e.g. dot-product valuations), others such as Cobb-Douglas functions are not. Berger, Müller, and Naeemi [5, 6] have investigated such functions and given characterizations that suggest a more general result is possible. Another potential application is scoring rules for alternate representations of uncertainty, several of which result in a decision maker optimizing a convex function [14].

In one sense getting such a characterization is straightforward. In the affine case we want $S(t')(t)$ to be an affine function such that $S(t')(t) \leq G(t)$ and $S(t')(t') = G(t')$. Since we have fixed its value at a point, the only freedom we have is in the linear part of the function, and being a subgradient is exactly the definition of such a linear function. So while our characterization of affine scores is in some sense vacuous, it is also powerful in that it allows us to make use of the tools of convex analysis. A similarly vacuous characterization is possible for the convex case: $S(t')(t)$ is a convex function such that $S(t')(t) \leq G(t)$ and $S(t')(t') = G(t')$. The challenge is to find a way to state it that is useful and naturally handles constraints such as those imposed by the form of a utility function.

Theorem 10 shows that scoring rules for finite properties are essentially equivalent the weights and points that induce a power diagram. So the natural open problem of characterizing all scoring rules for a given finite property in some effective fashion is equivalent to the problem of characterizing all weights and points that induce a particular power diagram. We are uncertain about whether this problem has been studied, but it is a natural question in that context as well. As power diagrams are known to be connected to the spines of amoebas in algebraic geometry, tropical hypersurfaces in tropical geometry, and aspects of toric geometry used by string theorists [27] there may be interesting corresponding characterization questions in those fields as well.

While our examples have focused on mechanism design and scoring rules, another interesting direction to pursue is other settings where our results may be applicable. One natural domain that is closely related to scoring rules for properties is the literature on M-estimators in machine learning, statistics and economics. Essentially, this literature takes a loss function (i.e. a scoring rule) and asks what property it elicits. For example, the mean is an M-estimator induced by the squared error loss function. Some work in this literature (e.g. [19]) requires that the loss function satisfy certain properties, and our results may be useful in characterizing and supplying such loss functions.

Acknowledgments

This paper is based on work done while Rafael was an intern at Microsoft Research Cambridge. We would like to thank Gabriel Carroll, Yiling Chen, Peter Key, and Greg Stoddard for helpful suggestions.

7. REFERENCES

[1] A. Archer and R. Kleinberg. Truthful germs are contagious: a local to global characterization of truthfulness. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 21—30, 2008.

[2] I. Ashlagi, M. Braverman, A. Hassidim, and D. Monderer. Monotonicity and implementability. *Econometrica*, 78(5):1749—1772, 2010.

[3] F. Aurenhammer. Power diagrams: properties, algorithms and applications. *SIAM Journal on Computing*, 16(1):78—96, 1987.

[4] M. Babaioff, L. Blumrosen, N. S. Lambert, and O. Reingold. Only valuable experts can be valued. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 221—222, 2011.

[5] A. Berger, R. M\"uller, and S. Naeemi. Characterizing incentive compatibility for convex valuations. *Algorithmic Game Theory*, pages 24—35, 2009.

[6] A. Berger, R. Muller, and S. H. Naeemi. PathâĂŞMonotonicity and incentive compatibiliy. 2010.

[7] G. Carroll. When are local incentive constraints sufficient? *Econometrica*, 80(2):661—686, 2012.

[8] Y. Chen, I. Kash, M. Ruberry, and V. Shnayder. Decision markets with good incentives. *Internet and Network Economics*, pages 72—83, 2011.

[9] Y. Chen and I. A. Kash. Information elicitation for decision making. *AAMAS*, 2011.

[10] K. S. Chung and J. C. Ely. Ex-post incentive compatible mechanism design. *URL http://www. kellogg. northwestern. edu/research/math/dps/1339. pdf. Working Paper*, 2006.

[11] A. Fiat, A. R. Karlin, E. Koutsoupias, and A. Vidali. Approaching utopia: Strong truthfulness and externality-resistant mechanisms. *arXiv preprint arXiv:1208.3939*, 2012.

[12] T. Gneiting. Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494):746–762, 2011.

[13] T. Gneiting and A. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359—378, 2007.

[14] J. Halpern. *Reasoning about uncertainty*. Cambridge, MA: MIT Press, 2003.

[15] N. Lambert and Y. Shoham. Eliciting truthful answers to multiple-choice questions. In *Proceedings of the 10th ACM conference on Electronic commerce*, page 109âĂŞ118, 2009.

[16] N. S. Lambert, D. M. Pennock, and Y. Shoham. Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, page 129âĂŞ138, 2008.

[17] R. M\"uller, A. Perea, and S. Wolf. Weak monotonicity and BayesâĂŞNash incentive compatibility. *Games and Economic Behavior*, 61(2):344—358, 2007.

[18] R. B. Myerson. Optimal auction design. *Mathematics of operations research*, pages 58—73, 1981.

[19] S. Negahban, P. Ravikumar, M. J. Wainwright, and

B. Yu. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. *Statistical Science*, 2012.

[20] A. Othman and T. Sandholm. Decision rules and decision markets. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 625—632, 2010.

[21] J. C. Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16(2):191—200, 1987.

[22] R. Rockafellar. *Convex analysis*, volume 28 of *Princeton Mathematics Series*. Princeton University Press, 1970.

[23] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, Oct. 2011.

[24] M. Saks and L. Yu. Weak monotonicity suffices for truthfulness on convex domains. In *Proceedings of the 6th ACM conference on Electronic commerce*, pages 286—293, 2005.

[25] L. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, page 783âĂŞ801, 1971.

[26] M. J. Schervish. A general method for comparing probability assessors. *The Annals of Statistics*, 17(4):1856âĂŞ1879, 1989.

[27] M. Van Manen and D. Siersma. Power diagrams and their applications. *arXiv preprint math/0508037*, 2008.