Optimal Online Bookmaking for Binary Games

Alankrita Bhatt
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA, USA
Email: abhatt@caltech.edu

Or Ordentlich and Oron Sabag
Rachel and Selim Benin School of Computer Science and engineering,
Hebrew University of Jerusalem, Israel
Email: {or.ordentlich, oron.sabag}@mail.huji.ac.il

***Abstract*—In online betting, the bookmaker can update the payoffs it offers on a particular event many times before the event takes place, and the updated payoffs may depend on the bets accumulated thus far. We study the problem of bookmaking with the goal of maximizing the return in the worst-case, with respect to the gamblers' behavior and the event's outcome. We formalize this problem as the *Optimal Online Bookmaking game*, and provide the exact solution for the binary case. To this end, we develop the optimal bookmaking strategy, which relies on a new technique called bi-balancing trees, that assures that the house loss is the same for all *decisive* betting sequences, where the gambler bets all its money on a single outcome in each round.**

I. Introduction

Consider an experiment I with $m \in \mathbb{N}$ possible outcomes, say $[m] = \{0, 1, \ldots, m-1\}$. Bookmakers offer bets of the following form on the outcome of the experiment: A gambler may invest money in any of the m outcomes. For any 1\$ invested in outcome i, the gambler receives $\gamma(i)$\$ ($\gamma(i) \geq 1$) if the experiment's result is $I = i$, and 0\$ if the experiment's result is not i. Such bets are often referred to as "horse races" in the literature [1], though the underlying experiment need not be a horse race.

Let

$$\Gamma = \sum_{i=0}^{m-1} \frac{1}{\gamma(i)} \tag{1}$$

be the *overround* parameter associated with the offered bet, and

$$r(i) = \frac{1}{\Gamma} \frac{1}{\gamma(i)}, \ i = 0, \ldots, m-1, \tag{2}$$

be the probability distribution the bookmaker "assigns" to the experiment. Assume further that the gambler invests $q(i)$\$ in each outcome $i = 0, \ldots, m-1$, where $\sum_{i=0}^{m-1} q(i) = 1$, such that both $r = (r(0), \ldots, r(m-1))^\top$ and $q = (q(0), \ldots, q(m-1))^\top$ are vectors in the $(m-1)$-dimensional probability simplex $\Delta_m = \left\{p \in [0,1]^m \ : \ \sum_{i=0}^{m-1} p(i) = 1\right\}$. Throughout the paper, the term "gambler" does not necessarily refer to a single individual but rather to the collective group of gamblers placing bets on the experiment's outcome through the bookmaker. From the bookmaker's perspective, the number of participants is irrelevant; only the total amount wagered on each of the m possible outcomes matters.

The bookmaker has collected 1\$ from the gambler before the experiment, and needs to pay the gambler

$$L = \sum_{i=1}^{m-1} \mathbb{1}\{I = i\} q(i)\gamma(i) = \frac{1}{\Gamma} \sum_{i=0}^{m-1} \mathbb{1}\{I = i\} f_i(r, q) \tag{3}$$

dollars after the experiment has taken place, where $f_i(r, q) = \frac{q(i)}{r(i)}$. The worst-case scenario, from the bookmaker's perspective, is that the experiment's outcome is $i^* = \text{argmax}_i \, f_i(r, q)$, and in this case $L = \frac{1}{\Gamma}\|f(r, q)\|_\infty$, where $f(r, q) = (f_0(r, q), \ldots, f_{m-1}(r, q))^\top$. Since $r, q \in \Delta_m$ we clearly have that $\|f(r, q)\|_\infty = \max_i \frac{q(i)}{r(i)} \geq 1$ and this is attained with equality iff $r = q$. Thus, a risk-averse bookmaker, whose objective is maximal gain in the worst-case, should aim to choose r in a way that will cause the gambler to distribute its budget as $q = r$. If it succeeds, the house will collect a gain of $1 - \frac{1}{\Gamma}$ regardless of the experiment's outcome. Thus, when $\Gamma = 1$ the bet is fair, and for $q = r$ both the house and the gambler get zero gain, but when $\Gamma > 1$, the house has a positive gain and the gambler a negative gain, when $r = q$.[1] Note, however, that the bookmaker first declares r and the gambler places its bets according to q only afterwards. It is therefore impossible for the bookmaker to guarantee that $r = q$. On the other hand, if the gambler places its entire bet on the correct outcome of the experiment, say outcome i, the house can gain only if the corresponding offered payoff is smaller than the placed bet, that is, $\gamma(i) \leq 1$. Together with the payoffs requirement $\gamma(i) \geq 1$, it simply means that unless the house payoffs are $\gamma(i) = 1$ (achieved with $\Gamma = m, r(i) = \frac{1}{m}$), the house can lose under particular scenarios.

The game described above consists of a single round: before the experiment the bookmaker declares $\Gamma > 0$ and $r \in \Delta_m$ (which correspond to $\gamma(0), \ldots, \gamma(m-1)$ via (2)) and those cannot be updated until the experiment takes place. The gambler then chooses $q \in \Delta_m$ and distributes its budget accordingly. Today, however, the gambling process is typically much more dynamic. Sports betting is mostly performed through websites, which employ algorithms for updating the odds they offer as the event approaches, and sometimes

[1] There is clearly no motivation for a single gambler to distribute its budget with $q = r$ in this case, as this will result in negative gain with probability 1. However, recall that our "gambler" is composed of many individual gamblers and q represents their combined distribution, so some of them may obtain a positive gain even if $q = r$.

even throughout the event [2]. The algorithms computing the updated odds may rely on the bets accumulated on the event thus far. The update algorithms may further rely on other factors such as new information related to the event [3] (for example, it may be revealed that a certain player is injured), but in this paper we ignore such additional opportunities, and we take a worst-case approach.

We therefore consider the following online setup. Assume the bet happens in T rounds, and let $\Gamma \geq 1$ be a fixed overround parameter that remains constant throughout the T rounds. In each round $t = 1, \ldots, T$, the bookmaker declares $r_t \in \Delta_m$, such that the payoff it offers for each outcome is $\gamma_t(i) = \frac{1}{\Gamma \cdot r_t(i)}$, and the gambler responds by investing a budget of 1\$ distributed on the m possible outcomes according to q_t. The experiment happens at the end of the Tth round. The bookmaker has collected T\$ and needs to pay the gambler

$$L = \frac{1}{\Gamma} \sum_{i=0}^{m-1} \mathbb{1}\{I = i\} \sum_{t=1}^{T} f_i(r_t, q_t) \tag{4}$$

dollars after the experiment have taken place. The question we pursue is:

"What is the largest gain the bookmaker can guarantee regardless of the gambler's behavior and the experiment's outcome?"

We restrict attention to the case of a binary experiment ($m = 2$) and provide an *exact* answer: the largest gain that can be guaranteed is $T\left(1 - \frac{1+T^{-1/2}}{\Gamma}\right)$. Recall that $(1 - \frac{1}{\Gamma})$ is an upper bound for the gain in a single-round attained by the non-causal choice $r = q$. Our result shows that in an online bet with T rounds we can come close to this gain, up to a (normalized) penalty of $T^{-1/2}/\Gamma$. In particular, the bookmaker can guarantee a positive gain whenever the overround parameter satisfies $\Gamma > 1 + T^{-1/2}$. We also provide the precise algorithm for updating r_t based on $q_1, \ldots, q_{t-1}$ which attains at least this gain.

The problem of designing the update policy and computing the optimal gain is a sequential/online optimization problem. Very few problems in this family have been solved *exactly* for any horizon T, see e.g. [4]–[8], and it is quite remarkable that this particular problem does admit an exact solution. Moreover, we find an efficient algorithm to compute the update policy. Furthermore, while the $O(T^{-1/2})$ behavior of the penalty term is common to online optimization problems, we could not find any of-the-shelf online optimization algorithm attaining this rate of convergence for our problem. The reason for this is that the loss functions $f_i(r, q)$ are unbounded. See more in Section II-A.

Paper structure: Section II provides the problem formulation and draws connections to related problems, while Section III presents our main results regarding the optimal bookmaking loss and the algorithms that achieve this loss. Finally, Section IV concludes the paper.

II. Problem Formulation and Related Work

We formalize our problem as a repeated vector-valued game with T-rounds, dubbed *the online bookmaking game*, with two players, the house/bookmaker and the gambler. The game takes place in T rounds, where in each round the house chooses a point in the simplex, which represents the returns it offers for each one of the m outcomes of the event gambled on, and then the gambler chooses a point in the simplex, which represents how it distributes its betting money on the m outcomes. Define:

- **Loss vector**: The loss vector is a mapping $f : \Delta_m \times \Delta_m \to \mathbb{R}_+^m$. In particular,
$$f(r, q) = (f_0(r, q), \ldots, f_{m-1}(r, q))^\top, \tag{5}$$
where
$$f_i(r, q) = \frac{q(i)}{r(i)}, \quad i = 0, \ldots, m-1. \tag{6}$$
- **House Strategy/Algorithm**: A strategy for the first player (the house) is a set of T functions
$$\phi_t : (\Delta_m)^{t-1} \to \Delta_m, \quad t = 1, \ldots, T, \tag{7}$$
such that at the tth round, the house chooses $r_t = \phi_t(q^{t-1})$, where $q^{t-1} = (q_1, \ldots, q_{t-1})$ are the $t-1$ points in the simplex that the gambler chose in the previous rounds.
- **Gambler's Strategy/Algorithm**: A strategy for the second player (the gambler) is a set of T functions
$$\psi_t : (\Delta_m)^t \to \Delta_m, \quad t = 1, \ldots, T, \tag{8}$$
such that at the tth round, the gambler chooses $q_t = \psi_t(r^t)$.
- **Individual accumulated loss vector**: The individual accumulated loss vector of a given house strategy $\{\phi_t\} = \{\phi_t\}_{t=1}^T$ and T points $q^T = (q_1, \ldots, q_T)$ chosen by the gambler is
$$\varphi_T(\{\phi_t\}, q^T) = \sum_{t=1}^{T} f(r_t = \phi_t(q^{t-1}), q_t). \tag{9}$$
- **Individual game loss**: The individual game loss of a given house strategy $\{\phi_t\}$ and the T points q^T chosen by the gambler corresponds to the largest return the gambler can make, and is defined as
$$\|\varphi_T(\{\phi_t\}, q^T)\|_\infty = \max_{i=0,\ldots,m-1} \left\{ \sum_{t=1}^{T} f_i(r_t = \phi_t(q^{t-1}), q_t) \right\}. \tag{10}$$
- **House strategy's loss**: The loss of a given house strategy $\{\phi_t\}$ is
$$L_T(\{\phi_t\}) = \max_{q^T \in (\Delta_m)^T} \|\varphi_T(\{\phi_t\}, q^T)\|_\infty. \tag{11}$$
- **Optimal house loss**: The optimal house loss is
$$L_T^* = \min_{\{\phi_t\}} L_T(\{\phi_t\}) = \min_{\{\phi_t\}} \max_{q^T \in (\Delta_m)^T} \|\varphi_T(\{\phi_t\}, q^T)\|_\infty. \tag{12}$$

When the underlying event has only $m = 2$ possible outcomes, we refer to the game as the *binary online bookmaking game*. This paper considers only the binary case. Note that for the case $m = 2$, the simplex Δ_m corresponds to the interval $[0, 1]$, so the probability vectors $r_t, q_t \in \Delta_m$ may be represented by a single number. With some abuse of notation, in the binary case we therefore use $r_t, q_t \in [0, 1]$ to denote $r_t(1), q_t(1)$, respectively, and the optimal house loss is defined via

$$L_T^* = \min_{\{r_t\}} \max_{q^T \in [0,1]^T} \max \left\{ \sum_{t=1}^{T} \frac{q_t}{r_t(q^{t-1})}, \sum_{t=1}^{T} \frac{1-q_t}{1-r_t(q^{t-1})} \right\} \tag{13}$$

where in the binary case we denote the house strategy at time t, that assigns a number $r_t \in [0, 1]$ to each vector $q^{t-1} \in [0, 1]^T$, by $r_t(q^{t-1})$.

We note that the online bookmaking game is related to the dynamic betting setup described above as follows: the largest gain the bookmaker can guarantee, regardless of the gambler's behavior and of the event's outcome, is $T(1 - \frac{L_T^*}{T\Gamma})$. In particular, if $\Gamma > \frac{1}{T} L_T^*$, the bookmaker can guarantee a positive gain.

Remark 1: Note that the definition of a gambler's strategy $\{\psi_t\}$ is not really needed for defining the house strategy's loss as well as the optimal house loss, as those quantities are computed by taking the worst-case sequence $q^T \in \Delta_m^T$. We nevertheless chose to include the gambler's strategy in the online bookmaking game definition to facilitate the operational interpretation of this game as representing the online gambling procedure described above.

A. Related Works

Competing with fixed strategies: Observe that the (non-causal) fixed strategy

$$r_t = \hat{q} = \frac{1}{T} \sum_{i=1}^{T} q_t, \quad t = 1, \ldots, T \tag{14}$$

attains the individual accumulated loss vector $\varphi_T(\hat{q}, q^T) = (T, \ldots, T)^\top$. Thus, the objective of designing an algorithm $\{\phi_t\}$ that attains house strategy loss $L_T(\{\phi_t\})$ close to T, is identical to that of designing an algorithm that attains minimal regret with respect to the class of all fixed (time-invariant) strategies $r_t = r \ : \ \forall t = 1, \ldots, T$, where r runs through all points in the simplex Δ_m. One can also consider a discretization of Δ_m with M points $r^1, \ldots, r^M \in \Delta_m$ and aim for a minimal regret with respect to only those M "experts". For scalar bounded loss functions , there are many online optimization algorithms that attain regret of $O(\sqrt{T \log M})$ with respect to any class of M experts [9] [10], the multiplicative weights updates (MWU) algorithm being a canonical representative [11]. Note, however, that our online optimization problem is *vector*-valued, and furthermore, the loss functions $f_0(r, q), \ldots, f_{m-1}(r, q)$ are unbounded in $\Delta_m \times \Delta_m$. The problem of unbounded loss may be handled by choosing $0 < \delta_T < 1$ and restricting the class of competing fixed strategies to the region $\Delta_m \cap [\delta_T, 1]^m$, such that for any r in this region and any $q \in \Delta_m$, we have that $\|f(r, q)\|_\infty \leq \frac{1}{\delta_T}$. The value of δ_T should be chosen small enough such that for any point $r \in \Delta_m$ there is a "close enough" allowed point in $\Delta_m \cap [\delta_T, 1]^m$, but large enough so that the upper bound $\frac{1}{\delta_T}$ on $\|f(r, q)\|_\infty$ is not too large. When such a restriction is employed, our problem becomes a vector-valued online optimization problem with bounded loss. While methods such as MWU, designed for scalar online optimization, are not suitable for such a task, the Blackwell approachability framework may be applied.

Blackwell approachability: For a vector-valued game with loss function $f(r, q) \in \mathbb{R}^m$, Blackwell [12] have posed and answered the following question (see [10, Definition 13.3]): *Given a convex region $\mathcal{S} \subset \mathbb{R}^m$, can we find a strategy $\{\phi_t(q^{t-1})\}_{t=1}^T$ such that the Euclidean distance between $\frac{1}{T} \sum_{i=1}^{T} f(r_t = \phi_t(q^{t-1}), q_t)$ and $\mathcal{S}$ vanishes with T, uniformly in q^T?* A set $\mathcal{S}$ for which such a strategy exists is called *approachable*, and Blackwell provided a simple necessary and sufficient condition for approachability, and provided an algorithm whose average loss vector $\frac{1}{T} \sum_{i=1}^{T} f(r_t = \phi_t(q^{t-1}), q_t)$ approaches $\mathcal{S}$ uniformly for all q^T, provided that $\mathcal{S}$ is approachable. Whenever the loss function is bounded, the convergence rate is $O(T^{-1/2})$.

In the online bookmaking game, we are interested in minimizing $\|\frac{1}{T} \sum_{i=1}^{T} f(r_t = \phi_t(q^{t-1}), q_t)\|_\infty$. Denoting the closed ℓ_∞ ball in $\mathbb{R}^m$ by $B_\infty = \{x \in \mathbb{R}^m \ : \ \|x\|_\infty \leq 1\}$, the question of characterizing the optimal house loss L_T^* is equivalent to finding the smallest $\beta > 0$ such that there exists a strategy $\{\phi_t(q^{t-1})\}_{t=1}^T$ for which $\frac{1}{T} \sum_{i=1}^{T} f(r_t = \phi_t(q^{t-1}), q_t) \in \beta \cdot B_\infty$. Thus, in principle, one can use Blackwell's algorithm for designing a "good" house strategy. The problem, however, is that the loss function is not bounded for the online bookmaking game, and one must therefore restrict the strategy space to vectors in $\Delta_m \cap [\delta_T, 1]^m$ as described above.

In the extended version [13], we show that this restricted strategy provides the following upper bound:

Theorem 1 (Th. 5 of [13]): The Blackwell approachability algorithm can be adapted to the online bookmaking problem, achieving in the binary case $\frac{1}{T} L_T^* \leq 1 + O(T^{-1/4})$ for all $q^T \in \Delta_m^T$.

Theorem 1 highlights that, to the best of our understanding, standard tools are insufficient for proving our main result and fail to recover the optimal convergence rate. However, it is worth noting that the Blackwell algorithm does not require knowledge of the time horizon T.

Dynamic programming: The online bookmaking problem can be formulated as a Markov decision process (MDP) with a state defined as the vector of accumulated bets on each team [14]. At each time, the action involves a min-max over r_t, q_t, while at the final time, an additional maximization over the accumulated bets' vector is performed. This formulation ensures that actions based on the defined state are optimal, with the advantage that the (normalized) states and actions lie in time-invariant spaces. However, the state space is continuous, requiring approximation methods, such as grid quantization, to evaluate the optimal gain numerically. Interestingly, as we demonstrate, considering actions based on the entire history rather than just the current state can simplify solutions to certain sequential problems.

Connection to universal compression: The online bookmaking game is closely related to the universal compression problem. In particular, consider the universal compression problem [15, Section IV] of designing a variable-length prefix-free source code $g : [m]^T \to \{0,1\}^*$ such that for any vector $x^T \in [m]^T$ we have that

$$\ell(x^T) - T \cdot H(\hat{q}_{x^T}) = o(T). \tag{15}$$

Here, $\ell(x^T)$ is the length (number of bits) of the codeword $g(x^T)$ representing x^T, $\hat{q}_{x^T}$ is the empirical distribution (normalized histogram) of the sequence x^T, and $H(\cdot)$ is the entropy function (defined with logarithm taken to base 2, such that the entropy is measured in bits). It is well-known that the problem of designing a variable-length prefix-free code for sequences in $[m]^T$ is equivalent (up to constant number of bits) to setting a probability assignment on $[m]^T$ [1], which in turn is equivalent to a sequence of T conditional probability assignments $p_{X_t|X^{t-1}} : [m]^{t-1} \to \Delta_m$, $t = 1, \dots, T$. Any strategy $\{\phi_t\}$ for the online bookmaking game induces a sequence of T conditional probability assignments $r_{X_t|X^{t-1}} : [m]^{t-1} \to \Delta_m$, $t = 1, \dots, T$. To see this, assign to each value $i \in [m]$ the corresponding standard basis vector $\mathbf{e}_i$ (we index the coordinates from 0 to $m-1$, and $\mathbf{e}_i$ is the vector whose ith coordinate equals 1 and all the rest are zero). These vectors are points in Δ_m, so we may define (with some abuse of notation) the conditional probability assignments

$$r_{X_t|X^{t-1}}(\cdot|x^{t-1}) = \phi_t(q_1 = \mathbf{e}_{x_1}, \cdots, q_{t-1} = \mathbf{e}_{x_{t-1}}),\ t \in [T]. \tag{16}$$

In fact, we will see in the next section that the problem of designing an optimal strategy $\{\phi_t\}$ for the online bookmaking game is equivalent to that of designing an optimal strategy for the case where all points $q_1, \dots, q_T$ are chosen from $\{\mathbf{e}_0, \dots, \mathbf{e}_{m-1}\}$. With the probability assignment (16) we can utilize Jensen's inequality to show (see [13])

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{i=0}^{m-1} \mathbb{1}\{x_t = i\} \log_2 \frac{1}{r_{X_t|X^{t-1}}(i|x^{t-1})} \quad -T \cdot H(\hat{q}_{x^T}) \\ &\le T \log_2 \left(\frac{L_T(\{\phi_t\})}{T}\right).\end{aligned} \tag{17}$$

Thus, any strategy $\{\phi_t\}$ for the online bookmaking game with house strategy loss $L_T(\{\phi_t\}) = T + o(T)$ can be translated to a universal compression scheme that attains redundancy of $o(T)$ bits with respect to the class of lossless compressors corresponding to an i.i.d. distribution. As it turns out, the opposite is not true. For instance, using the Shtarkov/Krichevsky-Trofimov probability assignment [8], [15], [16] for $r_{X_t|X^{t-1}}$, which yields redundancy of $O(\log T)$ bits for compression of any $x^T \in \{0,1\}^T$ [17, Chapter 13.5], results in $\sum_{t=1}^{T} f_1(r_t, q_t) = 2T$ for some sequences $q^T \in \{0,1\}^T$. Thus, in this sense, designing strategies for the online bookmaking game is a harder problem than designing schemes for universal compression.

III. Main results

The following theorem summarizes our main results, presenting the optimal house loss and the performance of two algorithms based on whether the gambler is decisive. Specifically, a *decisive gambler* chooses $q^T \in \{0,1\}^T$, that is, at each step it places its entire bet on a single outcome of the experiment.

Theorem 2: The optimal house loss for the binary online bookmaking game is

$$L_T^* = T + \sqrt{T}. \tag{18}$$

Moreover, if the gambler is decisive, Algorithm 1 (Section III-A) defines a house strategy $\{r_t^{\mathtt{ALG}}\} = \{\phi_t^{\mathtt{ALG}}\}$ that can be computed with T simple operations, and achieves the optimal house loss

$$\|\varphi_T(\{r_t^{\mathtt{ALG}}\}, q^T)\|_\infty = L_T^* \tag{19}$$

for all $q^T \in \{0,1\}^T$. If the gambler is non-decisive, i.e., $q^T \in [0,1]^T$, the house strategy $\{\overline{r}_t^{\mathtt{ALG}}\} = \{\overline{\phi}_t^{\mathtt{ALG}}\}$ (defined in Eq. (20)) achieves an individual game loss that can be bounded as

$$\|\varphi_T(\{\overline{r}_t^{\mathtt{ALG}}\}, q^T)\|_\infty \le L_T^*$$

for all $q^T \in [0,1]^T$.

The proof is given in [13]. Recall that L_T^* corresponds to the scenario where both the house and the gambler play optimally, and the event's outcome is chosen to maximize the gambler's gain. Even if the gambler plays optimally, the gambler's gain can be smaller than L_T^* if the event's outcome does not play to its favor.

The main difference in the guarantees of Algorithm 1 and its expected version in (20) for continuous bets stems from the gambler's behavior. We show in [13] that decisive (binary) gamblers maximize the house's loss. Thus, when a bookmaker follows the optimal strategy in Algorithm 1 against an optimal gambler, its loss for the worst outcome is exactly L_T^*. However, if the gamblers are not decisive, the house may benefit from such behavior so that its loss can be decreased. For instance, in the extreme case where the gambler distributes its money according to the offered bet, i.e., $q_t = r_t$, the house's loss can be can be as low as T.

A. Algorithms for Optimal Bookmaking

In this section, we present two algorithms for the online bookmaking problem in which the experiment has two possible outcomes, i.e., the binary online bookmaking problem. We follow the terminology that the experiment's outcome is the winning team in a game between Team 0 and Team 1. The algorithms are sequential and take as input the sequence of bets $q_1 q_2 \dots q_{t-1}$ placed so far, and produce the sequence of odds $r_1 r_2 \dots r_t$ for $t = 1, \dots, T$. Recall that the odds and the bet, at each time, are $r_t \in \Delta_m$ and $q_t \in \Delta_m$, and in the binary case each of them can be described with a scalar, i.e., $r_t \overset{\text{def}}{=} r_t(1) \in [0,1]$ and $q_t \overset{\text{def}}{=} q_t(1) \in [0,1]$. The corresponding payoffs vector that will be published by the bookmaker at time t is $[\gamma_t(0), \gamma_t(1)] = [\frac{1}{\Gamma(1-r_t)}, \frac{1}{\Gamma r_t}]$.

The difference between the two algorithms is based on the gambler's behavior. First, we present the optimal algorithm for the case where the gambler is decisive, meaning their bets satisfy $q_t \in \{0,1\}$. For this case, Theorem 2 asserts that, regardless of the gamblers' bets, q^T, the algorithm achieves the optimal loss L_T^*. We then extend the strategy in Algorithm 1 to handle continuous bets $q_t \in [0,1]$.

B. An Optimal Algorithm for Decisive Gamblers

The algorithm for decisive bets is presented in Algorithm 1. The main idea is to track two numbers, (a,b), which serve as the state of the algorithm and are updated each time a new bet is placed. Initially, the state is set to the optimal loss, $(a,b) = (T + \sqrt{T}, T + \sqrt{T})$.

The state variables (a,b) correspond to the worst-case future losses, considering whether Team 0 or Team 1 will win the game, respectively. As new bets are placed, the state is updated in a way that ensures that the worst-case losses for both teams are minimized. This point is made clearer in [13].

Algorithm 1 Optimal Strategy For Decisive Gamblers

Inputs: T (Total rounds), $q_1 q_2 \dots q_{T-1}$ (Bets are revealed sequentially)
Output: r^T (House strategy)
Initialization: $a = T + \sqrt{T}$, $b = T + \sqrt{T}$
$r_1 \leftarrow \frac{1}{2}$
for $t = 1 : T-1$ **do**
  $d \leftarrow T - t$
  if $q_t = 0$ **then**
    $a \leftarrow d\frac{b-(d-1)}{b-d}$ ▷ Update the future loss for Team 0
  else if $q_t = 1$ **then**
    $b \leftarrow d\frac{a-(d-1)}{a-d}$ ▷ Update the future loss for Team 1
  end if
  $b^+ \leftarrow (d-1)\frac{a-(d-2)}{a-(d-1)}$ ▷ Hypothetical future loss for Team 1
  $r_{t+1} \leftarrow \frac{1}{b-b^+}$
end for

The strategy $r_{t+1}(q^t)$ only depends on q^t, and we denote the collection mappings produced by Algorithm 1 as $\{\phi_t^{\text{ALG}}\} = \{r_t^{\text{ALG}}\}$. The algorithm's implementation requires tracking two real numbers and at each cycle only a simple update is executed.

C. Algorithm for Non-Decisive Gamblers (Continuous Bets)

In practice, the gambler is not a single entity. Rather, it represents the accumulation of bets made by several gamblers, that may distribute their bets over the two teams. We present a modification of Algorithm 1 to handle continuous bets $q_t \in [0,1]$.

The algorithm for continuous bets $q_t \in [0,1]$ is based on the optimal policy for binary bets, generated by Algorithm 1. In particular, let $r_t^{\text{ALG}}(x^{t-1})$ denote the odds produced by Algorithm 1 for some input $x^{t-1} \in \{0,1\}^t$. The idea here is to view $q_t \in [0,1]$ as the expected value of a binary random variable $X_t \sim \text{Ber}(q_t)$. The strategy for continuous bets is

$$\overline{r}_t^{\text{ALG}}(q^{t-1}) = \mathbb{E}\left[r_t^{\text{ALG}}(X^{t-1})\right], \tag{20}$$

where the expected value is taken with respect to the sequence of independent random variables $X_i \sim \text{Ber}(q_i)$. Note that (20) defines a deterministic, instantaneous mapping, and we denote the sequence of mappings by $\{\overline{r}_t^{\text{ALG}}\}$. Using the fact that $q \mapsto f_i(r,q)$ is linear and $r \mapsto f_i(r,q)$ is convex, we show in [13] that the loss for the strategy $\{\overline{r}_t^{\text{ALG}}(q^{t-1})\}$ from (20) is maximized by q^T on the boundary of $[0,1]^T$, that is, on $\{0,1\}^T$.

Algorithm 1 involves carrying out simple operations per round to compute the action $r_t^{\text{ALG}}(q^{t-1})$ when the gamblers are decisive, leading to an efficient total runtime of $\mathcal{O}(T)$. Unfortunately, this efficiency is lost when computing the odds for non-decisive gamblers. In this case, the odds at time t are computed via (20) as

$$\overline{r}_t^{\text{ALG}}(q^{t-1}) = \sum_{x^{t-1}\in\{0,1\}^{t-1}} r_t^{\text{ALG}}(x^{t-1}) \prod_{i=1}^{t-1} q_i^{x_i}(1-q_i)^{1-x_i}. \tag{21}$$

Therefore, calculating the odds when the gamblers place continuous bets involves a summation over 2^{t-1} terms which is computationally very expensive. In [13], we provide a low-complexity algorithm to approximately calculate $\overline{r}_t^{\text{ALG}}(q^{t-1})$ and quantify this algorithm's approximation error. This algorithm approximates the expectation operation in (20) via a Monte Carlo simulation.

IV. Conclusions

We have introduced the online bookmaking game, where the house/bookmaker updates the odds it offers based on the bets accumulated so far, with the goal of maximizing its worst-case return. We have shown that the optimal strategy of the gambler is decisive, meaning that in each round the gambler places all its bet on a single outcome. Consequently, the problem of designing a general optimal house strategy reduces to designing an optimal strategy for decisive gamblers. For the binary case, we have shown that the latter problem is a special case of the tree bi-balancing problem, that we introduce and solve. Consequently, we obtain the complete solution to the binary online bookmaking problem, and develop the optimal algorithm for updating the odds offered by the house.

Acknowledgment

The work of AB was supported by the Center for the Mathematics of Information Postdoctoral Fellowship at Caltech. The work of OO was supported by the Israel Science Foundation (ISF), grant No. 1641/21. The work of OS was supported by the Israel Science Foundation (ISF), grant No. 1096/23.

References

[1] T. Cover and J. A. Thomas, *Elements of information theory*. Hoboken, NJ: Wiley-Interscience, 2006.

[2] H. Lopez-Gonzalez and M. D. Griffiths, “Understanding the convergence of markets in online sports betting,” *International Review for the Sociology of Sport*, vol. 53, no. 7, pp. 807–823, 2018.

[3] H. Zhu, A. Soen, Y. K. Cheung, and L. Xie, “Online learning in betting markets: profit versus prediction,” in *Proceedings of the 41st International Conference on Machine Learning (ICML)*. PMLR 235, 2024.

[4] T. Cover, “Behaviour of sequential predictors of binary sequences,” in *Proc. 4th Prague Conf. Inform. Theory, Statistical Decision, Functions, Random Processes*, 1965.

[5] A. Rakhlin and K. Sridharan, “A tutorial on online supervised learning with applications to node classification in social networks,” *arXiv preprint arXiv:1608.09014*, 2016.

[6] N. Gravin, Y. Peres, and B. Sivan, “Towards optimal algorithms for prediction with expert advice,” in *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 2016, pp. 528–547.

[7] T. M. Cover, “Universal portfolios,” *Mathematical finance*, vol. 1, no. 1, pp. 1–29, 1991.

[8] Y. M. Shtar’kov, “Universal sequential coding of single messages,” *Problemy Peredachi Informatsii*, vol. 23, no. 3, pp. 3–17, 1987.

[9] N. Cesa-Bianchi, Y. Freund, D. Haussler, D. P. Helmbold, R. E. Schapire, and M. K. Warmuth, “How to use expert advice,” *Journal of the ACM (JACM)*, vol. 44, no. 3, pp. 427–485, 1997.

[10] E. Hazan, “Introduction to online convex optimization,” 2023. [Online]. Available: https://arxiv.org/abs/1909.05207

[11] S. Arora, E. Hazan, and S. Kale, “The multiplicative weights update method: a meta-algorithm and applications,” *Theory of computing*, vol. 8, no. 1, pp. 121–164, 2012.

[12] D. Blackwell, “An analog of the minimax theorem for vector payoffs.” 1956.

[13] A. Bhatt, O. Ordentlich, and O. Sabag, “Optimal online bookmaking for binary games,” 2025. [Online]. Available: https://arxiv.org/abs/2501.06923

[14] D. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 2001, vol. 1 and 2.

[15] N. Merhav and M. Feder, “Universal prediction,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2124–2147, 1998.

[16] R. Krichevsky and V. Trofimov, “The performance of universal encoding,” *IEEE Transactions on Information Theory*, vol. 27, no. 2, pp. 199–207, 1981.

[17] Y. Polyanskiy and Y. Wu, “Information theory: From coding to learning,” *Book draft*, 2022.