

Efficient Market Making via Convex Optimization, and a Connection to Online Learning

Jacob Abernethy, University of Pennsylvania

Yiling Chen, Harvard University

Jennifer Wortman Vaughan, University of California, Los Angeles

We propose a general framework for the design of securities markets over combinatorial or infinite state or outcome spaces. The framework enables the design of computationally efficient markets tailored to an arbitrary, yet relatively small, space of securities with bounded payoff. We prove that any market satisfying a set of intuitive conditions must price securities via a convex cost function, which is constructed via conjugate duality. Rather than deal with an exponentially large or infinite outcome space directly, our framework only requires optimization over a convex hull. By reducing the problem of automated market making to convex optimization, where many efficient algorithms exist, we arrive at a range of new polynomial-time pricing mechanisms for various problems. We demonstrate the advantages of this framework with the design of some particular markets. We also show that by relaxing the convex hull we can gain computational tractability without compromising the market institution's bounded budget. Although our framework was designed with the goal of deriving efficient automated market makers for markets with very large outcome spaces, this framework also provides new insights into the relationship between market design and machine learning, and into the complete market setting. Using our framework, we illustrate the mathematical parallels between cost function based markets and online learning and establish a correspondence between cost function based markets and market scoring rules for complete markets.

Categories and Subject Descriptors: F.0 [Theory of Computation]: General; J.4 [Computer Applications]: Social and Behavioral Sciences

General Terms: Algorithms, Economics, Theory

Additional Key Words and Phrases: Market design, securities market, prediction market, automated market maker, convex analysis, online linear optimization

ACM Reference Format:

Abernethy, J., Chen, Y., Vaughan, J. W. 2012. Efficient Market Making via Convex Optimization, and a Connection to Online Learning. ACM TEAC 1, 1, Article X (2012), 38 pages.

DOI = 10.1145/0000000.0000000 <http://doi.acm.org/10.1145/0000000.0000000>

Parts of this research initially appeared in Chen and Vaughan [2010] and Abernethy et al. [2011].

This work is supported NSF grants CCF-0953516, CCF-0915016, IIS-1054911, and DMS-070706, DARPA grant FA8750-05-2-0249, and a Yahoo! PhD Fellowship, and is based on work that was supported by NSF under CNS-0937060 to the CRA for the CIFellows Project. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors alone. The authors are grateful to David Pennock for useful discussions about this work and Xiaolong Li and Michael Ruberry for comments on an earlier draft.

Author's addresses: J. Abernethy, Computer and Information Science Department, University of Pennsylvania; Y. Chen, School of Engineering and Applied Sciences, Harvard University; J. W. Vaughan, Computer Science Department, University of California, Los Angeles.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2012 ACM 0000-0000/2012/-ARTX \$15.00

DOI 10.1145/0000000.0000000 <http://doi.acm.org/10.1145/0000000.0000000>

1. INTRODUCTION

Securities markets play a fundamental role in economics and finance. A securities market offers a set of *contingent securities* whose payoffs depend on the future state of the world. For example, an Arrow-Debreu security pays \$1 if a particular state of the world is reached and \$0 otherwise [Arrow 1964; 1970]. Consider an Arrow-Debreu security that will pay off in the event that a category 4 or higher hurricane passes through Florida in 2012. A Florida resident who worries about his home being damaged might buy this security as a form of insurance to hedge his risk; if there is a hurricane powerful enough to damage his home, he will be compensated. Additionally, a risk-neutral trader who has reason to believe that the probability of a category 4 or higher hurricane landing in Florida in 2012 is p should be willing to buy this security at any price below p or (short) sell it at any price above p to capitalize his information. For this reason, the market price of the security can be viewed as the traders' collective estimate of how likely it is that a powerful hurricane will occur. Securities markets thus have dual functions: *risk allocation* and *information aggregation*.

Insurance contracts, options, futures, and many other financial derivatives are examples of contingent securities. A securities market primarily focused on information aggregation is often referred to as a *prediction market*. The forecasts of prediction markets have proved to be accurate in a variety of domains [Ledyard et al. 2009; Berg et al. 2001; Wolfers and Zitzewitz 2004]. While our work builds on ideas from prediction market design [Chen and Vaughan 2010; Othman et al. 2010; Agrawal et al. 2011], our framework can be applied to any contingent securities.

A securities market is said to be *complete* if it offers at least $|\mathcal{O}|$ linearly independent securities over a set \mathcal{O} of mutually exclusive and exhaustive states of the world, which we refer to as *outcomes* [Arrow 1964; 1970; Mas-Colell et al. 1995]. For example, a prediction market with n Arrow-Debreu securities for n outcomes is complete. In a complete securities market without transaction fees, a trader may bet on any combination of the securities, allowing him to hedge any possible risk he may have. It is generally assumed that the trader may *short sell* a security, betting against the given outcome; in a market with short selling, the n th security is not strictly necessary, as a trader can substitute the purchase of this security by short selling all others. Furthermore, traders can change the market prices to reflect any valid probability distribution over the outcome space, allowing them to reveal any belief. Completeness therefore provides expressiveness for both risk allocation and information aggregation.

Unfortunately, completeness is not always achievable. In many real-world settings, the outcome space is exponentially large or even infinite. For instance, a competitive race between n athletes results in an outcome space of $n!$ rank orders, while the future price of a stock has an infinite outcome space, namely $\mathbb{R}_{\geq 0}$. In such situations operating a complete securities market is not practical for two reasons: (a) humans are notoriously bad at estimating small probabilities and (b) it is computationally intractable to manage such a large set of securities. Instead, it is natural to offer a smaller set of structured securities. For example, rather than offer a security corresponding to each rank ordering, in *pair betting* a market institution offers securities of the form “\$1 if candidate A beats candidate B” [Chen et al. 2007a; Chen et al. 2008a]. There has been a surge of recent research examining the tractability of running standard prediction market mechanisms (such as the popular Logarithmic Market Scoring Rule (LMSR) market maker [Hanson 2003]) over combinatorial outcome spaces by limiting the space of available securities [Pennock and Sami 2007]. While this line of research has led to a few positive results [Chen et al. 2007b; Chen et al. 2008b; Guo and Pennock 2009; Agrawal et al. 2008], it has led more often to hardness results [Chen et al. 2007b; Chen

et al. 2008a] or to markets with undesirable properties such as unbounded loss of the market institution [Gao et al. 2009].

In this paper, we propose a general framework to design *automated market makers* for securities markets. An automated market maker is a market institution that adaptively sets prices for each security and is always willing to accept trades at these prices. Unlike previous research aimed at finding a space of securities that can be efficiently priced using an existing market maker like the LMSR, we start with an arbitrary space of securities and design a *new* market maker tailored to this space. Our framework is therefore very general and includes existing market makers for complete markets, such as the LMSR and Quad-SCPM [Agrawal et al. 2011], as special cases.

We take an axiomatic approach. Given a relatively small space of securities with bounded payoff, we define a set of intuitive conditions that a reasonable market maker should satisfy. We prove that a market maker satisfying these conditions must price securities via a convex potential function (the *cost function*), and that the space of reachable security prices must be precisely the convex hull of the payoff vectors for each outcome (that is, the set of vectors, one per outcome, denoting the payoff for each security if that outcome occurs). We then incorporate ideas from online convex optimization [Hazan 2009; Rakhlin 2009] to define a convex cost function in terms of an optimization over this convex hull; the vector of prices is chosen as the optimizer of this convex objective. With this framework, instead of dealing with the exponentially large or infinite outcome space, we only need to deal with the lower-dimensional convex hull. The problem of automated market making is reduced to the problem of convex optimization, for which we have many efficient techniques to leverage.

To demonstrate the advantages of our framework, we provide two new computationally efficient markets. The first market can efficiently price *subset bets* on permutations, which are known to be #P-hard to price using the LMSR [Chen et al. 2008a]. The second market can be used to price bets on the landing location of an object on a sphere. For situations where the convex hull cannot be efficiently represented, we show that we can relax the convex hull to gain computational tractability without compromising the market maker's bounded budget. This allows us to provide a computationally efficient market maker for the aforementioned pair betting, which is also known to be #P-hard to price using the LMSR [Chen et al. 2008a].

Although our framework was designed with the goal of deriving novel, efficient automated market makers for markets with very large outcome spaces, this framework also provides new insights into the relationship between market design and machine learning, and into the complete market setting. With our framework, we illustrate the mathematical parallels between cost function based markets and online learning, and establish a correspondence between cost function based markets and market scoring rules for complete markets.

Roadmap of the paper: The rest of the paper is organized as follows. We begin in Section 2 with a review of the relevant literature on automated market makers and prediction market design. In Section 3 we describe the problem of market design for large outcome spaces, discuss the difficulties inherent to this problem, and introduce our axiomatic approach. In Section 4 we give a detailed framework for constructing pricing mechanisms based on convex optimization and conjugate duality. We give a couple of examples of efficient duality-based cost function market makers in Section 5. In Section 6 we consider the computational issues associated with our framework, and show how the proposed convex optimization problem can be relaxed to gain tractability without increasing the worst-case loss of the market maker. We illustrate the mathematical parallels between our framework and online learning in Section 7. Finally, in

Section 8, we describe how our framework can be used to establish a correspondence between cost function based markets and market scoring rules for complete markets.

2. BACKGROUND AND RELATED WORK

Automated market makers for complete markets are well studied in both economics and finance. Our work builds on the literature on cost function based markets [Hanson 2003; 2007; Chen and Pennock 2007]. A simple cost function based market maker offers $|\mathcal{O}|$ Arrow-Debreu securities, each corresponding to a potential outcome. The market maker determines how much each security should cost using a differentiable *cost function*, $C : \mathbb{R}^{|\mathcal{O}|} \rightarrow \mathbb{R}$, which is simply a potential function specifying the amount of money currently wagered in the market as a function of the number of shares of each security that have been purchased. If q_o is the number of shares of security o currently held by traders, and a trader would like to purchase a bundle of r_o shares for each security $o \in \mathcal{O}$ (where each r_o could be positive, representing a purchase, zero, or even negative, representing a sale), the trader must pay $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})$ to the market maker. The instantaneous price of security o (that is, the price per share of an infinitesimal portion of a security) is then $\partial C(\mathbf{q})/\partial q_o$, and is denoted $p_o(\mathbf{q})$.

One example of a cost function based market that has received considerable attention is Hanson’s Logarithmic Market Scoring Rule (LMSR) [Hanson 2003; 2007; Chen and Pennock 2007]. The cost function of the LMSR is

$$C(\mathbf{q}) = b \log \sum_{o \in \mathcal{O}} e^{q_o/b}, \quad (1)$$

where $b > 0$ is a parameter of the market controlling the rate at which prices change. The corresponding price function for each security o is

$$p_o(\mathbf{q}) = \frac{\partial C(\mathbf{q})}{\partial q_o} = \frac{e^{q_o/b}}{\sum_{o' \in \mathcal{O}} e^{q_{o'}/b}}. \quad (2)$$

It is well known that the monetary loss of an automated market maker using the LMSR is upperbounded by $b \log |\mathcal{O}|$. Additionally, the LMSR satisfies several other desirable properties, which are discussed in more detail in Section 3.1.

When $|\mathcal{O}|$ is large or infinite, calculating the cost of a purchase becomes intractable in general. Recent research on automated market makers for large outcome spaces has focused on restricting the allowable securities over a combinatorial outcome space and examining whether the LMSR prices can be computed efficiently in the restricted space. If the outcome space contains $n!$ rank orders of n competing candidates, it is #P-hard to price *pair bets* (securities of the form “\$1 if and only if candidate A beats candidate B”) or *subset bets* (for example, “\$1 if one of the candidates in subset C finishes at position k ”) using the LMSR on the full set of permutations [Chen et al. 2008a]. If the outcome space contains 2^n Boolean values of n binary base events, it is #P-hard to price securities on conjunctions of any two base events (for example, “\$1 if and only if a Democrat wins Florida and Ohio”) using the LMSR [Chen et al. 2008a]. This line of research has led to some positive results when the uncertain event enforces particular structure on the outcome space. In particular, for a single-elimination tournament of n teams, securities such as “\$1 if and only if team A wins a k th round game” and “\$1 if and only if team A beats team B given they face off” can be priced efficiently using the LMSR [Chen et al. 2008b]. The tractability of these securities is due to a structure-preserving property — the market probability can be represented by a Bayesian network and price updating does not change the structure of the network. Pennock and Xia [2011] significantly generalized this result and characterize all structure-preserving securities. For a taxonomy tree on some statistic where the value

of the statistic of a parent node is the sum of those of its children, securities such as “\$1 if and only if the value of the statistic at node A belongs to $[x, y]$ ” can be priced efficiently using the LMSR [Guo and Pennock 2009].

One approach to combat the computational intractability of pricing over combinatorial spaces is to approximate the market prices using sampling techniques. Yahoo!’s Predictalot,¹ a play-money combinatorial prediction market for the NCAA Men’s Basketball playoff, allows traders to bet on almost any combination of the 2^{63} outcomes of the tournament. Predictalot is based on the LMSR. Instead of calculating the exact prices for securities, it uses importance sampling to approximate the prices. Xia and Pennock [2011] devised a Monte-Carlo algorithm that can efficiently compute the price of any security in disjunctive or conjunctive normal form with guaranteed error bounds. However, using sampling techniques brings a new problem to pricing. The sampling algorithm in general won’t give the same prices if quoted twice, even if the market status remains the same. Because of this, traders can exploit the market to make a profit, which increases the loss of the market maker.

In this paper, we take a drastically different approach to combinatorial market design. Instead of searching for supportable spaces of securities for existing market makers, we design new market makers tailored to any security space of interest and with desirable theoretical properties. Additionally, rather than requiring that securities have a fixed (e.g., \$1) payoff when the underlying event happens, we allow more general contingent securities with arbitrary, efficiently computable and bounded payoffs.

Our approach makes use of powerful techniques from convex optimization. Agrawal et al. [2011] and Peters et al. [2007] also use convex optimization for automated market making. One major difference is that they only consider complete markets, while we consider markets with an arbitrary set of securities. They consider the setting in which traders submit limit orders, and formulate a convex optimization problem that can be solved by the market institution in order to decide what quantity of orders to accept. While formulating the problem in terms of limit orders leads to a syntactically different problem, their mechanisms can be turned into equivalent cost function based market makers. Agrawal et al. [2011] show that their mechanisms can be formulated as a risk minimization problem with an associated penalty function. Mathematically the penalty function plays a similar role as the conjugate function R in our framework, but they do not explicitly make a connection with conjugate duality.

This paper focuses on cost function based market makers. It is worth noting that there are other market mechanisms, with different properties, designed for securities markets. For complete markets, Dynamic Parimutuel Markets [Pennock 2004; Mangold et al. 2005] also use a cost function to price securities, however the securities are parimutuel bets whose future payoff is not fixed a priori, but depends on the market activities. Brahma et al. [2010] and Das and Magdon-Ismail [2008] design Bayesian learning market makers that maintain a belief distribution and update it based on the traders’ behavior. Call markets have been studied to trade securities over combinatorial spaces. In a call market, participants submit limit orders and the market institution determines what orders to accept or reject. Researchers have studied the computational complexity of operating call markets for both permutation [Chen et al. 2007b; Agrawal et al. 2008; Ghodsi et al. 2008] and Boolean [Fortnow et al. 2004] combinatorics.

Related work on online learning and related work on market scoring rules are discussed in Sections 7 and 8 respectively.

¹<http://labs.yahoo.com/project/336>

3. AN AXIOMATIC APPROACH TO MARKET DESIGN

In this work, we are primarily interested in a market-design scenario in which the outcome space \mathcal{O} is exponentially large, or even infinite, making it infeasible to run a complete market; not only is it generally intractable for the market maker to price an exponential number of securities, but it is notoriously difficult for human traders to reason about the probabilities of so many individually unlikely outcomes. To address both of these problems, we restrict the market maker to offer a menu of only K securities for some reasonably-sized K . These securities will be designed by the market maker and one can interpret each security as corresponding to some “interesting” or “useful” query that we might like to make about the future outcome. For example, if a set of players compete in a tournament, the market maker can offer a security for every question of the form “does player X survive beyond round Y ?”

We assume that the payoff of each security, clearly depending on the future outcome ω , can be described by an arbitrary but efficiently-computable function $\rho : \mathcal{O} \rightarrow \mathbb{R}_{\geq 0}^K$; if a trader purchases a share of security i and the true outcome is ω , then the trader is paid $\rho_i(\omega)$. We call such a security space *complex*. The complete security space is a special case of a complex security space in which $K = |\mathcal{O}|$ and for each $i \in \{1, \dots, K\}$, $\rho_i(\omega)$ equals 1 if ω is the i th outcome and 0 otherwise. The markets we design enable traders to purchase arbitrary *security bundles* $r \in \mathbb{R}^K$. A negative element of r encodes a sale of such a security. The payoff for r upon outcome ω is exactly $\rho(\omega) \cdot r$, where $\rho(\omega)$ denotes the vector of payoffs for each security for outcome ω . Let us define $\rho(\mathcal{O}) := \{\rho(\omega) \mid \omega \in \mathcal{O}\}$. *It will be assumed, throughout the paper, that $\rho(\mathcal{O})$ is closed and bounded.*

The first step in the design of automated market makers for complex security spaces is to determine an appropriate set of properties that we would like such market makers to satisfy. To build intuition about which properties might be desirable, we first step back and consider what it is that makes a market maker like the LMSR a good choice for complete markets.

3.1. What Makes A Market Maker Reasonable?

Consider the cost function associated with the Logarithmic Market Scoring Rule (Equation 1) and the corresponding instantaneous price functions (Equation 2). This cost function and the resulting market satisfy several natural properties that make the LMSR a “reasonable” choice:

- (1) The cost function is differentiable everywhere. As a result, an instantaneous price $p_\omega(\mathbf{q}) = \partial C(\mathbf{q}) / \partial q_\omega$ can always be obtained for the security associated with any outcome ω , regardless of the current quantity vector \mathbf{q} .
- (2) The market incorporates information from the traders, in the sense that the purchase of a security corresponding to outcome ω causes p_ω to increase.
- (3) The market does not provide explicit opportunities for arbitrage. Since instantaneous prices are never negative, traders are never paid to obtain securities. Additionally, the sum of the instantaneous prices of the securities is always 1. If the prices summed to something less than (respectively, greater than) 1, a trader could purchase (respectively, short sell) small equal quantities of each security for a guaranteed profit. This is prevented. In addition to preventing arbitrage, these properties also ensure that prices can be interpreted naturally as probabilities, representing the market’s current estimate of the distribution over outcomes.

- (4) The market is *expressive* in the sense that a trader with sufficient funds can always set the market prices to reflect his beliefs about the probability of each outcome.²

As described in Section 2, previous research on cost function based markets for combinatorial outcome spaces has focused on developing algorithms to efficiently implement or approximate LMSR pricing [Chen et al. 2008a; Chen et al. 2008b; Guo and Pennock 2009]. Because of this, there has been no need to explicitly extend these properties to complex markets; the properties hold automatically for any implementation of the LMSR. This is no longer the case when our goal is to design new markets tailored to custom sets of securities.

To gain intuition about what makes an arbitrary complex market “reasonable,” let us begin by considering the example of *pair betting* [Chen et al. 2007a; Chen et al. 2008a]. Suppose our outcome space consists of rankings of a set of n competitors, such as n horses in a race. The outcome of such a race is a permutation $\pi : [n] \rightarrow [n]$, where $[n]$ denotes the set $\{1, \dots, n\}$, and $\pi(i)$ is the final position of i , with $\pi(i) = 1$ being best. A typical market for this setting might offer n securities, with the i th security paying off \$1 $\pi(i) = 1$ and \$0 otherwise. Additionally, there might be separate, *independent* markets allowing bets on horses to place (come in first or second) or show (come in first, second, or third). However, running independent markets for sets of outcomes with clear correlations is wasteful in that information revealed in one market does not automatically propagate to the others. Instead, suppose that we would like to define a set of securities that allow traders to make arbitrary *pair bets*; that is, for every i, j , a trader can purchase a security which pays out \$1 whenever $\pi(i) < \pi(j)$. What properties would make a market for pair bets reasonable?

The first two properties described above have straight-forward interpretations in this setting. We would still like the instantaneous price of each security to be well-defined at all times; intuitively, the instantaneous price of the security for $\pi(i) < \pi(j)$ should represent the traders’ collective belief about the probability that horse i finishes ahead of horse j . Call this price $p_{i,j}$. We would still like the market to incorporate information, in the sense that buying the security corresponding to $\pi(i) < \pi(j)$ should never cause the price $p_{i,j}$ to drop.

The remaining two properties are more tricky to quantify. Intuitively, these properties require us to define a set of constraints over the prices achievable in the market (to prevent arbitrage), and to ensure that any prices reflecting consistent beliefs about the distribution over outcomes can be achieved (for expressiveness). One can come up with various logical constraints that prices should satisfy. For example, $p_{i,j}$ must be nonnegative at all times for all i and j , and $p_{i,j} + p_{j,i}$ must always equal 1 since exactly one of the two securities corresponding to $\pi(i) < \pi(j)$ and $\pi(j) < \pi(i)$ respectively will pay out \$1. Similar reasoning gives us the additional constraint that for all i, j , and k , $p_{i,j} + p_{j,k} + p_{k,i}$ must be at least 1 and no more than 2. But are these constraints enough to prevent arbitrage? Are they too strong to allow the expression of arbitrary consistent beliefs?

In general, this type of ad hoc reasoning can lead us to many apparently reasonable constraints, but does not yield an algorithm to determine whether or not we have generated the full set of constraints necessary to prevent arbitrage, and cannot be applied easily to more complicated security spaces. We address this problem in the next section. We start by formalizing the desirable market properties described above in the context of complex markets. We then provide a precise mathematical characterization of all cost functions that satisfy these properties.

²Othman et al. [2010] introduced a similar property for complete markets, which they called *surjectivity*.

3.2. An Axiomatic Characterization of Complex Markets

We are now ready to formalize a set of conditions or axioms that one might expect a market to satisfy, and show that these conditions lead to some natural mathematical restrictions on the costs of security bundles. (We consider relaxations of these conditions in Section 6.) We do not presuppose a cost function based market. However, we show that the use of a convex cost function is *necessary* given the assumption of path independence on the security purchases.

3.2.1. Path Independence and the Use of Cost Functions. Imagine a sequence of traders entering the marketplace and purchasing security bundles. Let $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots$ be the sequence of security bundles purchased. After $t-1$ such purchases, the t -th trader should be able to enter the marketplace and query the market maker for the cost of arbitrary bundles. The market maker must be able to furnish a cost, denoted $\text{Cost}(\mathbf{r}|\mathbf{r}_1, \dots, \mathbf{r}_{t-1})$, for any bundle \mathbf{r} given a previous trade sequence $\mathbf{r}_1, \dots, \mathbf{r}_{t-1}$. If the trader chooses to purchase \mathbf{r}_t at a cost of $\text{Cost}(\mathbf{r}_t|\mathbf{r}_1, \dots, \mathbf{r}_{t-1})$, the market maker may update the costs of each bundle accordingly. Our first condition requires that the cost of acquiring a bundle \mathbf{r} must be the same regardless of how the trader splits up the purchase.

CONDITION 1 (PATH INDEPENDENCE). *For any \mathbf{r}, \mathbf{r}' , and \mathbf{r}'' such that $\mathbf{r} = \mathbf{r}' + \mathbf{r}''$, for any $\mathbf{r}_1, \dots, \mathbf{r}_t$,*

$$\text{Cost}(\mathbf{r}|\mathbf{r}_1, \dots, \mathbf{r}_t) = \text{Cost}(\mathbf{r}'|\mathbf{r}_1, \dots, \mathbf{r}_t) + \text{Cost}(\mathbf{r}''|\mathbf{r}_1, \dots, \mathbf{r}_t, \mathbf{r}').$$

Path independence helps to reduce both arbitrage opportunities and the strategic play of traders, as traders need not reason about the optimal path leading to some target position. However, it is worth pointing out that there are interesting markets that do not satisfy this condition, such as the continuous double auction and the market maker for continuous double auctions considered by Brahma et al. [2010] and Das and Magdon-Ismael [2008]. These markets do not fall into our framework and deserve separate treatment.

It turns out that the path independence alone implies that prices can be represented by a cost function C , as illustrated in the following theorem.

THEOREM 3.1. *Under Condition 1, there exists a cost function $C : \mathbb{R}^K \rightarrow \mathbb{R}$ such that we may always write*

$$\text{Cost}(\mathbf{r}_t|\mathbf{r}_1, \dots, \mathbf{r}_{t-1}) = C(\mathbf{r}_1 + \dots + \mathbf{r}_{t-1} + \mathbf{r}_t) - C(\mathbf{r}_1 + \dots + \mathbf{r}_{t-1}).$$

PROOF. Let $C(\mathbf{q}) := \text{Cost}(\mathbf{q}|\emptyset)$. Clearly $C(\mathbf{0}) = \text{Cost}(\mathbf{0}|\emptyset) = 0$. We will show, via induction on t , that for any t and any bundle sequence $\mathbf{r}_1, \dots, \mathbf{r}_t$,

$$\text{Cost}(\mathbf{r}_t|\mathbf{r}_1, \dots, \mathbf{r}_{t-1}) = C(\mathbf{r}_1 + \dots + \mathbf{r}_{t-1} + \mathbf{r}_t) - C(\mathbf{r}_1 + \dots + \mathbf{r}_{t-1}). \quad (3)$$

When $t = 1$, this holds trivially. Assume that Equation 3 holds for all bundle sequences of any length $t \leq T$. By Condition 1,

$$\begin{aligned} & \text{Cost}(\mathbf{r}_{T+1}|\mathbf{r}_1, \dots, \mathbf{r}_T) \\ &= \text{Cost}(\mathbf{r}_{T+1} + \mathbf{r}_T|\mathbf{r}_1, \dots, \mathbf{r}_{T-1}) - \text{Cost}(\mathbf{r}_T|\mathbf{r}_1, \dots, \mathbf{r}_{T-1}) \\ &= C\left(\mathbf{r}_{T+1} + \mathbf{r}_T + \sum_{t=1}^{T-1} \mathbf{r}_t\right) - C\left(\sum_{t=1}^{T-1} \mathbf{r}_t\right) - \left(C\left(\mathbf{r}_T + \sum_{t=1}^{T-1} \mathbf{r}_t\right) - C\left(\sum_{t=1}^{T-1} \mathbf{r}_t\right)\right) \\ &= C\left(\sum_{t=1}^{T+1} \mathbf{r}_t\right) - C\left(\sum_{t=1}^T \mathbf{r}_t\right), \end{aligned}$$

and we see that Equation 3 holds for $t = T + 1$ too. \square

With this theorem in mind, we drop the cumbersome $\text{Cost}(\mathbf{r}|\mathbf{r}_1, \dots, \mathbf{r}_t)$ notation from now on, and write the cost of a bundle \mathbf{r} as $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})$, where $\mathbf{q} = \mathbf{r}_1 + \dots + \mathbf{r}_t$ is the vector of previous purchases.

3.2.2. Formalizing the Properties of a Reasonable Market. Recall that one of the functions of a securities market is to aggregate traders' beliefs into an accurate prediction. Each trader may have his own (potentially secret) information about the future, which we represent as a distribution $\mathbf{p} \in \Delta_{|\mathcal{O}|}$ over the outcome space, where $\Delta_n = \{\mathbf{x} \in \mathbb{R}_{\geq 0}^n : \sum_{i=1}^n x_i = 1\}$, the n -simplex. The pricing mechanism should therefore incentivize the traders to reveal \mathbf{p} , but simultaneously avoid providing arbitrage opportunities. Towards this goal, we now revisit the relevant properties of the LMSR discussed in Section 3.2, and show how the ideas behind each of these properties can be extended to the complex market setting, yielding four additional conditions on our pricing mechanism.

The first condition ensures that the gradient of C , $\nabla C(\mathbf{q})$, is always well-defined. If we imagine that a trader can buy or sell an arbitrarily small bundle, we would like the cost of buying and selling an infinitesimal quantity of any particular bundle to be the same. If $\nabla C(\mathbf{q})$ is well-defined, it can be interpreted as a vector of instantaneous prices for each security, with $\partial C(\mathbf{q})/\partial q_i$ representing the price per share of an infinitesimal amount of security i . Additionally, we can interpret $\nabla C(\mathbf{q})$ as the traders' current estimates of the expected payoff of each security, in the same way that $\partial C(\mathbf{q})/\partial q_o$ was interpreted as the probability of outcome o when considering the complete security space.

CONDITION 2 (EXISTENCE OF INSTANTANEOUS PRICES). *C is continuous and differentiable everywhere on \mathbb{R}^K .*

The next condition encompasses the idea that the market should react to trades in a sensible way in order to incorporate the private information of the traders. In particular, it says that the purchase of a security bundle \mathbf{r} should never cause the market to lower the price of \mathbf{r} . This condition is closely related to incentive compatibility for a myopic trader. It is equivalent to requiring that a trader with a distribution $\mathbf{p} \in \Delta_{|\mathcal{O}|}$ can never find it profitable (in expectation) to buy a bundle \mathbf{r} and at the same time find it profitable to buy the bundle $-\mathbf{r}$. In other words, there can not be more than one way to express one's information.

CONDITION 3 (INFORMATION INCORPORATION). *For any \mathbf{q} and $\mathbf{r} \in \mathbb{R}^K$, $C(\mathbf{q} + 2\mathbf{r}) - C(\mathbf{q} + \mathbf{r}) \geq C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})$.*

The no arbitrage condition states that it is never possible for a trader to purchase a security bundle \mathbf{r} and receive a positive profit regardless of the outcome. Without this property, the market maker would occasionally offer traders a chance to obtain a guaranteed profit, which is clearly suboptimal in terms of the market maker's loss. However, we do consider the relaxation of this property in Section 6.

CONDITION 4 (NO ARBITRAGE). *For all $\mathbf{q}, \mathbf{r} \in \mathbb{R}^K$, there exists an $o \in \mathcal{O}$ such that $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q}) \geq \mathbf{r} \cdot \rho(o)$.*

Finally, the expressiveness condition specifies that any trader can set the market prices to reflect his beliefs, within any ϵ error, about the expected payoffs of each security if arbitrarily small portions of shares may be purchased. The ϵ approximation factor is necessary because the trader's beliefs may only be expressible in the limit;

note that the LMSR does not allow a trader to express the belief that an outcome will occur with probability 1 except in the limit.

CONDITION 5 (EXPRESSIVENESS). *For any $\mathbf{p} \in \Delta_{|\mathcal{O}|}$, we write $\mathbf{x}^{\mathbf{p}} := \mathbb{E}_{\sigma \sim \mathbf{p}}[\rho(\sigma)]$. Then for any $\mathbf{p} \in \Delta_{|\mathcal{O}|}$ and any $\epsilon > 0$ there is some $\mathbf{q} \in \mathbb{R}^K$ for which $\|\nabla C(\mathbf{q}) - \mathbf{x}^{\mathbf{p}}\| < \epsilon$.*

Having formalized our set of conditions, we must now address the question of how to determine whether or not these conditions are satisfied for a particular cost function C . The following theorem precisely characterizes the set of all cost functions that satisfy these conditions. The statement and proof require the use of a few pieces of terminology from convex optimization, which will be our main tool for designing cost functions that satisfy Conditions 2-5; for more on why this is necessary, see the note in Section 4. In particular, the *relative boundary* of a convex set S is its boundary in the ‘‘ambient’’ dimension of S . For example, if we consider the n -dimensional probability simplex $\Delta_n := \{\mathbf{x} \in \mathbb{R}^n : \sum_i x_i = 1, \forall i x_i \geq 0\}$, then the relative boundary of Δ_n is the set $\{\mathbf{x} \in \Delta_n : x_i = 0 \text{ for some } i\}$. We use $\text{relint}(S)$ to refer to the *relative interior* of a convex set S , which is the set S minus all of the points on the relative boundary. The interior of a square in 3-dimensional space is empty, but the relative interior is not. We will use $\text{closure}(S)$ to refer to the closure of S , the smallest closed set containing all of the limit points of S . For any subset S of \mathbb{R}^d , let $\mathcal{H}(S)$ denote the convex hull of S . An important object, which we will use throughout the paper, is $\mathcal{H}(\rho(\mathcal{O}))$ the convex hull of the set of outcome payoffs. (Recall that $\rho(\mathcal{O}) := \{\rho(\sigma) | \sigma \in \mathcal{O}\}$.) As we have assumed that $\rho(\mathcal{O})$ is a closed set, it follows easily that $\mathcal{H}(\rho(\mathcal{O}))$ is also closed, and hence $\text{closure}(\mathcal{H}(\rho(\mathcal{O}))) = \mathcal{H}(\rho(\mathcal{O}))$.

THEOREM 3.2. *Under Conditions 2-5, C must be convex with*

$$\text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}) = \mathcal{H}(\rho(\mathcal{O})). \quad (4)$$

Moreover, any convex differentiable function $C : \mathbb{R}^K \rightarrow \mathbb{R}$ respecting (4) must also satisfy Conditions 2-5.

PROOF. We begin with the first direction. Take any C satisfying Conditions 2-5. We first establish that C is convex everywhere. Assume C is non-convex somewhere. Then there must exist some \mathbf{q} and \mathbf{r} such that $C(\mathbf{q}) > (1/2)C(\mathbf{q} + \mathbf{r}) + (1/2)C(\mathbf{q} - \mathbf{r})$. This means $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q}) < C(\mathbf{q}) - C(\mathbf{q} - \mathbf{r})$, which contradicts Condition 3, so C must be convex.

To prove the equality, we will establish containment in both directions. We first prove that $\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\} \subset \mathcal{H}(\rho(\mathcal{O}))$, from which it follows that $\text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}) \subseteq \mathcal{H}(\rho(\mathcal{O}))$ because $\mathcal{H}(\rho(\mathcal{O}))$ is already closed by assumption. Notice that Condition 2 trivially guarantees that $\nabla C(\mathbf{q})$ is well-defined for any \mathbf{q} . Towards a contradiction, let us assume there exists some \mathbf{q}' for which $\nabla C(\mathbf{q}') \notin \mathcal{H}(\rho(\mathcal{O}))$. Because the hull is a convex set, this can be reformulated in the following way: There must exist some halfspace, defined by a normal vector \mathbf{r} , that separates $\nabla C(\mathbf{q}')$ from every member of $\rho(\mathcal{O})$. More precisely

$$\nabla C(\mathbf{q}') \notin \mathcal{H}(\rho(\mathcal{O})) \iff \exists \mathbf{r} \forall \sigma \in \mathcal{O} : \nabla C(\mathbf{q}') \cdot \mathbf{r} < \rho(\sigma) \cdot \mathbf{r}.$$

The strict inequality in this equation is due to the assumption that $\mathcal{H}(\rho(\mathcal{O}))$ is a closed convex set. On the other hand, letting $\mathbf{q} := \mathbf{q}' - \mathbf{r}$, we see by convexity of C that $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q}) \leq \nabla C(\mathbf{q}') \cdot \mathbf{r}$. Combining these last two inequalities, we see that the price of bundle \mathbf{r} purchased with history \mathbf{q} is always smaller than the payoff for *any* outcome. This implies that there exists some arbitrage opportunity, contradicting Condition 4.

We now show that $\mathcal{H}(\rho(\mathcal{O})) \subseteq \text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\})$. The statement of Condition 5 is equivalent to the statement that every element $\mathbf{x}^{\mathbf{p}} \in \mathcal{H}(\rho(\mathcal{O}))$ is a limit point

of the set $\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}$. But then we are done, as the $\text{closure}(S)$ is defined as the S including all of its limit points.

We now prove the final statement, which is that (4) is also sufficient to achieve Conditions 2-5. Take some convex differentiable $C : \mathbb{R}^K \rightarrow \mathbb{R}$ for which (4) is true. Condition 2 follows by definition. As previously argued, Condition 3 is equivalent to the convexity of C . Condition 5 is equivalent to the statement that $\mathcal{H}(\rho(\mathcal{O})) \subseteq \text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\})$. Finally, to establish Condition 4, we have to reverse our previous argument. The existence of an arbitrage opportunity means that there exist some \mathbf{q}, \mathbf{r} such that $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q}) < \rho(\mathfrak{o}) \cdot \mathbf{r}$ for each $\mathfrak{o} \in \mathcal{O}$. Using convexity of C , we also have that $\nabla C(\mathbf{q}) \cdot \mathbf{r} \leq C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})$. Combining gives us that $\nabla C(\mathbf{q}) \cdot \mathbf{r} \leq \rho(\mathfrak{o}) \cdot \mathbf{r}$ for all $\mathfrak{o} \in \mathcal{O}$, but this last statement is equivalent to the statement that $\nabla C(\mathbf{q}) \notin \mathcal{H}(\rho(\mathcal{O}))$. This is a contradiction and thus Condition 4 is satisfied. \square

What we have arrived at from the set of proposed conditions is that (a) a pricing mechanism can always be described precisely in terms of a convex cost function C and (b) the set of reachable prices of a mechanism, that is the set $\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}$, must be *identically* the convex hull of the payoff vectors for each outcome $\mathcal{H}(\rho(\mathcal{O}))$ except possibly differing at the relative boundary of $\mathcal{H}(\rho(\mathcal{O}))$. For complete markets, this would imply that the set of achievable prices should be the convex hull of the n standard basis vectors. Indeed, this comports exactly with the natural assumption that the vector of security prices in complete markets should represent a probability distribution, or equivalently that it should lie in the n -simplex [Agrawal et al. 2011].

4. DESIGNING THE COST FUNCTION VIA CONJUGATE DUALITY

The natural conditions we introduced above imply that to design a market for a set of K securities with payoffs specified by an arbitrary payoff function $\rho : \mathcal{O} \rightarrow \mathbb{R}_{\geq 0}^K$, we should use a cost function based market with a convex, differentiable cost function such that $\text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}) = \mathcal{H}(\rho(\mathcal{O}))$. We now provide a general technique that can be used to design and compare properties of cost functions that satisfy these criteria. Our proposed framework uses the notion *conjugate duality* to construct cost functions. The aim here is to simplify the task of designing a function C which satisfies Conditions 2-5. We refer to any market mechanism belonging to our framework as a *Duality-based Cost Function Market Maker*.

Duality-based Cost Function Market Maker

Input: Outcome space \mathcal{O}

Input: K securities specified by a payoff function $\rho : \mathcal{O} \rightarrow \mathbb{R}_{\geq 0}^K$

Input: Convex compact price space Π (typically $\Pi \equiv \mathcal{H}(\rho(\mathcal{O}))$)

Input: Strictly convex R with $\text{reint}(\Pi) \subseteq \text{dom}(R)$

Output: Market mechanism specified by the cost function $C : \mathbb{R}^K \rightarrow \mathbb{R}$ with

$$C(\mathbf{q}) := \sup_{\mathbf{x} \in \text{reint}(\Pi)} \mathbf{x} \cdot \mathbf{q} - R(\mathbf{x})$$

To understand this framework, we begin by reviewing the definition of a convex conjugate. Here and throughout the paper we use the notation $\text{dom}(f)$ to refer to the domain of a function f , i.e., where it is defined and finite valued.

Definition 4.1 (Rockafellar [1970], Section 12). For any convex function $f : \mathbb{R}^K \rightarrow [-\infty, \infty]$, the *convex conjugate* f^* of f is defined as

$$f^*(\mathbf{z}) := \sup_{\mathbf{x} \in \mathbb{R}^K} \mathbf{z} \cdot \mathbf{x} - f(\mathbf{x}).$$

The curious reader can find good discussions of conjugate functions in, e.g., Boyd and Vandenberghe [2004] or Hiriart-Urruty and Lemaréchal [2001]. Rockafellar [1970] further shows that if f is convex and *proper*³ then f^* is also convex and proper. Properness shall be assumed throughout; that is, when we introduce a function and refer to it as *convex* we mean *convex and proper*.

The notion of convex duality has several nice features. For example, under weak conditions it holds that $f^{**} \equiv f$ for a convex f . We need more tools from convex analysis to give precise proofs of the results needed for the present discussion, however we save the technical details for the appendix. We now state the key result that justifies the duality-based framework. The proof of this theorem can also be found in the appendix.

THEOREM 4.2. *Assume we have an outcome space \mathcal{O} and a payoff function ρ such that $\rho(\mathcal{O})$ is a bounded subset of \mathbb{R}^K . Then for any cost function $C : \mathbb{R}^K \rightarrow \mathbb{R}$ satisfying Conditions 2-5 and where C is closed⁴, there exists a strictly convex function $R : \mathbb{R}^K \rightarrow [-\infty, \infty]$ such that*

$$C(\mathbf{q}) = \sup_{\mathbf{x} \in \text{relint}(\mathcal{H}(\rho(\mathcal{O})))} \mathbf{x} \cdot \mathbf{q} - R(\mathbf{x}). \quad (5)$$

Furthermore, for any convex function R defined on $\text{relint}(\mathcal{H}(\rho(\mathcal{O})))$, if R is strictly convex on its domain then the cost function defined by the conjugate, $C := R^$, satisfies Conditions 2-5.*

This theorem is the key result that will guide us in designing a market pricing mechanism. This mechanism relies on constructing a cost function $C : \mathbb{R}^K \rightarrow \mathbb{R}$ that satisfies Conditions 2-5, and we are now given ingredients to achieve this: pick *any* strictly convex function R with domain containing $\mathcal{H}(\rho(\mathcal{O}))$, and let C be defined as in (5). Moreover, any C satisfying the desired conditions can be constructed in this fashion.

4.1. Properties of Duality-based Cost Functions

We now devote a few paragraphs to some important details regarding the proposed duality-based pricing mechanism.

In our definition, we introduce the concept of a “price space” denoted by Π . For the conditions of Theorem 4.2 to hold, we need $\Pi \equiv \mathcal{H}(\rho(\mathcal{O}))$. One might ask why we even introduce a price space Π when it is already given by ρ . Indeed, we give the more general definition because, as we will discuss, there can be computational benefits to allowing Π to be larger. We also require that R be differentiable which, while not strictly necessary, is a reasonable condition and eases the notation as we can now discuss the gradient $\nabla R(\mathbf{x})$.

This duality based approach to designing the market mechanism is convenient for several reasons. First, it leads to markets that are efficient to implement whenever $\mathcal{H}(\rho(\mathcal{O}))$ can be described by a polynomial number of simple constraints.⁵ The difficulty with combinatorial outcome spaces is that actually enumerating the set of out-

³The properness of a function is defined in appendix. This is not to be confused with the properness of a scoring rule that we will discuss in Section 8.

⁴See the appendix for the definition of closed convex functions.

⁵Under reasonable assumptions, a convex program can be solved with error ϵ in time polynomial in $1/\epsilon$ and the size of the problem input using standard techniques, e.g., the ellipsoid method and interior point

comes can be challenging or impossible. In our proposed framework we need only work with the convex hull of the payoff vectors for each outcome when represented by a low-dimensional payoff function $\rho(\cdot)$. This has significant benefits, as one often encounters convex sets which contain exponentially many vertices yet can be described by polynomially many constraints. Moreover, as the construction of C is based entirely on convex programming, we reduce the problem of automated market making to the problem of optimization for which we have a wealth of efficient algorithms. Second, this method yields simple formulas for properties of markets that help us choose the best market to run. Two of these properties, worst-case monetary loss and worst-case information loss, are analyzed in Section 4.2.

In order to establish precise statements, our discussions about certain convex sets – e.g., $\{\nabla C\}$, $\mathcal{H}(\rho(\mathcal{O}))$, and Π – have required precise definitions like the relative boundary and interior, and the closure of a set. One might ask whether this is necessary, as we might be focusing too heavily on “boundary cases.” While these details are occasionally cumbersome, they are important and do arise for very simple markets. For example, for the case of a complete market on n outcomes using the LMSR cost function $C(\mathbf{q}) = b \log \sum_i \exp(q_i/b)$, we have that $\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^n\} = \text{relint}(\Delta_n)$; prices of 0 and 1 can be reached only in the limit.

For the remainder of the paper, we shall further assume that our chosen R is continuous and defined everywhere on $\mathcal{H}(\rho(\mathcal{O}))$; that is, not just on the relative interior. It is not entirely unreasonable to consider functions R for which this is not the case, for example we could imagine an R which asymptotes towards the boundary of $\mathcal{H}(\rho(\mathcal{O}))$. However, there are practical reasons why this is undesirable as we will show such cases lead to unbounded loss for the market maker. Notice also that if R is defined on the compact set $\mathcal{H}(\rho(\mathcal{O}))$ it follows immediately that R is also bounded on $\mathcal{H}(\rho(\mathcal{O}))$. Furthermore, we can always write

$$C(\mathbf{q}) = \max_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} \mathbf{x} \cdot \mathbf{q} - R(\mathbf{x}); \quad (6)$$

that is, where we have replaced \sup with \max . Equation 6 is often convenient as we need to consider the maximizer of the optimization.

LEMMA 4.3. *If R is continuous and defined on all of $\mathcal{H}(\rho(\mathcal{O}))$, the price vector at any $\mathbf{q} \in \mathbb{R}^K$ satisfies*

$$\nabla C(\mathbf{q}) = \arg \max_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} \mathbf{q} \cdot \mathbf{x} - R(\mathbf{x}). \quad (7)$$

PROOF. We first note that the optimization problem in Equation 7 has a unique maximizer because R is strictly convex. We know via conjugate duality that for any $\mathbf{q} \in \mathbb{R}^K$,

$$R(\nabla C(\mathbf{q})) = \sup_{\mathbf{q}' \in \mathbb{R}^K} \mathbf{q}' \cdot \nabla C(\mathbf{q}) - C(\mathbf{q}').$$

Since the supremum is over all of \mathbb{R}^K , it is achieved anywhere the derivative of the objective function (with respect to \mathbf{q}') vanishes. This holds when $\mathbf{q}' = \mathbf{q}$, which gives us that

$$R(\nabla C(\mathbf{q})) + C(\mathbf{q}) = \mathbf{q} \cdot \nabla C(\mathbf{q}), \quad (8)$$

for every \mathbf{q} . Equation 7 follows immediately from Equation 8. \square

methods. Efficient techniques for convex optimization have been thoroughly studied and can be found in standard texts; hence we omit such discussions in the present work.

Lemma 4.3 shows that given any \mathbf{q} , the instantaneous prices are simply the maximizer of the convex optimization problem (6) for any R that is bounded and defined on $\mathcal{H}(\rho(\mathcal{O}))$. This convenient fact will be used throughout the paper.

Given an arbitrary smooth convex function f , we can define the *Legendre Transformation* which maps a point $\mathbf{x} \in \text{dom}(f)$ via the rule $\mathbf{x} \mapsto \nabla f(\mathbf{x})$. Indeed, under certain circumstances we get that this map is the *inverse* of the Legendre transformation of the conjugate f^* , i.e., $\nabla f^*(\nabla f(\mathbf{x})) = \mathbf{x}$ and $\nabla f(\nabla f^*(\mathbf{y})) = \mathbf{y}$ for every $\mathbf{x} \in \text{dom}(f)$ and $\mathbf{y} \in \text{dom}(f^*)$. Unfortunately the required conditions are quite strong: we need that f is strictly convex, the interior of $\text{dom}(f)$ is non-empty, and ∇f always diverges towards the boundary of $\text{dom}(f)$ (see Chapter 26 of Rockafellar [1970]). So while we would like to argue that ∇C is the inverse of the map ∇R for our framework, this will generally not be true. Assuming R is differentiable then given any $\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))$, according to Lemma 4.3 we always have $\nabla C(\nabla R(\mathbf{x})) = \mathbf{x}$ by setting $\mathbf{q} = \nabla R(\mathbf{x})$. However, $\nabla R(\nabla C(\mathbf{q})) = \mathbf{q}$ does not hold in general. On the other hand, if $\mathcal{H}(\rho(\mathcal{O}))$ has a non-empty interior, and the optimal solution to Equation 6 is always contained within the interior, then the statement $\nabla R(\nabla C(\mathbf{q})) = \mathbf{q}$ will hold. Note, however, that these conditions are not satisfied for a complete market on n outcomes, where $\mathcal{H}(\rho(\mathcal{O}))$ is the n -simplex Δ_n which has an empty interior (even though the relative interior is non-empty). Thus, cost function based market makers for complete markets do not satisfy $\nabla R(\nabla C(\mathbf{q})) = \mathbf{q}$. In fact, while each \mathbf{q} maps to a single price $\mathbf{x} = \nabla C(\mathbf{q})$, each price \mathbf{x} can be achieved at multiple values of \mathbf{q} in these markets.

4.2. Bounding the Market Maker's Loss and Loss of Information

We now discuss two key properties of our proposed market framework. We will make use of the notion of a Bregman divergence. The *Bregman divergence* with respect to a differentiable convex function f is given by

$$D_f(\mathbf{x}, \mathbf{y}) := f(\mathbf{x}) - f(\mathbf{y}) - \nabla f(\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}).$$

It is clear by convexity that $D_f(\mathbf{x}, \mathbf{y}) \geq 0$ for all \mathbf{x} and \mathbf{y} .

4.2.1. Bounding the Market Maker's Monetary Loss. When comparing market mechanisms, it is useful to consider the market maker's worst-case monetary loss,

$$\sup_{\mathbf{q} \in \mathbb{R}^K} \left(\sup_{\mathfrak{o} \in \mathcal{O}} (\rho(\mathfrak{o}) \cdot \mathbf{q}) - C(\mathbf{q}) + C(\mathbf{0}) \right).$$

This quantity is simply the worst-case difference between the maximum amount that the market maker might have to pay the traders ($\sup_{\mathfrak{o} \in \mathcal{O}} \rho(\mathfrak{o}) \cdot \mathbf{q}$) and the amount of money collected by the market maker ($C(\mathbf{q}) - C(\mathbf{0})$). The following theorem provides a bound on this loss in terms of the conjugate function.

THEOREM 4.4. *Consider any duality-based cost function market maker with $\Pi = \mathcal{H}(\rho(\mathcal{O}))$. The worst-case monetary loss of the market maker is no more than*

$$\sup_{\mathbf{x} \in \rho(\mathcal{O})} R(\mathbf{x}) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}). \quad (9)$$

Furthermore, the above bound is tight, as the supremum of the market maker's loss is exactly the value in Equation 9.

PROOF. Let \mathbf{q} denote the final vector of quantities sold, $\nabla C(\mathbf{q})$ denote the final vector of instantaneous prices, and \mathfrak{o} denote the true outcome. From Equations 6 and 7, we have that $C(\mathbf{q}) = \nabla C(\mathbf{q}) \cdot \mathbf{q} - R(\nabla C(\mathbf{q}))$ and $C(\mathbf{0}) = -\min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x})$. The difference between the amount that the market maker must pay out and the amount that

the market maker has previously collected when outcome \circ happens is

$$\begin{aligned}
& \rho(\circ) \cdot \mathbf{q} - C(\mathbf{q}) + C(\mathbf{0}) \\
&= \rho(\circ) \cdot \mathbf{q} - (\nabla C(\mathbf{q}) \cdot \mathbf{q} - R(\nabla C(\mathbf{q}))) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) \\
&= \mathbf{q} \cdot (\rho(\circ) - \nabla C(\mathbf{q})) + R(\nabla C(\mathbf{q})) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) + R(\rho(\circ)) - R(\rho(\circ)) \\
&= R(\rho(\circ)) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) - (R(\rho(\circ)) - R(\nabla C(\mathbf{q})) - \mathbf{q} \cdot (\rho(\circ) - \nabla C(\mathbf{q}))) \quad (10) \\
&\leq R(\rho(\circ)) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) - (R(\rho(\circ)) - R(\nabla C(\mathbf{q})) - \nabla R(\nabla C(\mathbf{q})) \cdot (\rho(\circ) - \nabla C(\mathbf{q}))) \\
&= R(\rho(\circ)) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) - D_R(\rho(\circ), \nabla C(\mathbf{q})),
\end{aligned}$$

where D_R is the Bregman divergence with respect to R , as defined above. The first equality follows from Equation 8. The inequality follows from the first-order optimality condition for convex optimization, which says that for any convex and differentiable f defined on the domain Π , if f is minimized at \mathbf{x} , then

$$\nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) \geq 0 \quad \text{for any } \mathbf{y} \in \Pi.$$

Consider $f(\mathbf{x}) = R(\mathbf{x}) - \mathbf{q} \cdot \mathbf{x}$. The minimum of this function occurs at $\mathbf{x} = \nabla C(\mathbf{q})$ via the duality assumption. Plugging in $\mathbf{y} = \rho(\circ)$ yields the inequality.

Since the divergence is always nonnegative, this is upperbounded by $R(\rho(\circ)) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x})$, which is in turn upperbounded by $\sup_{\mathbf{x} \in \rho(\mathcal{O})} R(\mathbf{x}) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x})$.

Finally, we show that this loss bound is tight. First, select any $\epsilon > 0$. Choose an outcome \circ so that $\sup_{\circ' \in \mathcal{O}} R(\rho(\circ')) - R(\rho(\circ)) < \epsilon/2$. Next, choose some \mathbf{q}' so that $D_R(\rho(\circ), \nabla C(\mathbf{q}')) < \epsilon/2$. This is achievable because the space of gradients of C is assumed to span $\text{relint}(\mathcal{H}(\rho(\mathcal{O})))$ via Theorem 3.2, and so we can ensure that $\nabla C(\mathbf{q}')$ is arbitrarily close to $\rho(\circ)$. Finally, let $\mathbf{q} := \nabla R(\nabla C(\mathbf{q}'))$, and observe that by construction we have $\nabla C(\mathbf{q}) = \nabla C(\mathbf{q}')$. To compute the market maker's loss for this particular choice of \mathbf{q} and \circ , we apply Equation 10 to obtain:

$$\begin{aligned}
& R(\rho(\circ)) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) - (R(\rho(\circ)) - R(\nabla C(\mathbf{q})) - \mathbf{q} \cdot (\rho(\circ) - \nabla C(\mathbf{q}))) \\
&= R(\rho(\circ)) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) - D_R(\rho(\circ), \nabla C(\mathbf{q})) \\
&> \sup_{\circ' \in \mathcal{O}} R(\rho(\circ')) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) - \epsilon
\end{aligned}$$

where the first equality holds by the definition of the Bregman divergence, because $\mathbf{q} = \nabla R(\nabla C(\mathbf{q}))$. \square

This theorem tells us that as long as the conjugate function is bounded on $\mathcal{H}(\rho(\mathcal{O}))$, the market maker's worst-case loss is also bounded.⁶ It says further that this loss is actually realized, for a particular outcome \circ , at least when the price vector approaches $\rho(\circ)$. This suggests that loss to the market maker is worst when the traders are the most certain about the outcome.

4.2.2. Bounding Information Loss. Information loss can occur when securities are sold in discrete quantities (for example, single units), as they are in most real-world markets.

⁶In Section 6, we will state a more general, stronger bound on market maker's loss capturing the intuitive notion that the market maker's profits should be higher when the distance between the final vector of prices and the payoff vector $\rho(\circ)$ of the true outcome \circ is large; see Theorem 6.2.

Without the ability to purchase arbitrarily small bundles, traders may not be able to change the market prices to reflect their true beliefs about the expected payoff of each security, even if expressiveness is satisfied. We will argue that the amount of information loss is captured by the market's bid-ask spread for the smallest trading unit. Given some \mathbf{q} , the current bid-ask spread of security bundle \mathbf{r} is defined to be $(C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})) - (C(\mathbf{q}) - C(\mathbf{q} - \mathbf{r}))$. This is simply the difference between the current cost of buying the bundle \mathbf{r} and the current price at which \mathbf{r} could be sold.

To see how the bid-ask spread relates to information loss, suppose that the current vector of quantities sold is \mathbf{q} . If securities must be sold in unit chunks, a rational, risk-neutral trader will not buy security i unless she believes the expected payoff of this security is at least $C(\mathbf{q} + \mathbf{e}^i) - C(\mathbf{q})$, where \mathbf{e}^i is the vector that has value 1 at its i th element and 0 everywhere else. Similarly, she will not sell security i unless she believes the expected payoff is at most $C(\mathbf{q}) - C(\mathbf{q} - \mathbf{e}^i)$. If her estimate of the expected payoff of the security is between these two values, she has no incentive to buy or sell the security. In this case, it is only possible to infer that the trader believes the true expected payoff lies somewhere in the range $[C(\mathbf{q}) - C(\mathbf{q} - \mathbf{e}^i), C(\mathbf{q} + \mathbf{e}^i) - C(\mathbf{q})]$. The bid-ask spread is precisely the size of this range.

The bid-ask spread depends on how fast instantaneous prices change as securities are bought or sold. Intuitively, the bid-ask spread relates to the depth of the market. When the bid-ask spread is large, new purchases or sales can change the prices of the securities dramatically; essentially, the market is shallow. When the bid-ask spread is small, purchases or sales may only move the prices slightly; the market is deep. Based on this intuition, for complete markets, Chen and Pennock [2007] use the inverse of $\partial^2 C(\mathbf{q}) / \partial q_i^2$ to capture the notion of market depth for each security i independently. In a similar spirit, we define a *market depth parameter* β . Larger values of β correspond to deeper markets. We will bound the bid-ask spread in terms of this parameter, and use this parameter to show that there exists a clear tradeoff between worst-case monetary loss and information loss; this will be formalized in Theorem 4.7 below.

To simplify discussion, assume that C is twice-differentiable. Our parameter β is related to the curvature of C . Given any unit vector \mathbf{v} , the curvature (i.e., second derivative) of C at \mathbf{q} in the direction of \mathbf{v} can be calculated as $\mathbf{v}^\top \nabla^2 C(\mathbf{q}) \mathbf{v}$, where $\nabla^2 C(\mathbf{q})$ is the Hessian of C at \mathbf{q} . Furthermore, for any unit vector \mathbf{v} , $\mathbf{v}^\top \nabla^2 C(\mathbf{q}) \mathbf{v}$ is lower bounded by the smallest eigenvalue and upper bounded by the largest eigenvalue of $\nabla^2 C(\mathbf{q})$. To see this, note that the Hessian is a symmetric matrix, and therefore has K linearly independent eigenvectors, each normalized to have length one. Let \mathbf{u}_i be the i th unit eigenvector of $\nabla^2 C(\mathbf{q})$ corresponding to eigenvalue λ_i . λ_i is nonnegative due to convexity of C . Any unit vector \mathbf{v} can be represented as a linear combination of the K unit eigenvectors, $\mathbf{v} = \sum_i a_i \mathbf{u}_i$ with $\sum_i a_i^2 = 1$. For any orthogonal eigenvectors \mathbf{u}_i and \mathbf{u}_j , it is easy to see that $\mathbf{u}_i^\top \nabla^2 C(\mathbf{q}) \mathbf{u}_i = \lambda_i$ and $\mathbf{u}_i^\top \nabla^2 C(\mathbf{q}) \mathbf{u}_j = 0$. Thus, $\mathbf{v}^\top \nabla^2 C(\mathbf{q}) \mathbf{v} = \sum_i a_i^2 \lambda_i$, which lies in $[\min_i \lambda_i, \max_i \lambda_i]$.

Definition 4.5. For any duality-based cost function market maker with twice-differentiable cost function C , the *market depth parameter* $\beta(\mathbf{q})$ for a quantity vector \mathbf{q} is defined as $\beta(\mathbf{q}) = 1/\lambda_C(\mathbf{q})$, where $\lambda_C(\mathbf{q})$ is the largest eigenvalue of $\nabla^2 C(\mathbf{q})$, the Hessian of C at \mathbf{q} . The worst-case market depth is $\beta = \inf_{\mathbf{q} \in \mathbb{R}^K} \beta(\mathbf{q})$.

As described above, this definition of worst-case market depth implies that $1/\beta$ is an upper bound on the curvature of C . We will derive the upper bound of the bid-ask spread and the lower bound of the worst-case loss of the market maker in terms of β . Our derivation makes use of the following lemma that establishes a convenient relationship between the Bregman divergence of a convex function f and the eigenvalues of the Hessian of f . The proof of the lemma is in Appendix B.

LEMMA 4.6. *Let $f(\mathbf{x})$ be a twice-differentiable convex function. If for all $\mathbf{x} \in \text{dom}(f)$, every eigenvalue of $\nabla^2 f(\mathbf{x})$ falls in the set $[a, b]$, $a \leq b$, then for any $\mathbf{x}, \mathbf{x}' \in \text{dom}(f)$,*

$$\frac{a\|\mathbf{x} - \mathbf{x}'\|^2}{2} \leq D_f(\mathbf{x}, \mathbf{x}') \leq \frac{b\|\mathbf{x} - \mathbf{x}'\|^2}{2}. \quad (11)$$

We now present a theorem showing an inherent tension between worst-case monetary loss and information loss. Here $\text{diam}(\mathcal{H}(\rho(\mathcal{O})))$ denotes the diameter of the hull of the payoff vectors for each outcome.

THEOREM 4.7. *For any duality-based cost function market maker with twice differentiable C and worst-case market depth β , the bid-ask spread for bundle \mathbf{r} with previous purchases \mathbf{q} is no more than $\|\mathbf{r}\|^2/\beta$. The worst-case monetary loss of the market maker is at least $\beta \cdot \text{diam}^2(\mathcal{H}(\rho(\mathcal{O}))) / 8$.*

PROOF. The bid-ask spread can be written in terms of Bregman divergences. In particular, $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q}) - (C(\mathbf{q}) - C(\mathbf{q} - \mathbf{r})) = D_C(\mathbf{q} + \mathbf{r}, \mathbf{q}) + D_C(\mathbf{q} - \mathbf{r}, \mathbf{q})$. According to Lemma 4.6, because $1/\beta$ is the upper bound of the eigenvalues of $\nabla^2 C(\mathbf{q})$ at any \mathbf{q} , both $D_C(\mathbf{q} + \mathbf{r}, \mathbf{q})$ and $D_C(\mathbf{q} - \mathbf{r}, \mathbf{q})$ are upper bounded by $\|\mathbf{r}\|^2/2\beta$. Thus, $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q}) - (C(\mathbf{q}) - C(\mathbf{q} - \mathbf{r})) \leq \|\mathbf{r}\|^2/\beta$.

Let $\mathbf{x}_0 = \arg \min_{\mathbf{x} \in \Pi} R(\mathbf{x})$. The first-order optimality condition for convex optimization gives that $\nabla R(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \geq 0$ for all $\mathbf{x} \in \Pi$. According to Theorem 4.4, the worst-case loss of the market maker is

$$\begin{aligned} \sup_{\mathbf{x} \in \rho(\mathcal{O})} R(\mathbf{x}) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) &= \sup_{\mathbf{x} \in \rho(\mathcal{O})} (R(\mathbf{x}) - R(\mathbf{x}_0)) \\ &= \sup_{\mathbf{x} \in \rho(\mathcal{O})} (D_R(\mathbf{x}, \mathbf{x}_0) + \nabla R(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)) \\ &\geq \sup_{\mathbf{x} \in \rho(\mathcal{O})} D_R(\mathbf{x}, \mathbf{x}_0). \end{aligned}$$

Because C is twice-differentiable, for any \mathbf{q} such that $\nabla C(\mathbf{q}) \in \text{relint}(\Pi)$, we have a correspondence between the Hessian of C at \mathbf{q} and the Hessian of R at $\nabla C(\mathbf{q})$. More precisely, we have that $\mathbf{u}^\top \nabla^2 C(\mathbf{q}) \mathbf{u} = \mathbf{u}^\top \nabla^2 R(\nabla C(\mathbf{q})) \mathbf{u}$ for any $\mathbf{u} = \mathbf{x} - \mathbf{x}'$ with $\mathbf{x}, \mathbf{x}' \in \Pi$, where $\nabla^{-2} R(\nabla C(\mathbf{q}))$ denotes the inverse of the Hessian of R at $\nabla C(\mathbf{q})$. (See, for example, Gorni [1991] for more on the second-order properties of convex functions.) This means that $\beta(\mathbf{q})$ is equivalently defined as the smallest eigenvalue of $\nabla^2 R(\nabla C(\mathbf{q}))|_{\Pi}$; that is, where we consider the second derivative only within the price region Π . Thus, β lower bounds the eigenvalues of $\nabla^2 R(\mathbf{x})$ for all $\mathbf{x} \in \Pi$.

Applying Lemma 4.6, we have $D_R(\mathbf{x}, \mathbf{x}_0) \geq \frac{\beta}{2} \|\mathbf{x} - \mathbf{x}_0\|^2$. In the worst-case, \mathbf{x}_0 is in the center of $\mathcal{H}(\rho(\mathcal{O}))$ and $\|\mathbf{x} - \mathbf{x}_0\|$ is at least $\text{diam}(\mathcal{H}(\rho(\mathcal{O}))) / 2$, which finishes the proof. \square

We can see that there is a direct tradeoff between the upper bound⁷ of the bid-ask spread, which shrinks as β grows, and the lower bound of the worst-case loss of the market maker, which grows linearly in β . This tradeoff is very intuitive. When the market is shallow (small β), small trades have a large impact on market prices, and traders cannot purchase too many shares of the same security without paying a lot. When the market is deep (large β), prices change slowly, allowing the market maker

⁷Strictly speaking, as we are emphasizing the necessary tradeoff between bid-ask spread and worst-case loss, we should have a *lower bound* on the bid-ask spread. On the other hand, if the worst-case market depth parameter is β then there is some \mathbf{q} and \mathbf{r} such that $D_C(\mathbf{q} + \mathbf{r}, \mathbf{q}) / \|\mathbf{r}\|^2 \approx 1/(2\beta)$ and this approximation can be made arbitrarily tight for small enough \mathbf{r} when C is twice differentiable.

to gain more precise information, but simultaneously forcing the market maker to take on more risk since many shares of a security can be purchased at prices that are potentially too low. This tradeoff can be adjusted by scaling R , which scales β . This is analogous to adjusting the “liquidity parameter” b of the LMSR.

4.3. Selecting a Conjugate Function

We have seen that the choice of the conjugate function R impacts market properties such as worst-case loss and information loss. We now explore this choice in more detail.

In many situations, the ideal choice of the conjugate is a function of the form

$$R(\mathbf{x}) := \frac{\lambda}{2} \|\mathbf{x} - \mathbf{x}_0\|^2. \quad (12)$$

Here $R(\mathbf{x})$ is simply the squared Euclidean distance between \mathbf{x} and an initial price vector $\mathbf{x}_0 \in \Pi$, scaled by $\lambda/2$. By utilizing this quadratic conjugate function, we achieve a market depth $\beta(\mathbf{q})$ that is uniformly λ for any \mathbf{q} for which $\nabla C(\mathbf{q}) \in \text{relint}(\Pi)$. Furthermore, if \mathbf{x}_0 is chosen as the “center” of Π , namely $\mathbf{x}_0 = \arg \min_{\mathbf{x} \in \Pi} \max_{\mathbf{y} \in \Pi} \|\mathbf{x} - \mathbf{y}\|$, then the worst-case loss of the market maker is $\max_{\mathbf{x} \in \Pi} R(\mathbf{x}) = (\lambda/8) \text{diam}^2(\Pi)$. While the market maker can tune λ appropriately according to the desired tradeoff between worst-case market depth and worst-case loss, the tradeoff is tightest when R has a Hessian that is uniformly a scaled identity matrix, or more precisely where R takes the form in Equation 12.

Unfortunately, by selecting a conjugate of this form, or any R with bounded derivative, the market maker does inherit one potentially undesirable property: security prices may become constant when $\nabla C(\mathbf{q})$ reaches a point at $\text{relbnd}(\Pi)$, the relative boundary of Π . That is, if we arrive at a total demand \mathbf{q} where $\nabla C(\mathbf{q}) = \rho(\sigma)$ for some outcome σ , our mechanism begins offering securities at a price *equal to the best-case payoff*, akin to asking someone to bet a dollar for the chance to possibly win a dollar. The Quad-SCPM for complete markets is known to exhibit this behavior [Agrawal et al. 2011].

To avoid these undesirable pricing scenarios, it is sufficient to require that our conjugate function satisfies one condition. We say that a convex function R defined on Π is a *pseudo-barrier*⁸ for Π if $\|\nabla R(\mathbf{x}_t)\| \rightarrow \infty$ for any sequence of points $\mathbf{x}_1, \mathbf{x}_2, \dots \in \Pi$ which tends towards $\text{relbnd}(\Pi)$. If we require our conjugate function R to be a pseudo-barrier, we are guaranteed that the instantaneous price vector $\nabla C(\mathbf{q})$ always lies in $\text{relint}(\Pi)$, and does not become constant near the boundary.

It is important to note that, while it is desirable that $\|\nabla R(\mathbf{x}_t)\| \rightarrow \infty$ as \mathbf{x}_t approaches $\text{relbnd}(\Pi)$, it is generally not desirable that $R(\mathbf{x}_t) \rightarrow \infty$. Recall that the market maker’s worst-case loss grows with the maximum value of R on Π and thus we restrict a conjugate function that is bounded on the domain. A perfect example of convex function that is simultaneously bounded and a pseudo-barrier is the *negative entropy function* $H(\mathbf{x}) = \sum_i x_i \log x_i$, defined on the n -simplex Δ_n . It is perhaps no surprise that the LMSR, the most common market mechanism for complete security spaces, can be described by the choice $R(\mathbf{x}) := bH(\mathbf{x})$ where the price space $\Pi = \Delta_n$ [Agrawal et al. 2011; Chen and Vaughan 2010].

⁸We use the term pseudo-barrier to distinguish this from the typical definition of a barrier function on a set Π , which is a function that grows without bound towards the boundary of Π . The term *Legendre* was used by Cesa-Bianchi and Lugosi [2006] for a similar notion, which may have originated in Rockafellar [1970], yet this definition requires the stronger condition that Π contains a nonempty interior.

5. EXAMPLES OF COMPUTATIONALLY EFFICIENT MARKETS

In the previous section, we provided a general framework for designing markets on combinatorial or infinite outcome spaces. We now provide some examples of markets that can be operated efficiently using this framework.

5.1. Subset Betting

Recall the scenario described in Section 3.1 in which the outcome is a ranking of a set of n competitors, such as n horses in a race, represented as a permutation $\pi : [n] \rightarrow [n]$. Chen et al. [2007a] proposed a betting language, *subset betting*, in which traders can place bets (i, j) , for any candidate i and any slot j , that pay out \$1 in the event that $\pi(i) = j$ and \$0 otherwise.⁹ Chen et al. [2008a] showed that pricing bets of this form using the LMSR is #P-hard and provided an algorithm for approximating the prices by exploiting the structure of the market. Using our framework, it is simple to design a computationally efficient market for securities of this form.

In order to set up such a combinatorial market within our framework, we must be able to efficiently work with the convex hull of the payoff vectors for each outcome. Notice that, for an outcome π , the associated payoff can be described by a matrix M_π , with $M_\pi(i, j) = I[\pi(i) = j]$, where $I[\cdot]$ is the indicator function. Taking this one step further, it is easy to verify that the convex hull of the set of permutation matrices is precisely the set of *doubly stochastic matrices*, that is the set

$$\Pi = \left\{ X \in \mathbb{R}_{\geq 0}^{n \times n} : \sum_{i'=1}^n X(i', j) = \sum_{j'=1}^n X(i, j') = 1 \quad \forall i, j \right\},$$

where $X(i, j)$ represents the element at the i th row and j th column of the matrix X . Notice, importantly, that this set is described by only n^2 variables and $O(n)$ constraints.

To fully specify the market maker, we must also select a conjugate function R for our price space. While the quadratic conjugate function is an option, there is a natural extension of the negative entropy function, whose desirable properties were discussed in the previous section, for the space of stochastic matrices. For any $X \in \Pi$, let us set

$$R(X) = b \sum_{i,j} X(i, j) \log X(i, j)$$

for some parameter $b > 0$. The worst-case market depth is computed as the minimum of the smallest eigenvalue of the Hessian of R within $\text{relint}(\Pi)$. This occurs when the X matrix has all values $1/n$, hence the worst-case depth is nb . The worst-case loss, on the other hand, is easily computed as $bn \log n$. Note that this bound on worst-case loss is the same that would be obtained by running n independent markets, one for each slot j , using the LMSR.

5.2. Sphere Betting

One important challenge of operating a combinatorial prediction market is to always maintain the logical consistency of security prices. Our framework offers a way to incorporate the constraints on security prices into pricing. Hence, in addition to combinatorial prediction markets, our framework can be used to design markets where security prices have some natural constraints due to their problem domains.

⁹The original definition of subset betting allowed bets of the form “any candidate in set S will end up in slot j ” or “candidate i will end up in one of the slots in set S .” A bet of this form can be constructed easily using our betting language by bundling multiple securities.

We consider an example in which the outcome space is infinite. An object orbiting the planet, perhaps a satellite, is predicted to fall to earth in the near future and will land at an unknown location, which we would like to predict. We represent locations on the earth as unit vectors $\mathbf{u} \in \mathbb{R}^3$. The difficulty of this example arises from the fact that the outcome must be a unit vector, imposing constraints on the three coordinates. We will design a market with three securities, each corresponding to one coordinate of the final location of the object. In particular, security i will pay off $u_i + 1$ dollars if the object lands in location \mathbf{u} . (The addition of 1, while not strictly necessary, ensures that the payoffs, and therefore prices, remain positive, though it will be necessary for traders to sell securities to express certain beliefs.) This means that traders can purchase security bundles $\mathbf{r} \in \mathbb{R}^3$ and, when the object lands at a location \mathbf{u} , receive a payoff $(\mathbf{u} + \mathbf{1}) \cdot \mathbf{r}$. Note that in this example, the outcome space is infinite, but the security space is small.

The price space $\mathcal{H}(\rho(\mathcal{O}))$ for this market will be the 2-norm unit ball centered at $\mathbf{1}$. To construct a market for this scenario, let us make the simple choice of $R(\mathbf{x}) = \lambda \|\mathbf{x} - \mathbf{1}\|^2$ for some parameter $\lambda > 0$. When $\|\mathbf{q}\| \leq 2\lambda$, there exists an \mathbf{x} such that $\nabla R(\mathbf{x}) = \mathbf{q}$. In particular, this is true for $\mathbf{x} = (1/2)\mathbf{q}/\lambda + \mathbf{1}$, and $\mathbf{q} \cdot \mathbf{x} - R(\mathbf{x})$ is maximized at this point. When $\|\mathbf{q}\| > 2\lambda$, $\mathbf{q} \cdot \mathbf{x} - R(\mathbf{x})$ is maximized at an \mathbf{x} on the boundary of $\mathcal{H}(\rho(\mathcal{O}))$. Specifically, it is maximized at $\mathbf{x} = \mathbf{q}/\|\mathbf{q}\| + \mathbf{1}$. From this, we can compute

$$C(\mathbf{q}) = \begin{cases} \frac{1}{4\lambda} \|\mathbf{q}\|^2 + \mathbf{q} \cdot \mathbf{1}, & \text{when } \|\mathbf{q}\| \leq 2\lambda, \\ \|\mathbf{q}\| + \mathbf{q} \cdot \mathbf{1} - \lambda, & \text{when } \|\mathbf{q}\| > 2\lambda. \end{cases}$$

The market depth parameter β is 2λ ; in fact, $\beta(\mathbf{x}) = 2\lambda$ for any price vector \mathbf{x} in the interior of $\mathcal{H}(\rho(\mathcal{O}))$. By Theorem 4.4, the worst-case loss of the market maker is no more than λ , which is precisely the lower bound implied by Theorem 4.7. Finally, the divergence $D_C(\mathbf{q} + \mathbf{r}, \mathbf{q}) \leq \|\mathbf{r}\|^2/(4\lambda)$ for all \mathbf{q}, \mathbf{r} , with equality when $\|\mathbf{q}\|, \|\mathbf{q} + \mathbf{r}\| \leq 2\lambda$, implying that the bid-ask spread scales linearly with $\|\mathbf{r}\|^2/\lambda$.

We note that for this particular prediction problem, if we try to predict the latitude and longitude of the landing location, we don't have any constraints on prices. In particular, we can have two securities that pay off linearly with the latitude and longitude of the landing location respectively. These two securities are independent and can be traded in two independent markets.

6. COMPUTATIONAL COMPLEXITY AND RELAXATIONS

In Section 3, we argued that the space of feasible price vectors should be precisely $\mathcal{H}(\rho(\mathcal{O}))$, the convex hull of the payoff vectors for each outcome. In each of our examples, we have discussed market scenarios for which this hull has a polynomial number of constraints, allowing us to efficiently calculate prices via convex optimization. Unfortunately, one should not necessarily expect that a given payoff function and outcome space will lead to an efficiently describable convex hull. In this section, we explore a couple of approaches to overcome such complexity challenges. First, we discuss the case in which $\mathcal{H}(\rho(\mathcal{O}))$ has exponentially (or infinitely) many constraints yet gives rise to a *separation oracle*. Second, we show that the price space Π can indeed be *relaxed* beyond $\mathcal{H}(\rho(\mathcal{O}))$ without increasing the risk to the market maker. Finally, we show how this relaxation applies in practice.

6.1. Separation Oracles

If we encounter a convex hull $\mathcal{H}(\rho(\mathcal{O}))$ with exponentially-many constraints, all may not be lost. In order to calculate prices, we need to solve the optimization problem $\max_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} \mathbf{q} \cdot \mathbf{x} - R(\mathbf{x})$. Under certain circumstances this can still be solved efficiently.

Consider a convex optimization problem with a concave objective function $f(\mathbf{x})$ and constraints $g_i(\mathbf{x}) \leq 0$ for all i in some index set I . That is, we want to solve:

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{R}^d \\ & g_i(\mathbf{x}) \leq 0 \quad \forall i \in I \end{aligned}$$

This can be converted to a problem with a linear objective in the standard way:

$$\begin{aligned} \max \quad & c \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{R}^d, c \in \mathbb{R} \\ & f(\mathbf{x}) \geq c \\ & g_i(\mathbf{x}) \leq 0 \quad \forall i \in I \end{aligned}$$

Of course, if I is an exponentially or infinitely large set we will have trouble solving this problem directly. On the other hand, the constraint set may admit an efficient *separation oracle*, defined as a function that takes as input a point (\mathbf{x}, c) and returns true if all the necessary constraints are satisfied or, otherwise, returns false and specifies a violated constraint.¹⁰ Given an efficient separation oracle, one has access to alternative methods for optimization, the most famous being Khachiyan's *ellipsoid method*, that run in polynomial time. For more details see, for example, Grötschel et al. [1981].

This suggests that a fruitful direction for designing computationally efficient market makers is to examine the pricing problem on an instance-by-instance basis, and for a particular instance of interest, leverage the structure of the instance to develop an efficient algorithm for solving the specific separation problem. We leave this for future research.

6.2. Relaxation of the Price Space

When dealing with a convex hull $\mathcal{H}(\rho(\mathcal{O}))$ that has a prohibitively large constraint set and does not admit an efficient separation oracle, we still have one tool at our disposal: we can modify $\mathcal{H}(\rho(\mathcal{O}))$ to get an alternate price space Π which we *can* work with efficiently. Recall that in Section 3, we arrived at the requirement that $\Pi = \mathcal{H}(\rho(\mathcal{O}))$ as a necessary conclusion of the proposed conditions on our market maker. If we wish to violate this requirement, we need to consider which conditions must be weakened and revise the resulting guarantees from Section 3.

We will continue to construct duality-based cost function market makers in the usual way, via the tuple $(\mathcal{O}, \rho, \Pi, R)$. Π is still a convex compact set of feasible prices. But we now allow Π to be distinct from $\mathcal{H}(\rho(\mathcal{O}))$. Not surprisingly, the choice of Π will affect the interest of the traders and the market maker. We prove several claims which will aid us in our market design. Theorem 6.1 tells us that the expressiveness condition should not be relaxed, while Theorem 6.2 tells us that the no-arbitrage condition can be. Together, these imply that we may safely choose Π to be a *superset* of $\mathcal{H}(\rho(\mathcal{O}))$.

The first (perhaps surprising) theorem tells us that expressiveness is not only useful for information aggregation, it is actually *necessary* for the market maker to avoid unbounded loss. The proof involves showing that if \circ is the final outcome and $\rho(\circ) \notin \Pi$, then it is possible to make an infinite sequence of trades such that each trade causes a constant amount of loss to the market maker.

¹⁰More precisely, a separation oracle returns any *separating hyperplane* that divides the input from the feasible set.

THEOREM 6.1. *For any duality-based cost function market maker, the worst-case loss of the market maker is unbounded if $\rho(\mathcal{O}) \not\subseteq \Pi$.*

PROOF. Consider some outcome \circ such that $\rho(\circ) \notin \Pi$. By definition, the feasible price set $\Pi = \{\nabla C(\mathbf{q}) : \forall \mathbf{q}\}$ is compact. Because $\rho(\circ) \notin \Pi$, there exists a hyperplane that *strongly* separates Π and $\rho(\circ)$. In other words, there exists an $k > 0$ such that $\|\rho(\circ) - \nabla C(\mathbf{q})\| \geq k, \forall \mathbf{q}$.

When outcome \circ is realized, $B(\mathbf{q}) = \rho(\circ) \cdot \mathbf{q} - C(\mathbf{q}) + C(\mathbf{0})$ is the market maker's loss given \mathbf{q} . We have $\nabla B(\mathbf{q}) = \rho(\circ) - \nabla C(\mathbf{q})$, which represents the instantaneous change of the market maker's loss. For infinitesimal ϵ , let $\mathbf{q}' = \mathbf{q} + \epsilon(\rho(\circ) - \nabla C(\mathbf{q}))$. Then

$$\begin{aligned} B(\mathbf{q}') &= B(\mathbf{q}) + \nabla B(\mathbf{q}) \cdot [\epsilon(\rho(\circ) - \nabla C(\mathbf{q}))] \\ &= B(\mathbf{q}) + \epsilon \|\rho(\circ) - \nabla C(\mathbf{q})\|^2 \geq B(\mathbf{q}) + \epsilon k^2. \end{aligned}$$

This shows that for any \mathbf{q} we can find a \mathbf{q}' such that the market maker's worst-case loss is at least increased by ϵk^2 . This process can continue for infinite steps. Hence, we conclude that the market maker's loss is unbounded. \square

In the following theorem, which is a simple extension of Theorem 4.4, we see that including additional price vectors in Π does not adversely impact the market maker's worst-case loss, despite the fact that the no-arbitrage condition is violated.

THEOREM 6.2. *Consider any duality-based cost function market maker with R and Π satisfying $\sup_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) < \infty$ and $\mathcal{H}(\rho(\mathcal{O})) \subseteq \Pi$. Assume that the initial price vector satisfies $\nabla C(\mathbf{0}) \in \mathcal{H}(\rho(\mathcal{O}))$. Let \mathbf{q} denote the vector of quantities sold and \circ denote the true outcome. The monetary loss of the market maker is no more than*

$$R(\rho(\circ)) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) - D_R(\rho(\circ), \nabla C(\mathbf{q})).$$

PROOF. This proof is nearly identical to the proof of Theorem 4.4. The only major difference is that now $C(\mathbf{0}) = -\min_{\mathbf{x} \in \Pi} R(\mathbf{x})$ instead of $C(\mathbf{0}) = -\min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x})$, but this is equivalent since we have assumed that $\nabla C(\mathbf{0}) \in \mathcal{H}(\rho(\mathcal{O}))$. $R(\rho(\circ))$ is still well-defined and finite since we have assumed that $\mathcal{H}(\rho(\mathcal{O})) \subseteq \Pi$. \square

This tells us that expanding Π can only help the market maker; increasing the range of $\nabla C(\mathbf{q})$ can only increase the divergence term. This may seem somewhat counterintuitive. We originally required that $\Pi \subseteq \mathcal{H}(\rho(\mathcal{O}))$ as a consequence of the no-arbitrage condition, and by relaxing this condition, we are providing traders with potential arbitrage opportunities. However, these arbitrage opportunities do not hurt the market maker. As long as the initial price vector lies in $\mathcal{H}(\rho(\mathcal{O}))$, any such situations where a trader can earn a guaranteed profit are effectively created (and paid for) by other traders! In fact, if the final price vector $\nabla C(\mathbf{q})$ falls outside the convex hull, the divergence term will be strictly positive, improving the bound.

To elaborate on this point, let's consider an example where Π is strictly larger than $\mathcal{H}(\rho(\mathcal{O}))$. Let \mathbf{q} be the current vector of purchases, and assume the associated price vector $\mathbf{x} = \nabla C(\mathbf{q})$ lies in the interior of $\mathcal{H}(\rho(\mathcal{O}))$. Consider a trader who purchases a bundle \mathbf{r} such that the new price vector leaves this set, i.e., $\mathbf{y} := \nabla C(\mathbf{q} + \mathbf{r}) \notin \mathcal{H}(\rho(\mathcal{O}))$. We claim that this choice can be strictly improved in the sense that there is an alternative bundle \mathbf{r}' whose associated profit, for any outcome \circ , is strictly greater than the profit for \mathbf{r} .

For simplicity, assume \mathbf{y} is an interior point of $\Pi \setminus \mathcal{H}(\rho(\mathcal{O}))$ so that $\mathbf{q} + \mathbf{r} = \nabla R(\mathbf{y})$. Define $\pi(\mathbf{y}) := \arg \min_{\mathbf{y}' \in \mathcal{H}(\rho(\mathcal{O}))} D_R(\mathbf{y}', \mathbf{y})$, the minimum divergence projection of \mathbf{y} into $\mathcal{H}(\rho(\mathcal{O}))$. The alternative bundle we consider is $\mathbf{r}' = \nabla R(\pi(\mathbf{y})) - \mathbf{q}$. Our trader

pays $C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q} + \mathbf{r}')$ less to purchase \mathbf{r}' than to purchase \mathbf{r} . Hence, for any outcome \mathfrak{o} , we see that the increased profit for \mathbf{r}' over \mathbf{r} is

$$\begin{aligned} \rho(\mathfrak{o}) \cdot (\mathbf{r}' - \mathbf{r}) - C(\mathbf{q} + \mathbf{r}') + C(\mathbf{q} + \mathbf{r}) &> \rho(\mathfrak{o}) \cdot (\mathbf{r}' - \mathbf{r}) + \nabla C(\mathbf{q} + \mathbf{r}') \cdot (\mathbf{r} - \mathbf{r}') \\ &= (\rho(\mathfrak{o}) - \pi(\mathbf{y})) \cdot (\mathbf{r}' - \mathbf{r}). \end{aligned} \quad (13)$$

Notice that we achieve strict inequality precisely because $\nabla C(\mathbf{q} + \mathbf{r}') = \pi(\mathbf{y}) \neq \mathbf{y} = \nabla C(\mathbf{q} + \mathbf{r})$. Now use the optimality condition for $\pi(\mathbf{y})$ to see that, since $\rho(\mathfrak{o}) \in \mathcal{H}(\rho(\mathcal{O}))$, $\nabla_{\pi(\mathbf{y})}(D_R(\pi(\mathbf{y}), \mathbf{y})) \cdot (\rho(\mathfrak{o}) - \pi(\mathbf{y})) \geq 0$. It is easy to check that $\nabla_{\pi(\mathbf{y})}(D_R(\pi(\mathbf{y}), \mathbf{y})) = \nabla R(\pi(\mathbf{y})) - \nabla R(\mathbf{y}) = \mathbf{r}' - \mathbf{r}$. Combining this last expression with the inequality above and (13) tells us that the profit increase is strictly greater than $(\rho(\mathfrak{o}) - \pi(\mathbf{y})) \cdot (\mathbf{r}' - \mathbf{r}) \geq 0$. Simply put, the trader receives a guaranteed positive increase in profit for any outcome \mathfrak{o} .

The next theorem shows that any time the price vector lies outside of $\rho(\mathfrak{o})$, traders could profit by moving it back inside. The proof uses a nice application of minimax duality for convex-concave functions.

THEOREM 6.3. *For any duality-based cost function market maker, given a current quantity vector \mathbf{q}_0 with current price vector $\nabla C(\mathbf{q}_0) = \mathbf{x}_0$, a trader has the opportunity to earn a guaranteed profit of at least $\min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} D_R(\mathbf{x}, \mathbf{x}_0)$.*

PROOF. A trader looking to earn a guaranteed profit when the current quantity is \mathbf{q}_0 hopes to purchase a bundle \mathbf{r} so that the worst-case profit $\min_{\mathfrak{o} \in \mathcal{O}} \rho(\mathfrak{o}) \cdot \mathbf{r} - C(\mathbf{q}_0 + \mathbf{r}) + C(\mathbf{q}_0)$ is as large as possible. Notice that this quantity is strictly positive since $\mathbf{r} = \mathbf{0}$, which always has 0 profit, is one option. Thus, a trader would like to solve the following objective:

$$\begin{aligned} &\max_{\mathbf{r} \in \mathbb{R}^K} \min_{\mathfrak{o} \in \mathcal{O}} \rho(\mathfrak{o}) \cdot \mathbf{r} - C(\mathbf{q}_0 + \mathbf{r}) + C(\mathbf{q}_0) \\ &= \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} \max_{\mathbf{r} \in \mathbb{R}^K} \mathbf{x} \cdot \mathbf{r} - C(\mathbf{q}_0 + \mathbf{r}) + C(\mathbf{q}_0) \\ &= \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} \max_{\mathbf{r} \in \mathbb{R}^K} \mathbf{x} \cdot (\mathbf{q}_0 + \mathbf{r}) - C(\mathbf{q}_0 + \mathbf{r}) + C(\mathbf{q}_0) - \mathbf{x} \cdot \mathbf{q}_0 \\ &= \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) + C(\mathbf{q}_0) - \mathbf{x} \cdot \mathbf{q}_0 \\ &= \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) + \mathbf{x}_0 \cdot \mathbf{q}_0 - R(\mathbf{x}_0) - \mathbf{x} \cdot \mathbf{q}_0 \\ &\geq \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} D_R(\mathbf{x}, \mathbf{x}_0). \end{aligned}$$

The first equality with the min / max swap holds via Sion's Minimax Theorem [Sion 1958]. The last inequality was obtained using the first-order optimality condition of the solution $\mathbf{x}_0 = \arg \max_{\mathbf{x} \in \Pi} \mathbf{x} \cdot \mathbf{q}_0 - R(\mathbf{x})$ for the vector $\mathbf{x} - \mathbf{x}_0$ which holds since $\mathbf{x} \in \Pi$. \square

When $\mathbf{x}_0 \in \mathcal{H}(\rho(\mathcal{O}))$, $D_R(\mathbf{x}, \mathbf{x}_0)$ is minimized when $\mathbf{x} = \mathbf{x}_0$ and the bound is vacuous, as we would expect. The more interesting case occurs when the prices have fallen outside of $\mathcal{H}(\rho(\mathcal{O}))$, in which case a trader is guaranteed a riskless profit by moving $\nabla C(\mathbf{q})$ to the closest point in $\mathcal{H}(\rho(\mathcal{O}))$.

6.3. Pair Betting via Relaxation

We return our attention to the scenario in which the outcome is a ranking of n competitors, as described in Section 3.1. Consider a complex market in which traders make arbitrary pair bets: for every i, j , a trader can purchase a security which pays out \$1

whenever $\pi(i) < \pi(j)$. Like subset bets, pricing pair bets using the LMSR is known to be #P-hard [Chen et al. 2008a].

We can represent the payoff structure of any such outcome π by a matrix M_π defined by

$$M_\pi(i, j) = \begin{cases} 1, & \text{if } \pi(i) < \pi(j) \\ \frac{1}{2}, & \text{if } i = j \\ 0, & \text{if } \pi(i) > \pi(j). \end{cases}$$

We would like to choose our feasible price region as the set $\mathcal{H}(\{M_\pi : \pi \in S_n\})$, where S_n is the set of permutations on $[n]$. Unfortunately, the computation of this convex hull is necessarily hard: if given only a separation oracle for the set $\mathcal{H}(\{M_\pi : \pi \in S_n\})$, we could construct a linear program to solve the “minimum feedback arc set” problem, which is known to be NP-hard [Karp 1972].

On the positive side, we see from the previous section that the market maker can work in a larger feasible price space without risking a larger loss. We thus relax our feasible price region Π to the set of $n \times n$ real-valued matrices $X \in \mathbb{R}^{n^2}$ satisfying the intuitive set of constraints described in Section 3.1:

$$\begin{aligned} X(i, j) &\geq 0 && \forall i, j \in [n] \\ X(i, j) &= 1 - X(j, i) && \forall i, j \in [n] \\ X(i, j) + X(j, k) + X(k, i) &\geq 1 && \forall i, j, k \in [n] \end{aligned}$$

This relaxation was first discussed by Megiddo [1977], who referred to such matrices as *generalized order matrices*. He proved that, for $n \leq 4$, we do have $\Pi = \mathcal{H}(\{M_\pi : \pi \in S_n\})$, but gave a counterexample showing strict containment for $n = 13$. By using this relaxed price space, the market maker allows traders to bring the price vector outside of the convex hull, yet includes a set of basic (and natural) constraints on the prices. Such a market could be implemented with any strongly convex conjugate function (e.g., quadratic).

Notice that in this example, it is computationally hard in general for a trader to determine whether or not a particular price vector falls within the convex hull; if this were not the case, then we would be able to construct a separation oracle, and could price pair bets efficiently without the relaxation. Therefore, although arbitrage opportunities may be created, it is generally intractable for traders to find and exploit these opportunities.

7. RELATION TO ONLINE LEARNING

In this section, we use our framework to explore the striking mathematical connections that exist between automated market makers and the class of Follow the Regularized Leader algorithms for online learning. While the problem of learning in an online environment appears quite different semantically from the problem of pricing securities in a market, we show that the two frameworks have a strong syntactic correspondence. We begin with a brief overview of no-regret learning and the online linear optimization problem.

7.1. Online Learning and Regret-Minimizing Algorithms

Perhaps the most canonical example of online, no-regret learning is the problem of *learning from expert advice*. In the expert setting, we imagine an algorithm that must make a sequence of predictions based on the advice of a set of N *experts* and receive a

corresponding sequence of *losses*.¹¹ The goal of the algorithm is to achieve a *cumulative loss* that is “almost as low” as the cumulative loss of the best performing expert in hindsight. No statistical assumptions are made about these losses. Indeed, algorithms are expected to perform well even if the sequence of losses is chosen by an adversary.

Formally, at every time step $t \in \{1, \dots, T\}$, every expert $i \in \{1, \dots, N\}$ receives a loss $\ell_{i,t} \in [0, 1]$. The cumulative loss of expert i at time T is then defined as $L_{i,T} = \sum_{t=1}^T \ell_{i,t}$. An algorithm \mathcal{A} maintains a weight $w_{i,t}$ for each expert i at time t , where $\sum_{i=1}^N w_{i,t} = 1$. These weights can be viewed as a distribution over the experts. The algorithm then receives its own instantaneous loss $\ell_{\mathcal{A},t} = \sum_{i=1}^N w_{i,t} \ell_{i,t}$, which can be interpreted as the expected loss the algorithm would receive if it always chose an expert to follow according to the current distribution. The cumulative loss of \mathcal{A} up to time T is defined in the natural way as $L_{\mathcal{A},T} = \sum_{t=1}^T \ell_{\mathcal{A},t} = \sum_{t=1}^T \sum_{i=1}^N w_{i,t} \ell_{i,t}$. Below we use the symbols ℓ_t , \mathbf{L}_t , and \mathbf{w}_t to refer to the vector of losses, the vector of cumulative loss, and the vector of weights, respectively, for each expert on round t .

It is unreasonable to expect the algorithm to achieve a small cumulative loss if none of the experts perform well. For this reason, it is typical to measure the performance of an algorithm in terms of its *regret*, defined to be the difference between the cumulative loss of the algorithm and the loss of the best performing expert, that is,

$$L_{\mathcal{A},T} - \min_{i \in \{1, \dots, N\}} L_{i,T}.$$

An algorithm is said to have *no regret* if the average per time step regret approaches 0 as T approaches infinity.

The popular Randomized Weighted Majority (WM) algorithm [Littlestone and Warmuth 1994; Freund and Schapire 1997] is an example of a no-regret algorithm. Weighted Majority uses weights

$$w_{i,t} = \frac{e^{-\eta L_{i,t-1}}}{\sum_{j=1}^N e^{-\eta L_{j,t-1}}}, \quad (14)$$

where $\eta > 0$ is a tunable parameter known as the *learning rate*. It is well known that the regret of WM after T trials can be bounded as

$$L_{WM(\eta),T} - \min_{i \in \{1, \dots, N\}} L_{i,T} \leq \eta T + \frac{\log N}{\eta}.$$

When T is known in advance, setting $\eta = \sqrt{\log N/T}$ yields the standard $O(\sqrt{T \log N})$ regret bound.

It has been shown that the weights chosen by Weighted Majority are precisely those that minimize a combination of empirical loss and an entropy-based regularization term [Kivinen and Warmuth 1997; 1999; Helmbold and Warmuth 2009]. More specifically, the weight vector \mathbf{w}_t at time t is precisely the solution to the following minimization problem:

$$\min_{\mathbf{w} \in \Delta_N} \mathbf{w} \cdot \mathbf{L}_{t-1} - \frac{1}{\eta} \mathbf{H}(\mathbf{w})$$

where \mathbf{H} is the entropy function, $\mathbf{H}(\mathbf{w}) := -\sum_{i=1}^N w_i \log w_i$. Indeed, Weighted Majority is an example of broader class of algorithms collectively known as *Follow the Regularized Leader* (FTRL) algorithms [Shalev-Shwartz and Singer 2007; Hazan and Kale

¹¹This framework could be formalized equally well in terms of rewards, but losses are more common in the literature.

2010; Hazan 2009]. The FTRL template can be applied to a wide class of learning problems that fall under a general framework commonly known as *online convex optimization* [Zinkevich 2003]. Other problems that fall into this framework include online linear pattern classification [Kivinen and Warmuth 1997], online Gaussian density estimation [Azoury and Warmuth 2001], and online portfolio selection [Cover 1991]. In Algorithm 1, we present a version of FTRL tailored to the *online linear optimization* problem, an extension of the expert setting in which weights \mathbf{w}_t are chosen from a fixed bounded convex action space $\mathcal{K} \subset \mathbb{R}^N$. Notice that the experts setting is just a special case of online linear optimization, where the set \mathcal{K} is the N -simplex Δ_N .

ALGORITHM 1: Follow the Regularized Leader (FTRL)

- 1: Input: convex compact decision set $\mathcal{K} \subset \mathbb{R}^N$
- 2: Input: strictly convex differentiable regularization function $\mathcal{R}(\cdot)$ defined on \mathcal{K}
- 3: Parameter: $\eta > 0$
- 4: Initialize: $\mathbf{L}_0 = \langle 0, \dots, 0 \rangle$
- 5: **for** $t = 1, \dots, T$ **do**
- 6: The learner selects action $\mathbf{w}_t \in \mathcal{K}$ according to:

$$\mathbf{w}_t := \underset{\mathbf{w} \in \mathcal{K}}{\operatorname{argmin}} \mathbf{L}_{t-1} \cdot \mathbf{w} + \frac{1}{\eta} \mathcal{R}(\mathbf{w}) \quad (15)$$

- 7: Nature reveals ℓ_t , learner suffers loss $\ell_t \cdot \mathbf{w}_t$
 - 8: The learner updates $\mathbf{L}_t = \mathbf{L}_{t-1} + \ell_t$
 - 9: **end for**
-

For a complete description of the FTRL algorithm, we refer the reader to the excellent notes of Rakhlin [2009]. We will make use of a result from these notes, but we first include two additional assumptions that we will use to make the connection to duality-based cost function market makers. In the remainder of this section, we use $\|\cdot\|$ to denote the L2 norm.

ASSUMPTION 1. *For each time step t , $\|\ell_t\| \leq 1$.*

ASSUMPTION 2. *The regularizer $\mathcal{R}(\cdot)$ has the Legendre property defined in Section 11.2 of Cesa-Bianchi and Lugosi [2006]: \mathcal{R} is strictly convex on $\operatorname{relint}(\mathcal{K})$ and $\|\nabla \mathcal{R}(\mathbf{w})\| \rightarrow \infty$ as $\mathbf{w} \rightarrow \operatorname{relbnd}(\mathcal{K})$.*

Under the latter assumption, the solution to Equation 15 will always occur in the relative interior of \mathcal{K} , which implies that the optimization is effectively unconstrained. We can now utilize Corollary 9 of Rakhlin [2009] to obtain the following.

PROPOSITION 7.1. *Under Assumptions 1 and 2, the FTRL algorithm enjoys the following regret bound: For any $\mathbf{w}^* \in \mathcal{K}$,*

$$\sum_{t=1}^T \ell_t \cdot \mathbf{w}_t - \sum_{t=1}^T \ell_t \cdot \mathbf{w}^* \leq \frac{1}{\eta} \left(\mathcal{R}(\mathbf{w}^*) - \mathcal{R}(\mathbf{w}_1) - D_{\mathcal{R}}(\mathbf{w}^*, \mathbf{w}_{T+1}) + \sum_{t=1}^T D_{\mathcal{R}}(\mathbf{w}_t, \mathbf{w}_{t+1}) \right).$$

This proposition may not be so illuminating at first glance, but it expresses a fundamental tradeoff in the learning problem. If we choose a regularizer \mathcal{R} with heavy curvature, or equivalently if we choose a small η , then given the nature of the optimization problem in Equation 15, we ensure that the updates $\mathbf{w}_t \rightarrow \mathbf{w}_{t+1}$ are “small” and hence $D_{\mathcal{R}}(\mathbf{w}_t, \mathbf{w}_{t+1})$ will be small. On the other hand, we pay for either of these choices since (a) the bound is proportional to $1/\eta$, and (b) the difference $\mathcal{R}(\mathbf{w}^*) - \mathcal{R}(\mathbf{w}_1)$ grows larger when \mathcal{R} has more curvature.

Under certain reasonable assumptions on \mathcal{R} , it is possible to prove that $D_{\mathcal{R}}(\mathbf{w}_t, \mathbf{w}_{t+1}) \leq O(\eta^2)$. For example, if \mathcal{R} is strongly convex (with respect to the L2 norm), then $D_{\mathcal{R}}(\mathbf{w}_t, \mathbf{w}_{t+1}) \leq \eta^2 \|\ell_t\|^2$. See Rakhlin [2009] for more details.

COROLLARY 7.2. *Suppose that there exists $B > 0$ such that for every t , $D_{\mathcal{R}}(\mathbf{w}_t, \mathbf{w}_{t+1}) \leq B\eta^2$, and that there exists $C > 0$ such that $\mathcal{R}(\mathbf{w}^*) - \mathcal{R}(\mathbf{w}_1) \leq C$. Then $\text{Regret}(\text{FTRL}) \leq C/\eta + \eta BT$. If $\eta = \sqrt{C/BT}$, then $\text{Regret}(\text{FTRL}) \leq 2\sqrt{BCT}$.*

This final bound is quite powerful. It says that the regret of FTRL on any online linear optimization problem is always on the order of \sqrt{T} . The constant in front of this rate will depend on the total variation of the regularization function on \mathcal{K} (that is, $\mathcal{R}(\mathbf{w}^*) - \mathcal{R}(\mathbf{w}_1)$) as well as the stability of the updates (that is, the terms $D_{\mathcal{R}}(\mathbf{w}_t, \mathbf{w}_{t+1})$).

7.2. An Equivalence Between Online Learning and Market Making

Having reviewed much of the literature on the design of online learning algorithms, we now pivot back to the primary topic at hand, the design of market makers for complex security spaces. We will see that the tools that have been developed for the online learning setting are strikingly similar to those we have constructed for selecting pricing mechanisms. This is rather surprising, as the problem of learning in an online environment is semantically quite distinct from the problem of pricing securities in a prediction market: a learning algorithm receives *losses* and selects *weights* whereas a market maker manages *trades* and sets *prices*. We now show how these two problems can be viewed as two sides of the same coin. The two frameworks have very different semantics yet, in a very strong sense, have nearly identical *syntax*.

The relationship is described in full detail in Figure 1. We imagine that the learner uses the FTRL algorithm (Algorithm 1) to select weights, and the market uses the duality-based cost function market maker framework.

What we emphasize in Figure 1 is that, by identifying the objects Π , $R(\cdot)$, and $\{\mathbf{r}_t\}$ with the objects \mathcal{K} , $\mathcal{R}(\cdot)/\eta$, and $\{-\ell_t\}$, respectively, the mechanisms for choosing an instantaneous price vector $\mathbf{x}_t \in \Pi$ and selecting a weight vector $\mathbf{w}_t \in \mathcal{K}$ are identical. Put another way, if we consider security bundles \mathbf{r}_t as the negative loss vectors ℓ_t , then the duality-based cost function market maker becomes exactly FTRL.

The connection seems to break down when we arrive at the last pair of statements, as the FTRL regret and the market maker's worst-case loss do not appear to be identical. Strictly speaking this is true. However, these two quantities are not so far apart. Using the previous identification, we see that the term $\max_{\mathbf{x} \in \Pi} \mathbf{x} \cdot \mathbf{q}_T$, representing the worst-case payout of the market maker, matches exactly the term $-\min_{\mathbf{w} \in \mathcal{K}} \mathbf{w} \cdot L_T$. Now let us do a first-order approximation on the negation of the first term, i.e., the market maker's earnings from selling securities. We have

$$C(\mathbf{q}_T) - C(\mathbf{q}_0) = \sum_{t=1}^T C(\mathbf{q}_t) - C(\mathbf{q}_{t-1}) \approx \sum_{t=1}^T \nabla C(\mathbf{q}_{t-1}) \cdot (\mathbf{q}_t - \mathbf{q}_{t-1}) = \sum_{t=1}^T \mathbf{x}_t \cdot \mathbf{r}_t, \quad (16)$$

where we used the fact that the instantaneous price vector \mathbf{x}_t is equal to $\nabla C(\mathbf{q}_{t-1})$. This is not too surprising. Every trader will roughly pay the instantaneous prices \mathbf{x}_t for the securities times the quantities \mathbf{r}_t of each security sold. The total earned by the market maker $C(\mathbf{q}_T) - C(\mathbf{q}_0)$ is then roughly the sum of these payments over all trades.

How bad is this approximation? We can quantify this explicitly, since the difference between $C(\mathbf{q}_t) - C(\mathbf{q}_{t-1})$ and $\nabla C(\mathbf{q}_{t-1}) \cdot (\mathbf{q}_t - \mathbf{q}_{t-1})$ is exactly the value $D_C(\mathbf{q}_t, \mathbf{q}_{t-1})$. If R has the Legendre property (described in Assumption 2) then via standard arguments [Rockafellar 1970] we can also conclude that $D_C(\mathbf{q}_t, \mathbf{q}_{t-1}) = D_R(\mathbf{x}_t, \mathbf{x}_{t+1})$. Under this assumption, in other words, the worst-case loss of the market maker can be writ-

In an online linear optimization problem using FTRL:	In a market using a duality-based cost function market maker:
<ul style="list-style-type: none"> • the learner is given access to a fixed space of weights \mathcal{K} • the learning algorithm must select a weight vector $\mathbf{w} \in \mathcal{K}$ • the learner uses a convex regularizer $\mathcal{R}(\cdot)$, which is a parameter of FTRL • the learner receives loss vectors ℓ_t • the learning algorithm maintains a cumulative loss vector \mathbf{L}_t and updates according to $\mathbf{L}_t \leftarrow \mathbf{L}_{t-1} + \ell_t$ 	<ul style="list-style-type: none"> • the market maker has an outcome space \mathcal{O} and payoff function ρ, which define a feasible price space $\Pi = \mathcal{H}(\rho(\mathcal{O}))$ • the market maker must select instantaneous security prices $\mathbf{x} \in \Pi$ • the market maker uses a convex conjugate $R(\cdot)$, which is a parameter of the pricing function $C(\cdot)$ • the market maker receives security bundle purchases \mathbf{r}_t • the market maker maintains a quantity vector \mathbf{q}_t and updates according to $\mathbf{q}_t \leftarrow \mathbf{q}_{t-1} + \mathbf{r}_t$
<ul style="list-style-type: none"> • FTRL selects the weights by solving $\mathbf{w}_t = \operatorname{argmin}_{\mathbf{w} \in \mathcal{K}} \mathbf{w} \cdot \mathbf{L}_{t-1} + \frac{1}{\eta} \mathcal{R}(\mathbf{w})$ • the learner suffers regret $\sum_{t=1}^T \mathbf{w}_t \cdot \ell_t - \min_{\mathbf{w} \in \mathcal{K}} \mathbf{w} \cdot \mathbf{L}_T$ 	<ul style="list-style-type: none"> • the market mechanism sets prices via $\mathbf{x}_t = \operatorname{argmax}_{\mathbf{x} \in \Pi} \mathbf{x} \cdot \mathbf{q}_{t-1} - R(\mathbf{x})$ • the market maker suffers worst-case loss $C(\mathbf{q}_0) - C(\mathbf{q}_T) + \max_{\mathbf{x} \in \Pi} \mathbf{x} \cdot \mathbf{q}_T$

Fig. 1. The similarities between the duality-based cost function market maker framework and the Follow the Regularized Leader algorithm for online linear optimization.

ten as

$$\max_{\mathbf{x} \in \Pi} \mathbf{x} \cdot \mathbf{q}_T - \sum_{t=1}^T \mathbf{x}_t \cdot \mathbf{r}_t - \sum_{t=1}^T D_R(\mathbf{x}_t, \mathbf{x}_{t+1}).$$

Putting everything together, this final bound is exactly what we should expect. Look again at Theorem 6.2 and Proposition 7.1. The bounds in these theorems are nearly identical under the translation matching $\mathbf{w}^* \leftrightarrow \rho(\mathbf{o})$, $\mathbf{w}_{T+1} \leftrightarrow \nabla C(\mathbf{q})$, and $R(\mathbf{x}) \leftrightarrow \mathcal{R}(\mathbf{w})/\eta$, since by definition of FTRL, $\mathbf{w}_1 = \operatorname{argmin}_{\mathbf{w} \in \mathcal{K}} \mathcal{R}(\mathbf{w})$. The key difference is that the sum of divergence terms seems to get “lost in translation” when we look at Theorem 6.2. The above equation tells us why this is.

It is worth looking further into this key difference between the FTRL algorithm for online linear optimization and our proposed automated market maker. We could imagine a modified market maker with a different mechanism: after the $(t-1)$ th trade the market maker posts the (instantaneous) price vector \mathbf{x}_t , a trader arrives to purchase bundle \mathbf{r}_t , and the trader pays exactly $\mathbf{x}_t \cdot \mathbf{r}_t$. Notice this is different from the original

framework, where the trader would pay $C(\mathbf{q} + \mathbf{r}_t) - C(\mathbf{q})$, although we observed in Equation 16 that these two values are not so far apart.

Under the mapping outlined in Figure 1, algorithms for the expert setting ($\mathcal{K} = \Delta_n$) correspond to complete markets. Weighted Majority corresponds directly to the LMSR, with the learning rate η playing a similar role to the LMSR parameter b . The similarity between the Weighted Majority weights (Equation 14) and the LMSR prices (Equation 2) has been observed and exploited in the past [Chen et al. 2008a]. The Quad-SCPM market [Agrawal et al. 2011] can be mapped to online gradient descent, which is known to be equivalent to FTRL with a quadratic regularizer [Hazan et al. 2007; Hazan 2009].

8. RELATION TO MARKET SCORING RULES

We have described ways in which our optimization-based framework can be used to derive novel, efficient automated market makers for markets in which the outcome space is very large. Our framework also provides new insights into the complete market setting. In this section, we describe how our framework can be used to establish a correspondence between cost function based markets and market scoring rules.

Consider the special case of complete markets, and in particular, markets that offer n Arrow-Debreu securities for the n mutually exclusive and exhaustive outcomes. Our framework defines a set of market makers by equating the set of allowable prices Π to the n -simplex. That is, a market maker for a complete market that satisfies conditions 1–5 in Section 3 can use a cost function of the form

$$C(\mathbf{q}) = \sup_{\mathbf{x} \in \text{relint}(\Delta_n)} \mathbf{x} \cdot \mathbf{q} - R(\mathbf{x}), \quad (17)$$

where $R(\mathbf{x})$ is strictly convex over Δ_n . The market price $\mathbf{x}(\mathbf{q}) = \nabla C(\mathbf{q})$ is the optimal solution to the convex optimization. It is easy to check that when $R(\mathbf{x}) = b \sum_{i=1}^n x_i \log x_i$, the negative entropy function, we have the LMSR market maker. The LMSR is a popular example of a large class of market makers, called *market scoring rules* (MSR). In this section, after reviewing the notion of a proper scoring rule and describing the class of MSRs, we use Equation 17 to establish a correspondence between MSRs and cost function based market makers for complete markets.

8.1. Proper Scoring Rules

Scoring rules have long been used in the evaluation of probabilistic forecasts. In the context of information elicitation, scoring rules are used to encourage individuals to make careful assessments and truthfully report their beliefs [Savage 1971; Garthwaite et al. 2005; Lambert et al. 2008]. In the context of machine learning, scoring rules are used as loss functions to evaluate and compare the performance of different algorithms [Buja et al. 2005; Reid and Williamson 2009]. We briefly mention recent work of Abernethy and Frongillo [2011] who used a generalized notion of a scoring rule in order to construct a market mechanism for solving machine learning problems.

Formally, let $\{1, \dots, n\}$ be a set of mutually exclusive and exhaustive outcomes of a future event. A scoring rule s maps a probability distribution \mathbf{p} over outcomes to a score $s_i(\mathbf{p})$ for each outcome i , with $s_i(\mathbf{p})$ taking values in the range $[-\infty, \infty]$. Intuitively, this score represents the reward that a forecaster receives for predicting the distribution \mathbf{p} if the outcome turns out to be i . A scoring rule is said to be *regular* relative to the probability simplex Δ_n if $\sum_{i=1}^n p_i s_i(\mathbf{p}') \in [-\infty, \infty)$ for all $\mathbf{p}, \mathbf{p}' \in \Delta_n$, with $\sum_{i=1}^n p_i s_i(\mathbf{p}) \in (-\infty, \infty)$. This implies that $s_i(\mathbf{p}) \in (-\infty, \infty)$ whenever $p_i > 0$ and $s_i(\mathbf{p})$ may equal to $-\infty$ when $p_i = 0$. A scoring rule is said to be *proper* if a risk-neutral forecaster who believes the true distribution over outcomes to be \mathbf{p} has no incentive to report any alternate distribution \mathbf{p}' , that is, if $\sum_{i=1}^n p_i s_i(\mathbf{p}) \geq \sum_{i=1}^n p_i s_i(\mathbf{p}')$ for all

distributions $\mathbf{p}' \in \Delta_n$. The rule is *strictly proper* if this inequality holds with equality only when $\mathbf{p} = \mathbf{p}'$.

Two examples of regular, strictly proper scoring rules commonly used in both information elicitation and machine learning are the quadratic scoring rule [Brier 1950]:

$$s_i(\mathbf{p}) = a_i + b \left(2p_i - \sum_{i=1}^n p_i^2 \right) \quad (18)$$

and the logarithmic scoring rule [Good 1952]:

$$s_i(\mathbf{p}) = a_i + b \log(p_i) \quad (19)$$

where $b > 0$ and a_1, \dots, a_n are parameters.

Proper scoring rules are closely related to convex functions. In fact, the following characterization theorem of Gneiting and Raftery [2007], which is credited to McCarthy [1956] and Savage [1971], gives the precise relationship between convex functions and proper scoring rules.

THEOREM 8.1 (GNEITING AND RAFTERY [2007]). *A regular scoring rule is (strictly) proper if and only if there exists a (strictly) convex function $G : \Delta_n \rightarrow \mathbb{R}$ such that for all $i \in \{1, \dots, n\}$,*

$$s_i(\mathbf{p}) = G(\mathbf{p}) - G'(\mathbf{p}) \cdot \mathbf{p} + G'_i(\mathbf{p}),$$

where $G'(\mathbf{p})$ is any subgradient of G at the point \mathbf{p} and $G'_i(\mathbf{p})$ is the i -th element of $G'(\mathbf{p})$.

Note that for a scoring rule defined in terms of a function G ,

$$\sum_{i=1}^n p_i s_i(\mathbf{p}) = \sum_{i=1}^n p_i (G(\mathbf{p}) - G'(\mathbf{p}) \cdot \mathbf{p} + G'_i(\mathbf{p})) = G(\mathbf{p}).$$

Theorem 8.1 therefore indicates that a regular scoring rule is (strictly) proper if and only if its expected score function $G(\mathbf{p})$ is (strictly) convex on Δ_n , and the vector with elements $s_i(\mathbf{p})$ is a subgradient of G at the point \mathbf{p} . Hence, every bounded convex function G over Δ_n induces a proper scoring rule.

Define $S(\tilde{\mathbf{p}}, \mathbf{p}) = \sum_{i=1}^n p_i s_i(\tilde{\mathbf{p}})$ to be the expected score of a forecaster who has belief \mathbf{p} but predicts $\tilde{\mathbf{p}}$. Then, $G(\mathbf{p}) = S(\mathbf{p}, \mathbf{p})$. If a scoring rule is regular and proper, $d(\tilde{\mathbf{p}}, \mathbf{p}) = S(\mathbf{p}, \mathbf{p}) - S(\tilde{\mathbf{p}}, \mathbf{p})$ is the associated *divergence function* that captures the expected loss in score if a forecaster predicts $\tilde{\mathbf{p}}$ rather than his true belief \mathbf{p} . It is known that if $G(\mathbf{p})$ is differentiable, the divergence function is the Bregman divergence for G , that is, $d(\tilde{\mathbf{p}}, \mathbf{p}) = D_G(\tilde{\mathbf{p}}, \mathbf{p})$. For a nice survey on uses, properties, and characterizations of proper scoring rules, see Gneiting and Raftery [2007].

8.2. Market Scoring Rules

Market scoring rules (MSR) were developed by Hanson [2003; Hanson [2007] as a method of using scoring rules to pool opinions from many different forecasters. Market scoring rules are sequentially shared scoring rules. Formally, the market maintains a current probability distribution \mathbf{p} . At any time, a trader can enter the market and change this distribution to an arbitrary distribution \mathbf{p}' of her choice.¹² If the outcome turns out to be i , she receives a (possibly negative) payoff of $s_i(\mathbf{p}') - s_i(\mathbf{p})$. For example, in the MSR defined using the logarithmic scoring rule in Equation 19, a trader

¹²In some market scoring rules, such as the LMSR, distributions that place a weight of 0 on any outcome are not allowed since a trader would have to pay an infinite amount of money if the outcome with reported probability 0 actually occurred.

who changes the distribution from \mathbf{p} to \mathbf{p}' receives a payoff of $b \log(p'_i/p_i)$. This market formulation is equivalent to the cost function based formulation of the LMSR (hence its name) in the sense that a trader who changes the market probabilities from \mathbf{p} to \mathbf{p}' in the MSR formulation receives the same payoff for every outcome i as a trader who changes the quantity vectors from any \mathbf{q} to \mathbf{q}' such that market prices satisfy $\mathbf{x}(\mathbf{q}) = \mathbf{p}$ and $\mathbf{x}(\mathbf{q}') = \mathbf{p}'$ in the cost function based formulation. Using proper scoring rules, market scoring rules preserve the nice incentive compatible property of proper scoring rules for *myopic* traders. A trader who believes the true distribution to be \mathbf{p} and only cares about payoff of her current action, maximizes her expected payoff by changing the market's distribution to \mathbf{p} .

One advantage of the market scoring rule formulation is the ease of bounding the market maker's worst-case loss. Each trader in a market scoring rule is essentially responsible for paying the previous trader's score. Thus the market maker is responsible only for paying the score of the final trader. Let \mathbf{p}_0 be the initial probability distribution of the market. The worst-case loss of the market maker is then

$$\max_{i \in \{1, \dots, n\}} \sup_{\mathbf{p} \in \Delta_n} (s_i(\mathbf{p}) - s_i(\mathbf{p}_0)).$$

The LMSR market maker is not the only market that can be defined as either a market scoring rule or a cost function based market. The fact that there exists a correspondence between certain market scoring rules and certain cost function based markets was noted by Chen and Pennock [2007]. They pointed out that the MSR with scoring function s and the cost function based market with cost function C are equivalent if for all \mathbf{q} and all outcomes i , $C(\mathbf{q}) = q_i - s_i(\mathbf{x}(\mathbf{q}))$. However, they provide neither any guarantees about the circumstances under which this condition can be satisfied nor a general way to find the cost function given a market scoring rule; $\mathbf{x}(\mathbf{q})$ is the gradient of $C(\mathbf{q})$ and the condition defines a differential equation. Agrawal et al. [2011] also made use of the equivalence between markets when this strong condition holds. In the next section, we will give very general precise conditions under which an MSR is equivalent to a cost function based market and provide a way to translate a market scoring rule to a cost function based market and vice versa.

8.3. Equivalence between Market Scoring Rules and Cost Function Based Market Makers

Recall that a convex cost function C can be defined as $C(\mathbf{q}) = \sup_{\mathbf{x} \in \text{relint}(\Delta_n)} \sum_{i=1}^n x_i q_i - R(\mathbf{x})$ for a strictly convex function R , namely the convex conjugate of C . According to Theorem 8.1, there is a one-to-one and onto mapping between strictly convex and differentiable R and strictly proper, regular scoring rules with differentiable scoring functions $s_i(\mathbf{x})$, where for every pair we have

$$R(\mathbf{x}) = \sum_{i=1}^n x_i s_i(\mathbf{x}), \quad (20)$$

and

$$s_i(\mathbf{x}) = R(\mathbf{x}) - \sum_{j=1}^n \frac{\partial R(\mathbf{x})}{\partial x_j} x_j + \frac{\partial R(\mathbf{x})}{\partial x_i}. \quad (21)$$

Theorem 8.2 below shows that that the cost function based market using R in (20) and the market scoring rule market using $s_i(\mathbf{x})$ in (21) are equivalent in terms of traders' profits, reachable price vectors, and the market maker's worst-case loss under some mild conditions.

THEOREM 8.2. *Given a strictly convex, continuous conjugate function R and a strictly proper, regular scoring rule s with scoring functions s_i satisfying the relationships in Equations 20 and 21, if both R and s_i 's are differentiable everywhere in $\text{relint}(\Delta_n)$, the corresponding cost function based market and market scoring rule market are equivalent in the following three aspects:*

- (a) *A trade in the cost function based market bringing the quantity vector \mathbf{q} to \mathbf{q}' and the price vector $\mathbf{x}(\mathbf{q})$ to $\mathbf{x}(\mathbf{q}')$ gives the same profit as a trade in the MSR market bringing the market probability from $\mathbf{x}(\mathbf{q})$ to $\mathbf{x}(\mathbf{q}')$ for every outcome i as long as $\mathbf{x}(\mathbf{q}), \mathbf{x}(\mathbf{q}') > \mathbf{0}$.*
- (b) *Given any probability vector \mathbf{x} for which $s_i(\mathbf{x}) \in (-\infty, \infty) \forall i$ in the MSR market, there is always a quantity vector \mathbf{q} such that $\nabla C(\mathbf{q}) = \mathbf{x}$ in the cost function based market.*
- (c) *If the initial probability vector \mathbf{x}_0 in the MSR market is equal to the initial price vector $\nabla C(\mathbf{0})$ in the cost function based market, and $\mathbf{x}_0 \in \text{relint}(\Delta_n)$, then both markets have the same worst-case loss for the market maker.*

PROOF. Because R is continuous and defined on Δ_n , R is bounded on Δ_n . According to Lemma 4.3, for any $\mathbf{q} \in \mathbb{R}^n$,

$$\mathbf{x}(\mathbf{q}) = \nabla C(\mathbf{q}) = \operatorname{argmax}_{\mathbf{x} \in \Delta_n} \left(\sum_{i=1}^n x_i q_i - R(\mathbf{x}) \right) \quad (22)$$

in the cost function based market. Below, we prove each part in turn.

Part (a). Due to Equation 22, if $\mathbf{x}(\mathbf{q}) > \mathbf{0}$, $\mathbf{x}(\mathbf{q})$ must be the optimal solution to the unconstrained optimization problem $\max_{\mathbf{x}} \sum_{i=1}^n x_i q_i - R(\mathbf{x}) - \lambda_{\mathbf{q}} (\sum_{i=1}^n x_i - 1)$ for some $\lambda_{\mathbf{q}}$. Since R is differentiable in $\text{relint}(\Delta_n)$, this means that

$$q_i = \partial R(\mathbf{x}(\mathbf{q})) / \partial x_i(\mathbf{q}) + \lambda_{\mathbf{q}} \quad (23)$$

for some $\lambda_{\mathbf{q}}$.

Suppose in the cost function based market a trader changes the outstanding shares from \mathbf{q} to \mathbf{q}' and this trade changes the market price from $\mathbf{x}(\mathbf{q}) > \mathbf{0}$ to $\mathbf{x}(\mathbf{q}') > \mathbf{0}$. If outcome i occurs, the trader's profit is

$$\begin{aligned} & (q'_i - q_i) - (C(\mathbf{q}') - C(\mathbf{q})) \\ &= (q'_i - q_i) - \left(\sum_{j=1}^n x_j(\mathbf{q}') q'_j - R(\mathbf{x}(\mathbf{q}')) \right) + \left(\sum_{j=1}^n x_j(\mathbf{q}) q_j - R(\mathbf{x}(\mathbf{q})) \right) \\ &= \left(q'_i - \sum_{j=1}^n x_j(\mathbf{q}') q'_j + R(\mathbf{x}(\mathbf{q}')) \right) - \left(q_i - \sum_{j=1}^n x_j(\mathbf{q}) q_j + R(\mathbf{x}(\mathbf{q})) \right) \\ &= \left(\frac{\partial R(\mathbf{x}(\mathbf{q}'))}{\partial x_i(\mathbf{q}')} - \sum_{j=1}^n x_j(\mathbf{q}') \frac{\partial R(\mathbf{x}(\mathbf{q}'))}{\partial x_j(\mathbf{q}')} + R(\mathbf{x}(\mathbf{q}')) \right) \\ &\quad - \left(\frac{\partial R(\mathbf{x}(\mathbf{q}))}{\partial x_i(\mathbf{q})} - \sum_{j=1}^n x_j(\mathbf{q}) \frac{\partial R(\mathbf{x}(\mathbf{q}))}{\partial x_j(\mathbf{q})} + R(\mathbf{x}(\mathbf{q})) \right) \\ &= s_i(\mathbf{x}(\mathbf{q}')) - s_i(\mathbf{x}(\mathbf{q})). \end{aligned}$$

The first equality follows since $\mathbf{x}(\mathbf{q})$ is the solution to $\max_{\mathbf{x} \in \Delta_n} (\sum_{i=1}^n x_i q_i - R(\mathbf{x}))$ and the third equality follows from Equation 23. Since $s_i(\mathbf{x}(\mathbf{q}')) - s_i(\mathbf{x}(\mathbf{q}))$ is the profit of

a trader who changes the market probability from $\mathbf{x}(\mathbf{q})$ to $\mathbf{x}(\mathbf{q}')$ in the MSR market when outcome i occurs, this completes the proof of part (a).

Part (b). In the MSR market, only probability vectors in the set $Y = \{\mathbf{x} \in \Delta_n : s_i(\mathbf{x}) \in (-\infty, \infty) \forall i\}$ can possibly be reported by a trader with finite wealth. Since the scoring rule s is regular, $s_i(\mathbf{x}) \in [-\infty, \infty)$ and it can equal $-\infty$ only when x_i is 0. However, any \mathbf{x} that sets $s_i(\mathbf{x}) = -\infty$ for some i is not allowed, as it requires the trader to pay infinite amount of money when outcome i actually happens.

We show that in the cost function based market it is possible to achieve any price vector $\mathbf{x} \in Y$ by setting $q_i = s_i(\mathbf{x})$ for all i . By strict properness of the scoring rule s , we know that $\sum_{i=1}^n x'_i s_i(\mathbf{x}) - \sum_{i=1}^n x'_i s_i(\mathbf{x}') \leq 0$ for any \mathbf{x} and \mathbf{x}' and the equality holds only when $\mathbf{x}' = \mathbf{x}$. For any vector $\mathbf{x} \in Y$, $\mathbf{s}(\mathbf{x}) \in \mathbb{R}^n$. By Equation 22, we have $\nabla C(\mathbf{s}(\mathbf{x})) = \operatorname{argmax}_{\mathbf{x}' \in \Delta_n} \sum_{i=1}^n x'_i s_i(\mathbf{x}) - \sum_{i=1}^n x'_i s_i(\mathbf{x}') = \mathbf{x}$. Hence, the price vector in the cost function based market is exactly \mathbf{x} .

Part (c). We know that $C(\mathbf{0}) = \max_{\mathbf{x} \in \Delta_n} -R(\mathbf{x})$. If $\mathbf{x}_0 \in \operatorname{relint}(\Delta_n)$, we have $\mathbf{x}_0 = \nabla C(\mathbf{0}) = \operatorname{argmin}_{\mathbf{x} \in \Delta_n} R(\mathbf{x})$ and \mathbf{x}_0 must satisfy

$$\nabla R(\mathbf{x}_0) = \mathbf{0}. \quad (24)$$

Combining Equation 24 with Equation 21, we have

$$s_i(\mathbf{x}_0) = R(\mathbf{x}_0). \quad (25)$$

The worst-case loss of the cost function based market maker is

$$\begin{aligned} \sup_{\mathbf{x} \in \rho(\mathcal{O})} R(\mathbf{x}) - \min_{\mathbf{x} \in \mathcal{H}(\rho(\mathcal{O}))} R(\mathbf{x}) &= \sup_{\mathbf{x} \in \rho(\mathcal{O})} R(\mathbf{x}) - R(\mathbf{x}_0) \\ &= \sup_{\mathbf{x} \in \rho(\mathcal{O})} \sum_i x_i s_i(\mathbf{x}) - R(\mathbf{x}_0) \\ &= \max_{i \in \{1, \dots, n\}} s_i(\mathbf{e}^i) - R(\mathbf{x}_0) \end{aligned} \quad (26)$$

where \mathbf{e}^i is the n -dimensional vector that has 1 for its i -th element and 0 everywhere else. The second equality is due to Equation 20. The third equality is because $\rho(\mathcal{O}) = \{\mathbf{e}^1, \dots, \mathbf{e}^n\}$ for the complete market we consider.

The worst-case loss of the MSR market maker with the scoring functions $s_i(\mathbf{x})$ is

$$\begin{aligned} \max_{i \in \{1, \dots, n\}} \sup_{\mathbf{x} \in \Delta_n} (s_i(\mathbf{x}) - s_i(\mathbf{x}_0)) &= \max_{i \in \{1, \dots, n\}} \sup_{\mathbf{x} \in \Delta_n} s_i(\mathbf{x}) - R(\mathbf{x}_0) \\ &= \max_{i \in \{1, \dots, n\}} s_i(\mathbf{e}^i) - R(\mathbf{x}_0). \end{aligned} \quad (27)$$

The first equality is due to Equation 25. The second equality holds because for strictly proper scoring rule s

$$s_i(\mathbf{e}^i) = \sum_{j=1}^n e_j^i s_j(\mathbf{e}^i) \geq \sum_{j=1}^n e_j^i s_j(\mathbf{x}) = s_i(\mathbf{x})$$

for all $\mathbf{x} \in \Delta_n$.

Equations 26 and 27 are identical. Hence, the worst-case loss of the market maker is the same in these two markets. \square

Theorem 8.2 shows that a trader's profit for moving the prices from \mathbf{x} to \mathbf{x}' can be different in these two markets only when \mathbf{x} or \mathbf{x}' (or both) lie on the relative boundary of Δ_n , and the worst-case loss of the market maker can be different in these two markets only when the initial market price vector lies on the the relative boundary of

Δ_n . The reachable price vectors, however, are always the same. The LMSR market is an example where both the initial market price vector and market prices at any consequent time are in $\text{relint}(\Delta_n)$. The MSR market using a quadratic scoring rule is an example where the initial market price vector is in $\text{relint}(\Delta_n)$ but future market prices can reach the relative boundary of Δ_n . Its corresponding cost function based market maker is equivalent to the Quad-SCPM market introduced by Agrawal et al. [2011].

9. CONCLUSION

We conclude by mentioning one nice direction of work. As we discussed, there is an inherent tradeoff between the bid-ask spread and the worst-case loss of the market maker. But if the market maker chooses to sell securities with an additional *transaction cost* for each security sold, then this money can not only help to cover the worst-case loss, but can also lead to a profit. Furthermore, if a market becomes popular, the market-maker may wish to increase the market depth. This idea has been explored by Othman et al. [2010] for the case of complete markets, introducing a *liquidity sensitive* market maker, and they provide a new model with profit guarantees. Othman and Sandholm [2011] recently extend the work and characterize a family of market makers that are liquidity sensitive. Via our framework, we can define an alternative method for simultaneously including transaction costs and guaranteeing profit. In particular, this is achieved through relaxing the price space, as discussed in Section 6.2. We leave the details to future work.

REFERENCES

- ABERNETHY, J., CHEN, Y., AND VAUGHAN, J. W. 2011. An optimization-based framework for automated market-making. In *Proceedings of the 12th ACM Conference on Electronic Commerce*. 297–306.
- ABERNETHY, J. AND FRONGILLO, R. M. 2011. A collaborative mechanism for crowdsourcing prediction problems. In *Advances in Neural Information Processing Systems*.
- AGRAWAL, S., DELAGE, E., PETERS, M., WANG, Z., AND YE, Y. 2011. A unified framework for dynamic prediction market design. *Operations Research* 59, 3, 550–568.
- AGRAWAL, S., WANG, Z., AND YE, Y. 2008. Parimutuel betting on permutations. In *Proceedings of the 4th International Workshop On Internet And Network Economics*. 126–137.
- ARROW, K. J. 1964. The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies* 31, 2, 91–96.
- ARROW, K. J. 1970. *Essays in the Theory of Risk Bearing*. North Holland, Amsterdam.
- AZOURY, K. S. AND WARMUTH, M. K. 2001. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning* 43, 3, 211–246.
- BERG, J. E., FORSYTHE, R., NELSON, F. D., AND RIETZ, T. A. 2001. Results from a dozen years of election futures markets research. In *Handbook of Experimental Economic Results*, C. A. Plott and V. Smith, Eds.
- BOYD, S. AND VANDENBERGHE, L. 2004. *Convex Optimization*. Cambridge University Press.
- BRAHMA, A., DAS, S., AND MAGDON-ISMAIL, M. 2010. Comparing prediction market structures, with an application to market making. Working paper.
- BRIER, G. 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 78, 1, 1–3.
- BUJA, A., STUETZLE, W., AND SHEN, Y. 2005. Loss functions for binary class probability estimation and classification: Structure and applications. Working draft.
- CESA-BIANCHI, N. AND LUGOSI, G. 2006. *Prediction, Learning, and Games*. Cambridge University Press.
- CHEN, Y., FORTNOW, L., LAMBERT, N., PENNOCK, D. M., AND WORTMAN, J. 2008a. Complexity of combinatorial market makers. In *Proceedings of the 9th ACM Conference on Electronic Commerce*. 190–199.
- CHEN, Y., FORTNOW, L., NIKOLOVA, E., AND PENNOCK, D. M. 2007a. Betting on permutations. In *Proceedings of the 8th ACM Conference on Electronic Commerce*. ACM, 326–335.
- CHEN, Y., FORTNOW, L., NIKOLOVA, E., AND PENNOCK, D. M. 2007b. Betting on permutations. In *Proceedings of the 8th ACM conference on Electronic commerce*. 326–335.

- CHEN, Y., GOEL, S., AND PENNOCK, D. M. 2008b. Pricing combinatorial markets for tournaments. In *Proceedings of the 40th ACM Symposium on Theory of Computing*.
- CHEN, Y. AND PENNOCK, D. M. 2007. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence*. 49–56.
- CHEN, Y. AND VAUGHAN, J. W. 2010. A new understanding of prediction markets via no-regret learning. In *Proceedings of the 11th ACM Conference on Electronic Commerce*. 189–198.
- COVER, T. 1991. Universal portfolios. *Mathematical Finance* 1, 1, 1–29.
- DAS, S. AND MAGDON-ISMAIL, M. 2008. Adapting to a market shock: Optimal sequential market-making. In *Proceedings of the 21th Annual Conference on Neural Information Processing Systems*. 361–368.
- FORTNOW, L., KILIAN, J., PENNOCK, D. M., AND WELLMAN, M. P. 2004. Betting boolean-style: A framework for trading in securities based on logical formulas. *Decision Support Systems* 39, 1, 87–104.
- FREUND, Y. AND SCHAPIRE, R. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Comp. and System Sciences* 55, 1, 119–139.
- GAO, X., CHEN, Y., AND PENNOCK, D. M. 2009. Betting on the real line. In *Proceedings of the 5th Workshop on Internet and Network Economics*. 553–560.
- GARTHWAITE, P. H., KADANE, J. B., AND O’HAGAN, A. 2005. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association* 100, 680–701.
- GHODSI, M., MAHINI, H., MIRROKNI, V. S., AND ZADIMOGHADDAM, M. 2008. Permutation betting markets: Singleton betting with extra information. In *Proceedings of the 9th ACM conference on Electronic commerce*. 180–189.
- GNEITING, T. AND RAFTERY, A. 2007. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* 102, 477, 359–378.
- GOOD, I. J. 1952. Rational decisions. *Journal of the Royal Statistical Society, Series B (Methodological)* 14, 1, 107–114.
- GORNI, G. 1991. Conjugation and second-order properties of convex functions. *Journal of Mathematical Analysis and Applications* 158, 2, 293–315.
- GRÖTSCHEL, M., LOVÁSZ, L., AND SCHRIJVER, A. 1981. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica* 1, 2, 169–197.
- GUO, M. AND PENNOCK, D. M. 2009. Combinatorial prediction markets for event hierarchies. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*. 201–208.
- HANSON, R. 2003. Combinatorial information market design. *Information Systems Frontiers* 5, 1, 105–119.
- HANSON, R. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets* 1, 1, 3–15.
- HAZAN, E. 2009. A survey: The convex optimization approach to regret minimization. Draft.
- HAZAN, E., AGARWAL, A., AND KALE, S. 2007. Logarithmic regret algorithms for online convex optimization. *Machine Learning* 69, 2–3, 169–192.
- HAZAN, E. AND KALE, S. 2010. Extracting certainty from uncertainty: regret bounded by variation in costs. *Machine Learning* 80, 165–188.
- HELMBOLD, D. AND WARMUTH, M. 2009. Learning permutations with exponential weights. *JMLR* 10, 1705–1736.
- HIRIART-URRUTY, J.-B. AND LEMARÉCHAL, C. 2001. *Fundamentals of Convex Analysis*. Springer.
- KARP, R. 1972. Reducibility among combinatorial problems. In *Complexity of Computer Computations (Symposium Proceedings)*. Plenum Press, 85–103.
- KIVINEN, J. AND WARMUTH, M. 1997. Exponentiated gradient versus gradient descent for linear predictors. *Journal of Information and Computation* 132, 1, 1–63.
- KIVINEN, J. AND WARMUTH, M. K. 1999. Averaging expert predictions. In *Computational Learning Theory: 4th European Conference (EuroCOLT 99)*. Springer, 153–167.
- LAMBERT, N., PENNOCK, D. M., AND SHOHAM, Y. 2008. Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*.
- LEDYARD, J., HANSON, R., AND ISHIKIDA, T. 2009. An experimental test of combinatorial information markets. *Journal of Economic Behavior and Organization* 69, 182–189.
- LITTLESTONE, N. AND WARMUTH, M. 1994. The weighted majority algorithm. *Info. and Computation* 108, 2, 212–261.
- MANGOLD, B., DOOLEY, M., DORNFEST, R., FLAKE, G. W., HOFFMAN, H., KASTURI, T., AND PENNOCK, D. M. 2005. The tech buzz game. *IEEE Computer* 38, 7, 94–97.
- MAS-COLELL, A., WHINSTON, M. D., AND GREEN, J. R. 1995. *Microeconomics Theory*. Oxford University Press, New York, NY.

- MCCARTHY, J. 1956. Measures of the value of information. *PNAS* 42, 654–655.
- MEGIDDO, N. 1977. Mixtures of order matrices and generalized order matrices. *Discrete Mathematics* 19, 2, 177–181.
- OTHMAN, A. AND SANDHOLM, T. 2011. Homogeneous risk measures and liquidity-sensitive automated market makers. In *Proceedings of the 7th Workshop on Internet and Network Economics*. 314–325.
- OTHMAN, A., SANDHOLM, T., PENNOCK, D. M., AND REEVES, D. M. 2010. A practical liquidity-sensitive automated market maker. In *Proceedings of the 11th ACM Conference on Electronic Commerce*. 377–386.
- PENNOCK, D. M. 2004. A dynamic pari-mutuel market for hedging, wagering, and information aggregation. In *Proceedings of the Fifth ACM Conference on Electronic Commerce (EC'04)*.
- PENNOCK, D. M. AND SAMI, R. 2007. Computational aspects of prediction markets. In *Algorithmic Game Theory*, N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, Eds. Cambridge University Press.
- PENNOCK, D. M. AND XIA, L. 2011. Price updating in combinatorial prediction markets with bayesian networks. In *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence*. 581–588.
- PETERS, M., SO, A. M.-C., AND YE, Y. 2007. Pari-mutuel markets: Mechanisms and performance. In *Proceedings of the 3rd International Workshop on Internet and Network Economics*. 82–95.
- RAKHLIN, A. 2009. Lecture notes on online learning. Draft.
- REID, M. D. AND WILLIAMSON, R. C. 2009. Surrogate regret bounds for proper losses. In *ICML*.
- ROCKAFELLAR, R. T. 1970. *Convex analysis*. Princeton Univ Press.
- SAVAGE, L. J. 1971. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association* 66, 336, 783–801.
- SHALEV-SHWARTZ, S. AND SINGER, Y. 2007. A primal-dual perspective of online learning algorithms. *Machine Learning* 69, 2–3, 115–142.
- SION, M. 1958. On general minimax theorems. *Pacific Journal of Mathematics* 8, 1, 171–176.
- WOLFERS, J. AND ZITZEWITZ, E. 2004. Prediction markets. *Journal of Economic Perspective* 18, 2, 107–126.
- XIA, L. AND PENNOCK, D. M. 2011. An efficient monte-carlo algorithm for pricing combinatorial prediction markets for tournaments. In *Proceedings of the International Joint Conferences on Artificial Intelligence*. 305–314.
- ZINKEVICH, M. 2003. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*.

A. CONVEX ANALYSIS RESULTS AND PROOF OF THEOREM 4.2

Towards proving Theorem 4.2, we provide another definition and a couple of results from Rockafellar [1970].

Definition A.1 (Rockafellar [1970], Section 7). A convex function f is said to be *proper* if $f(\mathbf{x}) > -\infty$ for all \mathbf{x} and $f(\mathbf{x}) < +\infty$ for some \mathbf{x} . Also, $f : \mathbb{R}^K \rightarrow [-\infty, \infty]$ is said to be *closed* when the epigraph of f is a closed set, or equivalently, the set $\{\mathbf{x} : f(\mathbf{x}) \leq \alpha\}$ is closed for all $\alpha \in \mathbb{R}$.

THEOREM A.2 (ROCKAFELLAR [1970], THEOREM 12.2 AND COROLLARY 12.2.2). *For any closed convex function $f : \mathbb{R}^K \rightarrow [-\infty, \infty]$, the conjugate f^* is also closed and convex, and $f^{**} = f$. Furthermore, we can write*

$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in \text{relint}(\text{dom}(f))} \mathbf{y} \cdot \mathbf{x} - f(\mathbf{x}) .$$

The preceding theorem tells us that the convex conjugate, which is usually defined in terms of a sup over all of \mathbb{R}^K , can also be written as a sup over just the relative interior of the domain of the function. This is useful for our duality-based framework, as we want to optimize only inside of the convex hull of the payoff vectors.

THEOREM A.3 (ROCKAFELLAR [1970], THEOREM 26.3). *Given a proper closed convex function $f : \mathbb{R}^K \rightarrow [-\infty, \infty]$, f is finite and differentiable everywhere on \mathbb{R}^K if and only if its conjugate f^* is strictly convex on $\text{dom}(f^*)$.*

PROOF OF THEOREM 4.2. We begin with the first part of the Theorem, showing that for any $C : \mathbb{R}^K \rightarrow \mathbb{R}$ satisfying Conditions 2-5, there exists a function R such that Equation 5 is true for any $\mathbf{q} \in \mathbb{R}^K$.

Let $C : \mathbb{R}^K \rightarrow \mathbb{R}$ be some cost function satisfying Conditions 2-5. Theorem 3.2 implies that $\text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}) = \mathcal{H}(\rho(\mathcal{O}))$. It follows also that

$$\text{reint}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}) = \text{reint}(\mathcal{H}(\rho(\mathcal{O}))). \quad (28)$$

Let us now consider the convex conjugate of C ,

$$C^*(\mathbf{x}) := \sup_{\mathbf{q} \in \mathbb{R}^K} \mathbf{x} \cdot \mathbf{q} - C(\mathbf{q}). \quad (29)$$

Recall that we use the notation $\text{dom}(f)$ to refer to the domain of a function f , i.e., where it is defined and finite valued. We can show that

$$\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\} \subseteq \text{dom}(C^*) \subseteq \text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}). \quad (30)$$

For the first containment, it is clear that if we set $\mathbf{x} = \nabla C(\mathbf{q}')$ for any $\mathbf{q}' \in \mathbb{R}^K$ then the supremum in (29) is achieved for $\mathbf{q} = \mathbf{q}'$ and hence $C^*(\nabla C(\mathbf{q}')) = \mathbf{q}' \cdot \nabla C(\mathbf{q}') - C(\mathbf{q}')$. Since C^* is defined on $\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}$, we have $\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\} \subseteq \text{dom}(C^*)$. For the second containment, take some $\mathbf{x} \notin \text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\})$ and consider the derivative of the objective function in (29) with respect to any \mathbf{q} , which is $\mathbf{x} - \nabla C(\mathbf{q})$. This derivative will always have norm bigger than some $\epsilon > 0$ by construction for any \mathbf{q} , and hence the objective must increase without bound. Since \mathbf{x} that does not belong to $\text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\})$ must not in $\text{dom}(C^*)$, we establish $\text{dom}(C^*) \subseteq \text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\})$.

We now show that the choice of $R := C^*$ is strictly convex and satisfies (5). Indeed, strict convexity follows trivially from Theorem A.3. We establish (5) by observing that

$$C(\mathbf{q}) = C^{**}(\mathbf{q}) = \sup_{\mathbf{x} \in \mathbb{R}^K} \mathbf{x} \cdot \mathbf{q} - C^*(\mathbf{x}) = \sup_{\mathbf{x} \in \text{reint}(\text{dom}(C^*))} \mathbf{x} \cdot \mathbf{q} - C^*(\mathbf{x}),$$

where the last equality follows because of Theorem A.2. According to (28) and (30), we also have $\text{reint}(\text{dom}(C^*)) = \text{reint}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}) = \text{reint}(\mathcal{H}(\rho(\mathcal{O})))$ as desired.

We now prove the other direction. Take any strictly convex R defined on $\text{reint}(\mathcal{H}(\rho(\mathcal{O})))$ and let $C(\mathbf{q}) := R^*(\mathbf{q}) = \sup_{\mathbf{x} \in \text{dom}(R)} \mathbf{q} \cdot \mathbf{x} - R(\mathbf{x})$. To establish Conditions 2-5, Theorem 3.2 tells us that it is sufficient to establish three facts: (a) C is defined on all of \mathbb{R}^K , (b) C is everywhere differentiable, and (c) C has the property that $\text{closure}(\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}) = \mathcal{H}(\rho(\mathcal{O}))$. It is easy to establish (a), since for any $\mathbf{q} \in \mathbb{R}^K$, $C(\mathbf{q})$ is defined as a supremum of a concave function on a bounded domain, which always exists. For (b), Theorem A.3 gives us that R being strictly convex implies that C is everywhere differentiable. To prove (c), we note that we already proved that for any differentiable C , the sets $\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\}$ and $\text{dom}(C^*)$ are identical except possibly for points occurring at their respective relative boundaries. Thus, $\text{dom}(C^*) = \text{dom}(R) = \text{reint}(\mathcal{H}(\rho(\mathcal{O})))$, which implies (c). \square

B. PROOF OF LEMMA 4.6

Let $g(t) := D_f(\mathbf{x} + t\mathbf{r}/\|\mathbf{r}\|, \mathbf{x})$. Notice that $g(0) = 0$, and $g'(0) = 0$ since f is differentiable and $D_f(\mathbf{x}, \mathbf{x}')$ is minimized at $\mathbf{x} = \mathbf{x}'$ or, equivalently, $g(t)$ is minimized at $t = 0$. Using the fundamental theorem of calculus, it follows that

$$D_f(\mathbf{x} + \mathbf{r}, \mathbf{x}) = g(\|\mathbf{r}\|) - g(0) = \int_0^{\|\mathbf{r}\|} g'(s) ds = \int_0^{\|\mathbf{r}\|} (g'(s) - g'(0)) ds = \int_0^{\|\mathbf{r}\|} \int_0^s g''(t) dt ds.$$

Because

$$g(t) = D_f(\mathbf{x} + t\mathbf{r}/\|\mathbf{r}\|, \mathbf{x}) = f(\mathbf{x} + t\mathbf{r}/\|\mathbf{r}\|) - f(\mathbf{x}) - \nabla f(\mathbf{x}) \cdot (t\mathbf{r}/\|\mathbf{r}\|),$$

we obtain

$$g'(t) = \nabla f(\mathbf{x} + t\mathbf{r}/\|\mathbf{r}\|) \cdot (\mathbf{r}/\|\mathbf{r}\|) - \nabla f(\mathbf{x}) \cdot (\mathbf{r}/\|\mathbf{r}\|).$$

Taking the derivative of the above expression regarding t , we further have

$$g''(t) = (\mathbf{r}/\|\mathbf{r}\|)^\top \nabla^2 f(\mathbf{x} + t\mathbf{r}/\|\mathbf{r}\|) (\mathbf{r}/\|\mathbf{r}\|).$$

Because the curvature of f at $\mathbf{x} + t\mathbf{r}/\|\mathbf{r}\|$ is lower bounded by the smallest eigenvalue and upper bounded by the largest eigenvalue of $\nabla^2 f(\mathbf{x} + t\mathbf{r}/\|\mathbf{r}\|)$, it must be true that $a \leq g''(t) \leq b$. Thus,

$$\int_0^{\|\mathbf{r}\|} \int_0^s a \, dt \, ds \leq \int_0^{\|\mathbf{r}\|} \int_0^s g''(t) \, dt \, ds \leq \int_0^{\|\mathbf{r}\|} \int_0^s b \, dt \, ds \implies$$

$$\int_0^{\|\mathbf{r}\|} a s \, ds \leq \int_0^{\|\mathbf{r}\|} \int_0^s g''(t) \, dt \, ds \leq \int_0^{\|\mathbf{r}\|} b s \, ds \implies$$

$$\frac{a\|\mathbf{r}\|^2}{2} \leq \int_0^{\|\mathbf{r}\|} \int_0^s g''(t) \, dt \, ds \leq \frac{b\|\mathbf{r}\|^2}{2}.$$

As \mathbf{r} can be chosen arbitrarily as long as \mathbf{x} and $\mathbf{x} + \mathbf{r}$ are both in $\text{dom}(f)$, we establish Inequality 11.