Parimutuel Betting on Permutations*

Shipra Agrawal[1], Zizhuo Wang[2], and Yinyu Ye[3]

[1] Department of Computer Science, Stanford University
shipra@stanford.edu

[2] Department of Management Science and Engineering, Stanford University
zzwang@stanford.edu

[3] Department of Management Science and Engineering, Stanford University
yinyu-ye@stanford.edu

Abstract. We focus on a permutation betting market under parimutuel call auction model where traders bet on final rankings of n candidates. We present a *Proportional Betting* mechanism for this market. Our mechanism allows traders to bet on any subset of the n^2 'candidate-rank' pairs, and rewards them proportionally to the number of pairs that appear in the final outcome. We show that market organizer's decision problem for this mechanism can be formulated as a convex program of polynomial size. Further, the formulation yields a set of n^2 *unique* marginal prices that are sufficient to price the bets in this mechanism, and are computable in polynomial-time. These marginal prices reflect the traders' beliefs about the marginal distributions over outcomes. More importantly, we propose techniques to compute the joint distribution over $n!$ permutations from these marginal distributions. We show that using a maximum entropy criterion, we can obtain a concise parametric form (with only n^2 parameters) for the joint distribution which is defined over an exponentially large state space. We then present an approximation algorithm for computing the parameters of this distribution. In fact, our algorithm addresses a generic problem of finding the maximum entropy distribution over permutations that has a given mean, and is of independent interest.

1 Introduction

Prediction markets are increasingly used as an information aggregation device in academic research and public policy discussions. The fact that traders must "put their money where their mouth is" when they say things via markets helps to collect information. To take full advantage of this feature, however, we should ask markets the questions that would most inform our decisions, and encourage traders to say as many kinds of things as possible, so that a big picture can emerge from many pieces. Combinatorial betting markets hold great promise on this front. Here, the prices of contracts tied to the events have been shown to reflect the traders' belief about the probability of events. Thus, the pricing or ranking of possible outcomes in a combinatorial market is an important research topic.

* Research supported in part by NSF DMS-0604513 and Boeing.

C. Papadimitriou and S. Zhang (Eds.): WINE 2008, LNCS 5385, pp. 126–137, 2008.

We consider a permutation betting scenario where traders submit bids on final rankings of n candidates, for example, an election or a horse race. The possible outcomes are the $n!$ possible orderings among the candidates, and hence there are $2^{n!}$ subset of events to bid on. In order to aggregate information about the probability distribution over the entire outcome space, one would like to allow bets on all these event combinations. However, such betting mechanisms are not only intractable, but also exacerbate the thin market problems by dividing participants attention among an exponential number of outcomes [1][2]. Thus, there is a need for betting languages or mechanisms that could restrict the possible bid types to a tractable subset and at the same time provide substantial information about the traders' beliefs.

1.1 Previous Work

Previous work on parimutuel combinatorial markets can be categorized under two types of mechanisms: a) *posted price mechanisms* including the Logarithmic Market Scoring Rule (LMSR) of Hanson [2][3] and the Dynamic Pari-mutuel Market-Maker (DPM) of Pennock [4] b) *call auction models* developed by Lange and Economides [5], Peters et al. [6], in which all the orders are collected and processed together at once. An extension of the call auction mechanism to a dynamic setting similar to the posted price mechanisms, and a comparison between these models can be found in Peters et al. [7].

Chen et al. (2008) [8] analyze the computational complexity of market maker pricing algorithms for combinatorial prediction markets under LMSR model. They examine both permutation combinatorics, where outcomes are permutations of objects, and Boolean combinatorics, where outcomes are combinations of binary events. Even with severely limited languages, they find that LMSR pricing is #P-hard, even when the same language admits polynomial-time matching without the market maker. Chen, Goel, and Pennock [9] study a special case of Boolean combinatorics and provide a polynomial-time algorithm for LMSR pricing in this setting based on a Bayesian network representation of prices. They also show that LMSR pricing is NP-hard for a more general bidding language.

More closely related to our work are the studies by Fortnow et al. [10] and Chen et al. (2006) [11] on *call auction* combinatorial betting markets. Fortnow et al. [10] study the computational complexity of finding acceptable trades among a set of bids in a Boolean combinatorial market. Chen et al. (2006) [11] analyze the auctioneer's matching problem for betting on permutations, examining two bidding languages: *subset bets*, which are bets of the form candidate i finishes in positions x, y, or z or candidate i, j, or k finishes in position x, and *pair bets*, which take the form candidate i beats candidate j. They give a polynomial-time algorithm for matching divisible subset bets, but show that matching pair bets is NP-hard.

1.2 Our Contribution

In this paper, we focus on the problem of *pricing* a call auction under permutation betting scenario. We consider a new mechanism called *Proportional Betting* for betting on permutations, which is a slightly more generalized form of *Subset Betting* [11], and will be shown to include it as a special case (details in Section 3.2). In proportional betting mechanism, the traders bet on one or more of the n^2 'candidate-position' pairs, and receive rewards proportional to the number of pairs that appear in the final outcome. For example, a trader may place an order of the form "Horse A will finish in position 2 OR Horse B will finish in position 4". He [1] will receive a reward of \$2 if both Horse A & Horse B finish at the specified positions 2 & 4 respectively; and a reward of \$1 if only one horse finishes at the position specified. The market organizer collects all the orders and then decides which orders to accept in order to maximize his worst case profit.

We propose this proportional betting mechanism as a relaxation of *Fixed reward Betting* where a trader receives a fixed reward (say \$1) if *any* of his horse-position pairs appear in the outcome permutation. We show that the market organizer's problem is NP-hard for fixed reward betting. Note that a further relaxation of proportional betting would be to allow traders to bet only on individual candidate position pairs (or individual columns or rows like in subset betting [11]), and allow each trader to submit multiple bets. Here, a difference from our model is that in the relaxed model, a trader may place different bids for different bets and an arbitrary subset of his bets could be accepted, rather than all or nothing.

Our results for proportional betting model are described as follows:

- We show that the market organizer's decision problem for this mechanism can be formulated as a convex program with only $O(n^2 + m)$ variables and constraints, where m is the number of bidders. Further we show that we can obtain, in *polynomial-time*, a small set (n^2) of dual 'marginal prices' that satisfy the desired price consistency constraints, and are sufficient to price the bets in this mechanism. The polynomial-time computability of marginal prices in our call auction setting seems particularly interesting considering that computing the n^2 marginal prices that correspond to Hanson's logarithmic market scoring rule is #P-hard, even under a restricted form of "proportional betting" where traders are allowed to bet only on *individual* candidate-position pairs [8].
- In the second, and perhaps more interesting part of our work, we suggest a maximum entropy criteria to obtain a joint distribution over $n!$ outcomes from the n^2 marginal prices. Although defined over an exponential space, this distribution is shown to have a concise parametric form involving only n^2 parameters. Moreover, it is shown to agree with the maximum-likelihood distribution when prices are interpreted as observed statistics from the traders' beliefs.

  We present an approximation algorithm to compute the parameters of the maximum entropy joint distribution to any given accuracy in (pseudo)-polynomial time [2]. In fact, this algorithm can be directly applied to a generic problem of finding the maximum entropy distribution over permutations that has a given expected value, and is of independent interest.

[1] 'he' shall stand for 'he or she'.

[2] The approximation factors and running time will be established precisely in the text.

To the best of our knowledege, this is the first result on *pricing* a parimutuel call auction under permutation betting scenario.

2 Parimutuel Call Auction Model

In this section, we briefly describe the Convex Parimutuel Call Auction Model (CPCAM) developed by Peters et al. [6] that will form the basis of our betting mechanism. Consider a market with one organizer and m traders or bidders. There are S states of the world in the future on which the traders are submitting bids. For each bid that is accepted by the organizer and contains the realized future state, the organizer will pay the bidder some fixed amount of money, which is assumed to be \$1 without loss of generality. The organizer collects all the bids and decides which bids to accept in order to maximize his worst case profit.

Let $a_{ik} \in \{0,1\}$ denote the trader k's bid for state i. Let q_k and π_k denote the limit quantity and limit price for trader k, i.e., the maximum number of orders requested by trader k, and the maximum price he is willing to pay for the contract, respectively. The number of contracts accepted for trader k is denoted by x_k. x_k is allowed to take fractional values, that is, the orders are 'divisible' in the terminology of [11]. Also, let p_i denote the price computed for outcome state i. Below is the convex formulation of the market organizer's problem given by [6]:

$$
\begin{array}{lll}
\max\limits_{x,s,r} & \pi^T x - r + \sum_{i=1}^{S} \theta_i \log(s_i) & \\
\text{s. t.} & \sum_k a_{ik} x_k + s_i = r & 1 \le i \le S \\
 & 0 \le x \le q & \\
 & s \ge 0 &
\end{array}
\tag{1}
$$

The above convex program maximizes the worst case proft of the organizer which is given by the difference between the total amount of money collected ($\pi^T x$) and the worst case payment made (r). A "parimutuel" state price vector $\{p_i\}_{i=1}^S$ is given by the dual variables associated with the first set of constraints. The parimutuel property implies that if the bidders are charged a price of $\{\sum_i a_{ik} p_i\}$, instead of their limit price, the payouts made to the bidders are exactly funded by the money collected from the accepted orders in the worst-case outcome. $\theta > 0$ represents starting orders needed to guarantee uniqueness of the state price vector. They capture the prior belief of the organizer. The market organizer could actually lose this seed money in some outcomes. However, as shown in [6], infinitesimal quantity of starting orders are sufficient. That is, if we reduce θ uniformly to 0, the price vector converges to a unique limit.

3 Permutation Betting Mechanisms

In this section, we propose new mechanisms for betting on permutations under the parimutuel call auction model described above. Consider a permutation betting scenario with n candidates. Traders bet on rankings of the candidates in the final outcome. The final outcome is represented by an $n \times n$ permutation matrix, where ij^{th} entry of the matrix is 1 if the candidate i takes position j in the final outcome and 0 otherwise. We propose betting mechanisms that restrict the admissible bet types to 'set of candidate-position pairs'. Thus, trader k's bet will be specified by an $n \times n$ $(0, 1)$ matrix A_k, with 1 in the entries corresponding to the candidate-position pairs he is bidding on. We will refer to this matrix as the 'bidding matrix' of the trader. If a trader's bid is accepted, he will receive some payout in the event that his bid is a "winning bid".

Depending on how this payout is determined, two variations of this mechanism are examined: a) Fixed Reward Betting and b) Proportional Betting. The intractability of fixed reward betting will provide motivation to examine proportional betting more closely, which is the focus of this paper.

3.1 Fixed Reward Betting

In this mechanism, a trader receives a fixed payout (assume \$1 w.l.o.g.) if *any* entry in his bidding matrix matches with the corresponding entry in the outcome permutation matrix. That is, if M is the outcome permutation matrix, then the payout made to trader k is given by $I(A_k \bullet M > 0)$. Here, the operator '$\bullet$' denotes the Frobenius inner product[3], and $I(\cdot)$ denotes an indicator function. The market organizer must decide which bids to accept in order to maximize the worst case profit. Using the same notations as in the CPCAM model described in Section 2 for limit price, limit quantities, and accepted orders, the problem for the market organizer in this mechanism can be formulated as follows:

$$\begin{array}{ll} \max & \pi^T x - r \\ \text{s. t.} & r \geq \sum_{k=1}^{m} I(A_k \bullet M_\sigma > 0) x_k \ \forall \sigma \in \mathcal{S}_n \\ & 0 \leq x \leq q \end{array} \tag{2}$$

Here, $\mathcal{S}_n$ represents the set of n dimensional permutations, M_σ represents the permutation matrix corresponding to permutation σ. Note that this formulation encodes the problem of maximizing the worst-case profit of the organizer with no starting orders.

Above is a linear program with exponential number of constraints. We prove the following theorem regarding the complexity of solving this linear program.

Theorem 1. *The optimization problem in (2) is NP-hard even for the case when there are only two non-zero entries in each bidding matrix.*

Proof. The separation problem for the linear program in (2) corresponds to finding the permutation that "satisfies" maximum number of bidders. Here, an

[3] The Frobenius inner product, denoted as $A \bullet B$ in this paper, is the component-wise inner product of two matrices as though they are vectors. That is,

$$A \bullet B = \sum_{i,j} A_{ij} B_{ij}.$$

outcome permutation is said to "satisfy" a bidder, if his bidding matrix has at least one coincident entry with the permutation matrix. We show that the separation problem is NP-hard using a reduction from maximum satisfiability (MAX-2-SAT) problem. Then, using the result on equivalence of separation and optimization problem from [12], the theorem follows. A detailed proof can be found in our technical report [13].

This result motivates us to examine the following variation of this mechanism which makes payouts proportional to the number of winning entries in the bidding matrix.

3.2 Proportional Betting

In this mechanism, the trader receives a fixed payout (assume \$1 w.l.o.g.) *for each coincident entry* between the bidding matrix A_k and the outcome permutation matrix. Thus, the payoff of a trader is given by the Frobenius inner product of his bidding matrix and the outcome permutation matrix. The problem for the market organizer in this mechanism can be formulated as follows:

$$\begin{aligned} &\max \pi^T x - r \\ &\text{s. t. } r \geq \textstyle\sum_{k=1}^{m} (A_k \bullet M_\sigma) x_k \;\forall \sigma \in \mathcal{S}_n \\ &\qquad 0 \leq x \leq q \end{aligned} \tag{3}$$

The above linear program involves exponential number of constraints. However, the separation problem for this program is polynomial-time solvable, since it corresponds to finding the maximum weight matching in a complete bipartite graph, where weights of the edges are given by elements of the matrix $(\sum_k A_k x_k)$. Thus, the ellipsoid method with this separating oracle would give a polynomial-time algorithm for solving this problem. This approach is similar to the algorithm proposed in [11] for *Subset Betting*. Indeed, for the case of subset betting [11], the two mechanisms proposed here (fixed and proportional) are equivalent. This is because subset betting can be equivalently formulated under our framework, as a mechanism that allows non-zero entries only on a single row or column of the bidding matrix A_k. Hence, the number of entries that are coincident with the outcome permutation matrix can be either 0 or 1, resulting in $I(A_k \bullet M_\sigma > 0) = A_k \bullet M_\sigma$, for all permutations σ. Thus, subset betting forms a special case of the proportional betting mechanism proposed here, and all the results derived in the sequel for proportional betting will directly apply to it.

4 Pricing in Proportional Betting

In this section, we reformulate the market organizer's problem for Proportional Betting into a compact linear program involving only $O(n^2+m)$ constraints. The new formulation is not only faster to solve in practice (using interior point methods) but also generates a compact dual price vector of size n^2. These 'marginal prices' will be sufficient to price the bets in Proportional Betting, and are shown to satisfy some useful properties. The reformulation will also allow introducing n^2 starting orders in order to obtain unique prices.

Observe that the first constraint in (3) implicitly sets r as the worst case payoff over all possible permutations (or matchings). Since the matching polytope is integral [12], r can be equivalently set as the result of following linear program that computes maximum weight matching:

$$\begin{array}{rll} r = \max\limits_{M} & (\sum_{k=1}^{m} x_k A_k) \bullet M & \\ \text{s.t.} & M^T e = e & \\ & Me = e & \\ & M_{ij} \geq 0 & i,j \in \{1,\ldots,n\} \end{array} \tag{4}$$

Here e denotes the vector of all 1s (column vector). Taking dual, equivalently,

$$\begin{array}{rll} r = \min\limits_{v,w} & e^T v + e^T w & \\ \text{s.t.} & v_i + w_j \geq \sum_{k=1}^{m} (x_k A_k)_{ij} & i,j \in \{1,\ldots,n\} \end{array} \tag{5}$$

Here, $(x_k A_k)_{ij}$ denotes the ij^{th} element of the matrix $(x_k A_k)$. The market organizer's problem in (3) can now be formulated as:

$$\begin{array}{lll} \max\limits_{x,v,w} & \pi^T x - e^T v - e^T w & \\ \text{s.t.} & v_i + w_j \geq \sum_{k=1}^{m} (x_k A_k)_{ij} & i,j \in \{1,\ldots,n\} \\ & 0 \leq x \leq q & \end{array} \tag{6}$$

Observe that this problem involves only $n^2 + 2m$ constraints.

Let $Q \in \mathbb{R}^{n \times n}$ represent the dual variables corresponding to the first n^2 constraints in the above problem. It is easy to show that the dual matrix Q is well interpreted as a "parimutuel price". That is, $Q \geq 0$; and, if we charge each trader k a price of $A_k \bullet Q$ instead of their limit price (π_k), then the optimal decision remains unchanged and the total premium paid by the accepted orders will be equal to the total payout made in the worst case. Further, Q satisfies the following extended definition of "price consistency condition" introduced in [5].

Definition 1. *The price matrix Q satisfies price consistency constraints if and only if for all k:*

$$\begin{array}{ll} x_k = 0 & \Rightarrow Q \bullet A_k = c_k \geq \pi_k \\ 0 < x_k < q_k & \Rightarrow Q \bullet A_k = c_k = \pi_k \\ x_k = q_k & \Rightarrow Q \bullet A_k = c_k \leq \pi_k \end{array} \tag{7}$$

That is, a trader's bid is accepted only if his limit price is greater than the calculated price for the order.

These properties can be shown using the KKT conditions for (6), in a manner similar to [6] where a non-combinatorial setting is considered. However, the dual price Q thus computed is not guaranteed to be unique. To ensure uniqueness, we can use starting orders as discussed for the CPCAM model in Section 2. We introduce one starting order $\theta_{ij} > 0$ for each candidate-position pair (i,j). These starting orders can be of possibly infinitesimal quantity and represent the prior belief of organizer. Refer [13] for detailed proofs of properties of price matrix Q and the implications of introducing starting orders.

To summarize, we have shown that:

Theorem 2. *One can compute in polynomial-time, an $n \times n$ marginal price matrix Q which is sufficient to price the bets in the Proportional Betting mechanism. Further, the price matrix is unique, parimutuel, and satisfies the desired price-consistency constraints.*

5 Pricing the Outcome Permutations

There is analytical as well as empirical evidence that prediction market prices provide useful estimates of average beliefs about the probability that an event occurs [14][15][16]. Therefore, prices associated with contracts are typically treated as predictions of the probability of future events. The marginal price matrix Q derived in the previous section associates a price to each candidate-position pair. Also, it is easy to observe that Q is a doubly stochastic matrix (use KKT conditions of problem in (6)). Thus, the distributions given by a row (column) of Q could be interpreted as marginal distribution over positions for a given candidate (candidates for a given position). One would like to compute the complete price vector that assigns a price to each of the $n!$ outcome permutations. This price vector would provide information regarding the joint probability distribution over the entire outcome space. In this section, we discuss methods for computing this complete price vector from the marginal prices given by Q.

Let p_σ denote the price for permutation σ. Then, the marginal constraints on the price vector p are represented as:

$$\begin{aligned} \textstyle\sum_{\sigma \in \mathcal{S}_n} p_\sigma M_\sigma &= Q \\ p_\sigma &\geq 0 \ \ \forall \sigma \in \mathcal{S}_n \end{aligned} \tag{8}$$

Finding a feasible solution under these constraints is equivalent to finding a decomposition of doubly-stochastic matrix Q into a convex combination of $n \times n$ permutation matrices. There are multiple such decompositions possible. For example, one such solution can be obtained using Birkhoff-von Neumann decomposition [17][18]. Next, we propose a criterion to choose a meaningful distribution p from the set of distributions satisfying constraints in (8).

5.1 Maximum Entropy Criterion

Intuitively, we would like to use all the information about the marginal distributions that we have, but avoid including any information that we do not have. This intuition is captured by the 'Principle of Maximum Entropy'. It states that the least biased distribution that encodes certain given information is that which maximizes the information entropy. Therefore, we consider the problem of finding the maximum entropy distribution over the space of n dimensional permutations, satisfying the above constraints on the marginal distributions. The problem can be represented as follows:

$$\begin{array}{ll} \min & \sum_{\sigma\in\mathcal{S}_n} p_\sigma \log p_\sigma \\ \text{s.t.} & \sum_{\sigma\in\mathcal{S}_n} p_\sigma M_\sigma = Q \\ & p_\sigma \geq 0 \end{array} \tag{9}$$

The maximum entropy distribution obtained from above has many nice properties. Firstly, as we show next, the distribution has a concise representation in terms of only n^2 parameters. This property is crucial for combinatorial betting due to the exponential state space over which the distribution is defined. Let $Y \in R^{n\times n}$ be the Lagrangian dual variable corresponding to the marginal distribution constraints in (9), and s_σ be the dual variables corresponding to non-negativity constraints on p_σ. Then, the KKT conditions for (9) are given by:

$$\begin{array}{rl} \log(p_\sigma) + 1 - s_\sigma = Y \bullet M_\sigma & \\ \sum_\sigma p_\sigma M_\sigma = Q & \\ s_\sigma, p_\sigma \geq 0 & \forall\sigma \\ p_\sigma s_\sigma = 0 & \forall\sigma \end{array} \tag{10}$$

Assuming $p_\sigma > 0$ for all σ, this gives $p_\sigma = e^{Y\bullet M_\sigma - 1}$. Thus, the distribution is completely specified by the n^2 parameters given by Y. Once Y is known, it is possible to perform operations like computing the probability for a given set of outcome permutations, or finding the most probable outcomes.

Further, we show that the dual solution Y is a maximum likelihood estimator of distribution parameters under suitable interpretation of Q.

Maximum likelihood interpretation. For a fixed set of data and an *assumed* underlying probability model, maximum likelihood estimation method picks the values of the model parameters that make the data "more likely" than any other values of the parameters would make them. Let us assume in our model that the traders' beliefs about the outcome come from an exponential family of distributions D_η, with probability density function of the form $f_\eta \propto e^{\eta\bullet M_\sigma}$ for some parameter $\eta \in R^{n\times n}$. Suppose Q gives a summary statistics of s sample observations $\{M^1, M^2, \ldots, M^s\}$ from the traders' beliefs, i.e., $Q = \frac{1}{s}\sum_k M^k$. This assumption is inline with the interpretation of the prices in prediction markets as mean belief of the traders. Then, the maximum likelihood estimator is given by

$$\begin{aligned} \widehat{\eta} &= \arg\max_\eta \log f_\eta(M^1, M^2, \ldots, M^s) \\ &= \arg\max_\eta \log(\Pi_k \tfrac{e^{\eta\bullet M^k}}{\sum_\sigma e^{\eta\bullet M_\sigma}}) \end{aligned} \tag{11}$$

The optimality conditions for the above unconstrained convex program are:

$$\tfrac{1}{Z}\sum_\sigma e^{\eta\bullet M_\sigma} M_\sigma = \tfrac{1}{s}\sum_k M^k \tag{12}$$

where Z is the normalizing constant, $Z = \sum_\sigma e^{\eta\bullet M_\sigma}$. Since $\frac{1}{s}\sum_k M^k = Q$, observe from the KKT conditions for the maximum entropy model given in (10) that $\eta = Y$ satisfies the above optimality conditions. Hence, the parameter Y computed from the maximum entropy model is also the maximum likelihood estimator for the model parameters η.

5.2 Complexity of the Maximum Entropy Model

In this section, we analyze the complexity of solving the maximum entropy model in (9). As shown in the previous section, the solution to this model is given by the parametric distribution $p_\sigma = e^{Y \bullet M_\sigma - 1}$. The parameters Y are the dual variables given by the optimal solution to the following dual problem of (9)

$$\max_Y Q \bullet Y - \sum_\sigma e^{Y \bullet M_\sigma - 1} \tag{13}$$

We prove the following result regarding the complexity of computing the parameters Y:

Theorem 3. *It is #P-hard to compute the parameters of the maximum entropy distribution $\{p_\sigma\}$ over n dimensional permutations $\sigma \in \mathcal{S}_n$, that has a given marginal distribution.*

Proof. We make a reduction from the following problem:

Permanent of a $(0,1)$ matrix. The permanent of an $n \times n$ matrix B is defined as $\text{perm}(B) = \sum_{\sigma \in \mathcal{S}_n} \Pi_{i=1}^n B_{i,\sigma(i)}$. Computing permanent of a $(0,1)$ matrix is #P-hard [19].

We use the observation that $\sum_\sigma e^{Y \bullet M_\sigma} = \text{perm}(e^Y)$, where the notation e^Y is used to mean component-wise exponentiation: $(e^Y)_{ij} = e^{Y_{ij}}$. For complete proof, see [13].

Interestingly, there exists an FPTAS based on MCMC methods for computing the permanent of any non-negative matrix [20]. Next, we derive a polynomial-time algorithm for approximately computing the parameter Y that uses this FPTAS along with the ellipsoid method for optimization.

5.3 An Approximation Algorithm

Here, we give an outline of the algorithm and present main ideas involved in the analysis. The details along with a complete technical proof can be found in [13].

Using the KKT conditions for the problem, we show that computing optimal Y is equivalent to finding a feasible point in the following bounded convex set:

$$\mathbf{K}: \quad \begin{array}{l} Q \bullet Y - 1 \geq t \\ \sum e^{Y \bullet M_\sigma} M_\sigma \leq Q \\ 0 \geq Y_{ij} \geq -\gamma \qquad \forall i, j \end{array} \tag{14}$$

where $\gamma = \frac{n \log n}{q_{min}}$, $q_{min} = \min\{Q_{ij}\}$, and $t \in [-n \log n - 1, 0]$ is a fixed parameter. Showing this equivalence involves proving upper and lower bounds on optimal Y. Next, we use ellipsoid method to solve this feasibility problem. In each iteration, the ellipsoid method requires to determine if the given iterate Y is feasible, or compute a separating hyperplane, if infeasible. The gradient of a violated constraint forms a natural candidate for separating hyperplane. In the above problem, both these tasks pose a problem due to the intractability of second set of constraints. Checking feasibility requires computing the quantity:

$$f(Y) := \sum_\sigma e^{Y \bullet M_\sigma} M_\sigma$$

And, the gradient takes the form[4]:

$$\nabla f(Y) = \sum_\sigma e^{Y \bullet M_\sigma} (M_\sigma \otimes M_\sigma)$$

Both these quantities are #P-hard to compute. We use MCMC method for computing permanent [20] to compute an $(1+\epsilon)$-approximation of these quantities. For a fixed $\epsilon > 0$, each iteration of the resulting ellipsoid algorithm looks like this:

Algorithm

1. If Y violates any constraints other than the constraint on $f(Y)$, report $Y \notin$ **K**. The violated inequality gives the separating hyperplane.
2. Otherwise, compute a $(1 \pm \delta)$-approximation $\widehat{f}(Y)$ of $f(Y)$, where $\delta = \min\{\frac{\epsilon}{12}, 1\}$.
   (a) If $\widehat{f}(Y) \leq (1+3\delta)Q$, then report Y is feasible.
   (b) Otherwise, say ij^{th} constraint is violated. Compute a $(1 \pm \gamma)$-approximation $\hat{\nabla} f_{ij}(Y)$of the gradient $\nabla f_{ij}(Y)$, where $\gamma = \delta q_{min}/2n^4$. The approximate gradient $C = \hat{\nabla} f_{ij}(Y)$ gives the desired separating hyperplane.

We show that the above algorithm gives an approximate (pseudo-)polynomial time separating oracle for our problem, in the following sense:

Lemma 1. *Given any $Y \in R^{n \times n}$, and any parameter $\epsilon > 0$, the algorithm with runs time polynomial in n, $1/\epsilon$ and $1/\{\min Q_{ij}\}$ and does one of the following:*

- *asserts that $Y \in \mathbf{K}_\epsilon$, where $\mathbf{K}_\epsilon$ represents the set* **K** *with relaxed constraints $f(Y) \leq (1+\epsilon)Q$.*
- *or, finds $C \in R^{n \times n}$ such that $C \bullet X \leq C \bullet Y$ for every $X \in \boldsymbol{K}$.*

Thus, the ellipsoid algorithm using this oracle will terminate with either $Y \in \mathbf{K}_\epsilon$, or declares that there exists no Y in **K**. The proof of the lemma involves proving bounds on the diameter of set **K**, and gradient $\nabla f_{ij}(Y)$. The details are available in [13]. Overall, we prove the following theorem (refer [13] for proof):

Theorem 4. *Using the proposed approximate ellipsoid method, a distribution $\{p_\sigma \sim e^{Y \bullet M_\sigma}\}$ over permutations can be constructed in time poly$(n, \frac{1}{\epsilon}, \frac{1}{q_{min}})$, such that*

- $(1-\epsilon)Q \leq \sum_\sigma p_\sigma M_\sigma \leq Q$
- *p has close to maximum entropy, i.e., $\sum_\sigma p_\sigma \log p_\sigma \leq (1-\epsilon)OPT_E$, where $OPT_E (\leq 0)$ is the optimal value of (9).*

[4] $A \otimes B$ denotes 'Kronecker product' of matrix A and B.

Acknowledgements. We thank Arash Asadpour and Erick Delage for valuable insights and discussions.

References

1. Chen, Y., Fortnow, L., Nikolova, E., Pennock, D.M.: Combinatorial betting. SIGecom Exch. 7(1), 61–64 (2007)
2. Hanson, R.: Combinatorial information market design. Information Systems Frontiers 5(1), 107–119 (2003)
3. Hanson, R.D.: Logarithmic market scoring rules for modular combinatorial information aggregation. Journal of Prediction Markets (2007)
4. Pennock, D.M.: A dynamic pari-mutuel market for hedging, wagering, and information aggregation. In: ACM conference on Electronic commerce (2004)
5. Lange, J., Economides, N.: A parimutuel market microstructure for contingent claims. European Financial Management 11(1) (2005)
6. Peters, M., So, A.M.C., Ye, Y.: A convex parimutuel formulation for contingent claim markets. Working Paper (2005), `http://www.stanford.edu/~yyye/cpcam-ec.pdf`
7. Peters, M., So, A.M.C., Ye, Y.: Pari-mutuel markets: Mechanisms and performance. In: Workshop on Internet and Network Economics (2007)
8. Chen, Y., Fortnow, L., Lambert, N., Pennock, D.M., Wortman, J.: Complexity of combinatorial market makers. CoRR abs/0802.1362 (2008)
9. Chen, Y., Goel, S., Pennock, D.M.: Pricing combinatorial markets for tournaments. In: ACM Symposium on Theory of Computing (2008)
10. Fortnow, L., Kilian, J., Pennock, D.M., Wellman, M.P.: Betting boolean-style: a framework for trading in securities based on logical formulas. In: ACM conference on Electronic commerce (2003)
11. Chen, Y., Fortnow, L., Nikolova, E., Pennock, D.M.: Betting on permutations. ACM conference on Electronic commerce (2007)
12. Grötschel, M., Lovász, L., Schrijver, A.: Geometric Algorithms and Combinatorial Optimization. Springer, Heidelberg (1988)
13. Agrawal, S., Wang, Z., Ye, Y.: Parimutuel betting on permutations. Technical report (2008), `http://arxiv.org/abs/0804.2288`
14. Manski, C.: Interpreting the predictions of prediction markets. Economic Letters 91(3) (2006)
15. Ottaviani, M., Sørensen, P.N.: Aggregation of information and beliefs in prediction markets. Mimeo, London Business School (2006)
16. Wolfers, J., Zitzewitz, E.: Interpreting prediction market prices as probabilities. Working Paper 12200, National Bureau of Economic Research (May 2006), `http://www.nber.org/papers/w12200`
17. Birkhoff, G.: Three observations on linear algebra. Univ. Nac. Tucuman Rev. A 5 (1946)
18. Dulmage, L., Halperin, I.: On a theorem of Frobenius-König and J. von Neumann's game of hide and seek. Trans. Roy. Soc. Canada Sect. III 49 (1955)
19. Valiant, L.: The complexity of computing the permanent. Theoretical Computer Science (1979)
20. Jerrum, M., Sinclair, A., Vigoda, E.: A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. In: ACM symposium on Theory of computing (2001)