

Optimal Online Bookmaking for Binary Games

Alankrita Bhatt, Or Ordentlich and Oron Sabag

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 AND MAIN RESULTS

Consider an experiment I with $m \in \mathbb{N}$ possible outcomes, say $[m] = \{0, 1, \dots, m-1\}$. 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

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

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

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

be the probability distribution the bookmaker "assigns" to the experiment. Assume further that the gambler invests $q(i)$ \$ in each outcome $i = 0, \dots, m-1$, where $\sum_{i=0}^{m-1} q(i) = 1$, such that both $r = (r(0), \dots, r(m-1))^\top$ and $q = (q(0), \dots, q(m-1))^\top$ are vectors in the $(m-1)$ -dimensional probability simplex

$$\Delta_m = \left\{ p \in [0, 1]^m : \sum_{i=0}^{m-1} p(i) = 1 \right\}. \quad (3)$$

Throughout the paper, the word "gambler" need not refer to a single gambler, but to the accumulation of all gamblers placing their bets on the experiment's outcome through the bookmaker. From the bookmaker's perspective, the number of gamblers participating in the bet makes no difference, and it is only the total amount invested in each of the m possible outcomes that matters.

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

$$L = \sum_{i=1}^{m-1} \mathbb{1}\{I = i\} q(i) \gamma(i) = \frac{1}{\Gamma} \sum_{i=0}^{m-1} \mathbb{1}\{I = i\} f_i(r, q) \quad (4)$$

dollars after the experiment has taken place, where $f_i(r, q) = \frac{q(i)}{r(i)}$, $i = 0, \dots, m-1$. The worst-case scenario, from the bookmaker's perspective, is that the experiment's outcome is $i^* = \operatorname{argmax}_i f_i(r, q)$, and in this case $L = \frac{1}{\Gamma} \|f(r, q)\|_\infty$, where $f(r, q) = (f_0(r, q), \dots, f_{m-1}(r, q))^\top$. 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,

A. Bhatt is with the California Institute of Technology (abhhatt@caltech.edu). O. Ordentlich and O. Sabag are with the Rachel and Selim Benin School of Computer Science and engineering, Hebrew University of Jerusalem, Israel (or.ordentlich@mail.huji.ac.il, oron.sabag@mail.huji.ac.il). 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.

¹There 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 $q = r$.

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 experiment's outcome, the house can gain only if its return is smaller than the placed bet, that is, $\gamma(i) \leq 1$. Together with the odds requirement $\gamma(i) \geq 1$, it simply means that unless the house odds are $\gamma(i) = 1$ (achieved with $\Gamma = m, \gamma(i) = \frac{1}{m}$), 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 $\gamma(0), \dots, \gamma(m-1)$ 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 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 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 $t = 1, \dots, T$, the bookmaker declares $r_t \in \Delta_m$, such that the return 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 T th round. The bookmaker has collected T \$ and needs to pay the gambler

$$L = \frac{1}{\Gamma} \sum_{i=0}^{m-1} \mathbb{1}\{I = i\} \sum_{t=1}^T f_i(r_t, q_t) \quad (5)$$

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 the gain in a single-round for 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, \dots, q_{t-1} which attains this return, as well as an efficient algorithm that approximately attains this performance.

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], [5], [6], [7], [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 I-B.

A. Mathematical formulation of the online bookmaking game

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 $f : \Delta_m \times \Delta_m \rightarrow \mathbb{R}_+^m$. In particular,

$$f(r, q) = (f_0(r, q), \dots, f_{m-1}(r, q))^T, \quad (6)$$

where

$$f_i(r, q) = \frac{q(i)}{r(i)}, \quad i = 0, \dots, m-1. \quad (7)$$

- **House Strategy/Algorithm:** A strategy for the first player (the house) is a set of T functions

$$\phi_t : (\Delta_m)^{t-1} \rightarrow \Delta_m, \quad t = 1, \dots, T, \quad (8)$$

such that at the t th round, the house chooses $r_t = \phi_t(q^{t-1})$, where $q^{t-1} = (q_1, \dots, q_{t-1})$ 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

$$\psi_t : (\Delta_m)^t \rightarrow \Delta_m, \quad t = 1, \dots, T, \quad (9)$$

such that at the t th 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 $q^T = (q_1, \dots, q_T)$ chosen by the gambler is

$$\varphi_T(\{\phi_t\}, q^T) = \sum_{t=1}^T f(r_t = \phi_t(q^{t-1}), q_t). \quad (10)$$

- **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

$$\|\varphi_T(\{\phi_t\}, q^T)\|_\infty = \max_{i=0, \dots, m-1} \left\{ \sum_{t=1}^T f_i(r_t = \phi_t(q^{t-1}), q_t) \right\}. \quad (11)$$

- **House strategy's loss:** The loss of a given house strategy $\{\phi_t\}$ is

$$L_T(\{\phi_t\}) = \max_{q^T \in (\Delta_m)^T} \|\varphi_T(\{\phi_t\}, q^T)\|_\infty. \quad (12)$$

- **Optimal house loss:** The optimal house loss is

$$L_T^* = \min_{\{\phi_t\}} L_T(\{\phi_t\}) = \min_{\{\phi_t\}} \max_{q^T \in (\Delta_m)^T} \|\varphi_T(\{\phi_t\}, q^T)\|_\infty. \quad (13)$$

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

$$L_T^* = \min_{\{r_t\}} \max_{q^T \in [0, 1]^T} \max \left\{ \sum_{t=1}^T \frac{q_t}{r_t(q^{t-1})}, \sum_{t=1}^T \frac{1 - q_t}{1 - r_t(q^{t-1})} \right\} \quad (14)$$

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 $T(1 - \frac{L_T^*}{T})$. 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.

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

$$L_T^* = T + \sqrt{T}. \quad (15)$$

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

$$\|\varphi_T(\{r_t^{\text{ALG}}\}, q^T)\|_\infty = L_T^* \quad (16)$$

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 $\{\bar{r}_t^{\text{ALG}}\} = \{\bar{\phi}_t^{\text{ALG}}\}$ (defined in Eq. (25)) achieves an individual game loss that can be bounded as

$$\|\varphi_T(\{\bar{r}_t^{\text{ALG}}\}, q^T)\|_\infty \leq L_T^*$$

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

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 (25) for continuous bets stems from the gambler's behavior. We show in Section III 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 return 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 as low as T .

B. Related Work and Connections to Other Problems

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

$$r_t = \hat{q} = \frac{1}{T} \sum_{i=1}^T q_t, \quad t = 1, \dots, T \quad (17)$$

attains the individual accumulated loss vector $\varphi_T(\hat{q}, q^T) = (T, \dots, T)^\top$. 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 $r_t = r : \forall t = 1, \dots, T$, where r runs through all points in the simplex Δ_m . One can also consider a discretization of Δ_m with M points $r^1, \dots, r^M \in \Delta_m$ 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 $f_0(r, q), \dots, f_{m-1}(r, q)$ 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 ?* Here, a distance between a point $x \in \mathbb{R}^m$ and a set $\mathcal{S} \subset \mathbb{R}^m$ is defined as $\min_{s \in \mathcal{S}} \|x - s\|$. 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 Section V we derive an algorithm which follows Blackwell’s approach (taking into account the fact that $f(r, q)$ is unbounded). Such an algorithm only attains $\|\frac{1}{T} \sum_{i=1}^T f(r_t = \phi_t(q^{t-1}), q_t)\|_\infty = 1 + O(T^{-1/4})$ for all $q^T \in (\Delta_m)^T$. This demonstrates that, to the best of our understanding, standard tools are insufficient for proving our main result, Theorem 1, and do not even recover the optimal convergence rate.

We note in passing that the objective in online bookmaking game is somewhat reminiscent to the objective in *calibration* [13], [14], [15] which has been shown to be reducible to Blackwell approachability [16], [17], [18], as well as weak calibration [19], [20]. In particular, in the context of weak calibration, defining the two functions $w_1(r) = \frac{1}{r}$ and $w_2(r) = -\frac{1}{1-r}$, we see that a good strategy for the binary online bookmaking game guarantees that $\frac{1}{T} \sum_{t=1}^T w_i(r_t)(q_t - r_t)$ vanishes for both $i = 1, 2$. However, since these functions are not Lipschitz, an off-the-shelf strategy for weak calibration does not seem to imply a good strategy for the online bookmaking game, and vice versa.

Dynamic programming: The online bookmaking problem can be formulated as a Markov decision process (MDP) whose state is the vector that contains the accumulated bets on each team [21]. The action at each time includes a min-max over r_t, q_t , respectively, while the action in the last time has an additional maximum over the accumulated bets’ vector. It can be shown that this is a valid MDP so that actions that depend on the defined state are optimal. The defined MDP has the advantage that the actions and (normalized) states take values in time-invariant spaces. However, the MDP suffers from the fact that those spaces are continuous and therefore even to evaluate the optimal gain numerically we need to use some approximation method, e.g., grid quantization over the states and actions. Interestingly, as we demonstrate, actions that depend on the entire history of actions (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 [22, Section IV] of designing a variable-length prefix-free source code $g : [m]^T \rightarrow \{0, 1\}^*$ such that for any vector $x^T \in [m]^T$ we have that

$$\ell(x^T) - T \cdot H(\hat{q}_{x^T}) = o(T). \quad (18)$$

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 $p_{X_t|X^{t-1}} : [m]^{t-1} \rightarrow \Delta_m$, $t = 1, \dots, T$. In particular, we may use arithmetic coding with such conditional probability assignments and obtain [1]

$$\ell(x^t) \leq 2 + \sum_{t=1}^T \sum_{i=0}^{m-1} \mathbb{1}\{x_t = i\} \log_2 \frac{1}{p_{X_t|X^{t-1}}(i|x^{t-1})}. \quad (19)$$

Any strategy $\{\phi_t\}$ for the online bookmaking game induces a sequence of T conditional probability assignments $r_{X_t|X^{t-1}} : [m]^{t-1} \rightarrow \Delta_m$, $t = 1, \dots, T$. 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 i th 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

$$r_{X_t|X^{t-1}}(\cdot|x^{t-1}) = \phi_t(q_1 = \mathbf{e}_{x_1}, \dots, q_{t-1} = \mathbf{e}_{x_{t-1}}), \quad t = 1, \dots, T. \quad (20)$$

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 (20) we have

$$\sum_{t=1}^T \sum_{i=0}^{m-1} \mathbb{1}\{x_t = i\} \log_2 \frac{1}{r_{X_t|X^{t-1}}(i|x^{t-1})} - T \cdot H(\hat{q}_{x^T}) = \sum_{t=1}^T \sum_{i=0}^{m-1} \mathbb{1}\{x_t = i\} \log_2 \frac{\hat{q}_{x^T}(i)}{r_{X_t|X^{t-1}}(i|x^{t-1})} \quad (21)$$

$$\begin{aligned} &= \sum_{t=1}^T \sum_{i=0}^{m-1} \mathbf{e}_{x_t}(i) \log_2 \frac{\hat{q}_{x^T}(i)}{r_{X_t|X^{t-1}}(i|x^{t-1})} \\ &= \sum_{t=1}^T \log_2 \left(\sum_{i=0}^{m-1} \mathbf{e}_{x_t}(i) \frac{\hat{q}_{x^T}(i)}{r_{X_t|X^{t-1}}(i|x^{t-1})} \right) \end{aligned} \quad (22)$$

$$\leq T \log_2 \left(\sum_{i=0}^{m-1} \frac{1}{T} \sum_{t=1}^T \mathbf{e}_{x_t}(i) \frac{\hat{q}_{x^T}(i)}{r_{X_t|X^{t-1}}(i|x^{t-1})} \right) \quad (23)$$

$$\begin{aligned} &= T \log_2 \left(\sum_{i=0}^{m-1} \hat{q}_{x^T}(i) \cdot \frac{1}{T} \sum_{t=1}^T \frac{\mathbf{e}_{x_t}(i)}{r_{X_t|X^{t-1}}(i|x^{t-1})} \right) \\ &= T \log_2 \left(\sum_{i=0}^{m-1} \hat{q}_{x^T}(i) \cdot \frac{1}{T} \sum_{t=1}^T f_i(r_t = \phi_t(\mathbf{e}_{x_1}, \dots, \mathbf{e}_{x_{t-1}}, \mathbf{e}_{x_t})) \right) \\ &\leq T \log_2 \left(\frac{L_T(\{\phi_t\})}{T} \right), \end{aligned} \quad (24)$$

where (21) follows since $\hat{q}_{x^T}(i) = \frac{1}{T} \sum_{t=1}^T \mathbb{1}\{x_t = i\}$, (23) follows from Jensen's inequality and concavity of $y \mapsto \log_2(y)$, and (24) follows from the definition of $L_T(\{\phi_t\})$ in (12). 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.

The opposite is not true. Assume the gambler is restricted to a *decisive* strategy, where in each round it invests all its budget in a single outcome. This is equivalent to restricting the gambler to using $q^T \in \{\mathbf{e}_0, \dots, \mathbf{e}_{m-1}\}^T$ rather than $q^T \in \Delta_m^T$. A house strategy for such a gambler is a probability assignment $r_t(x^{t-1}) = r_{X_t|X^{t-1}}(\cdot|x^{t-1})$ where x_t is the outcome the gambler chose to bet on in round t . Probability assignments that give excellent redundancy for the universal compression problem may be very bad for the online bookmaking game. To see this, consider the binary case $m = 2$ with the Shtarkov/Krichevsky-Trofimov probability assignment [23], [8], [22] $p_{X_t|X^{t-1}}(i|x^{t-1}) = \frac{1/2 + \sum_{j=1}^{t-1} \mathbb{1}\{x_j = i\}}{t}$. While for universal compression this probability assignment yields redundancy of $O(\log T)$ bits for any $x^T \in \{0, 1\}^T$ [24, Chapter 13.5], for a gambling sequence $x^T = 0^{T-1}1$, this probability assignment gives

$$\sum_{t=1}^T f_1(r_t, q_t) = \frac{1}{p_{X_T|X^{T-1}}(1|0^{T-1})} = 2T.$$

To summarize, any asymptotically good house strategy for the online bookmaking game provides an asymptotically good universal compression scheme, but asymptotically good universal compression schemes do not necessarily lead to asymptotically good house strategies for the online bookmaking game. Thus, in this sense, designing strategies for the online bookmaking game is a harder problem than designing schemes for universal compression.

C. Paper Structure

In Section II we present our optimal algorithm for decisive gamblers, and the algorithm it induces for general (non-decisive) gamblers. Section III provides the optimality proof of these algorithms, along with several observations on the structure of an optimal algorithm. While our optimal algorithm for decisive gamblers runs in linear time (in the horizon T), a brute-force implementation of the general algorithm it induces requires $O(2^T)$ operations. In Section IV we develop a polynomial-time algorithm that approximates the optimal one and attains loss arbitrarily close to that of the optimal one. In Section V we develop a sub-optimal algorithm for the online bookmaking problem, based on the standard Blackwell-approachability approach. The loss of this algorithm is proved to be $T + O(T^{3/4})$, which is significantly larger than L_T^* . We nevertheless provide this algorithm for two reasons: 1) It demonstrates that standard approaches seem to be insufficient for attaining loss close to L_T^* ; 2) While it is sub-optimal, it has one advantage over the algorithm from Section II: It does not depend on the horizon. We conclude and list several directions for future work in Section VI.

II. 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., $r_t \stackrel{\text{def}}{=} r_t(1) \in [0, 1]$ and $q_t \stackrel{\text{def}}{=} q_t(1) \in [0, 1]$. The corresponding pay-offs 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 1 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 $q_t \in [0, 1]$.

A. 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. This point will be made clearer in the sequel. As new bets are placed, the state is updated using a specific formula that ensures that the worst-case losses for both teams are minimized. The derivation of the update formula and its optimality are provided in Theorem 4.

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.

B. 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

$$\bar{r}_t^{\text{ALG}}(q^{t-1}) = \mathbb{E} [r_t^{\text{ALG}}(X^{t-1})], \quad (25)$$

where the expected value is taken with respect to the sequence of independent random variables $X_i \sim \text{Ber}(q_i)$. Note that (25) defines a deterministic, instantaneous mapping, and we denote the sequence of mappings by $\{\bar{r}_t^{\text{ALG}}\}$.

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 (25) as

$$\bar{r}_t^{\text{ALG}}(q^{t-1}) = \sum_{x^{t-1} \in \{0, 1\}^{t-1}} r_t^{\text{ALG}}(x^{t-1}) \prod_{i=1}^{t-1} q_i^{x_i} (1 - q_i)^{1-x_i}. \quad (26)$$

Therefore, calculating the odds when the gamblers place continuous bets involves a summation over 2^{t-1} terms which is computationally very expensive. In Section IV, we provide a low-complexity algorithm to approximately calculate $\bar{r}_t^{\text{ALG}}(q^{t-1})$ and quantify this algorithm's approximation error. This algorithm approximates the expectation operation in (25) via a Monte Carlo simulation.

III. PROOF OF MAIN RESULTS

In this section, we prove our main results in Theorem 1: we derive the optimal house loss and show that Algorithm 1 is optimal for decisive gamblers, as well as the performance of its extension when generalizing to continuous gamblers. Our technical results are more general than required for the online bookmaking setup, so we will present them in their full generality and prove Theorem 1 at the end of this section.

A. Decisive Gamblers are Optimal

We first show that the house strategy's loss is achieved with gamblers that are decisive, that is, choosing a single team to bet on at each round is optimal from the gambler's perspective. In other words, if the gambler aims to maximize the house's loss it does not need to split its money among the teams and can place its bet on a single team in each round. It turns out that this claim holds not only for the online bookmaking problem, but for more general loss functions, defined as follows.

Problem 1 (Convex-Linear-Max Games): We consider the repeated game with two players as defined in Section I-A, but allow more general loss functions than the ones in (7). Specifically, each loss function is taken from a finite collection $f_i(r, q) = q(i)g_i(r)$, where $g_i : \mathbb{R}^m \rightarrow \mathbb{R}$ is a convex function and $i \in [m]$. The optimal loss is defined as before,

$$L_T^* = \min_{\{\phi_t\}} \max_{q^T \in \Delta_m^T} \|\varphi_T(\{\phi_t\}, q^T)\|_\infty. \quad (27)$$

The online bookmaking problem is a special case of Problem 1 by choosing $g_i(r) = \frac{1}{r(i)}$. The optimal loss in (27) is difficult to compute directly for several reasons such as the continuous domain of the players' strategies, the dependence of the mappings' domain on time, the inner maximization (the ℓ^∞ norm) etc. Our first question is whether the strategies' structure can be simplified for Problem 1.

One suggestion is to formulate the repeated game as a Markov decision process (MDP) whose state at time t is the vector of cumulative past losses $s_t(i) = \sum_{j=1}^{t-1} f_i(r_j, q_j)$ for $i = 1, \dots, m$. The MDP formulation benefits the fact that the actions r_t and q_t should not depend on the entire history of actions, r^{t-1}, q^{t-1} , whose lengths grow with time. Instead, they only depend on the vector s_t . However, the challenge in solving this MDP is that the state is continuous, and that the overall loss function cannot be expressed as a sum of immediate rewards due to the infinity norm. Intuitively, the inner maximization in (27) can be interpreted as being performed at the end of the game, as it depends on the vectors r^T and q^T . Consequently, the players do not know during the game which specific function will ultimately be selected at the end.

We take a different approach in which we aim to restrict the strategies' domains and images. We prove that the image of the second player's strategy, defined by ψ_t , can be restricted to the simplex vertices, i.e., the basis vectors $\{\mathbf{e}_i\}_{i=1}^m$. Formally, define the set of restricted strategies as

$$\psi_t^{\text{D}} : (\Delta_m)^t \rightarrow \{\mathbf{e}_1, \dots, \mathbf{e}_m\} \quad (28)$$

such that $q_t = \psi_t^{\text{D}}(r^t)$. We call a player whose image is restricted as (28) a *decisive player* as it allocates all its mass to a single coordinate of the vector q_t . If the second player is decisive, it follows directly that the domain of the first player's strategy can be restricted as

$$\phi_t^{\text{D}} : \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^{t-1} \rightarrow \Delta_m \quad (29)$$

with $r_t = \phi_t^{\text{D}}(q^{t-1})$ since values outside of this domain will not appear in q^{t-1} . We call a strategy of the form (29) a *skeleton strategy*.

As mentioned, it will be shown that decisive players are optimal so that skeleton strategies suffice to achieve the optimal loss. However, one may be interested in algorithms for non-decisive players, i.e., $q^T \in \Delta_m^T$. For this purpose, we can take any skeleton strategy, and define its expected skeleton strategy as

$$\bar{\phi}_t(q^{t-1}) = \mathbb{E}[\phi_t^{\text{D}}(X^{t-1})], \quad (30)$$

where X^{t-1} is a sequence of independent random variables with $\Pr(X_t = \mathbf{e}_i) = q_t(i)$. It will be shown that, if the second player is non-decisive, the first player can only gain from such behavior by utilizing the strategy in (30). The results of this section are formalized as follows.

Theorem 2: The optimal loss of the convex-linear-max game in Problem 1 can be achieved with a decisive player and skeleton strategies. That is, the optimal loss can be expressed as

$$L_T^* = \min_{\{\phi_t^{\text{D}}\}} \max_{q^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T} \|\varphi_T(\{\phi_t^{\text{D}}\}, q^T)\|_{\infty}. \quad (31)$$

Moreover, any skeleton strategy, $\{\phi_t^{\text{D}}\}$, and its expected skeleton strategy, $\{\bar{\phi}_t\}$, satisfy

$$\|\varphi_T(\{\bar{\phi}_t\}, q^T)\|_{\infty} \leq \max_{x^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T} \|\varphi_T(\{\phi_t^{\text{D}}\}, x^T)\|_{\infty} \quad (32)$$

for all $q^T \in \Delta_m^T$.

Proof of Theorem 2. We prove that the loss incurred by skeleton strategies and decisive players, i.e., the right-hand side (rhs) of (31), serves as both an upper and a lower bound for the optimal loss L_T^* . A byproduct of the upper bound derivation is the inequality in (32), which provides a performance bound on the expected skeleton strategy.

Restricting the maximization in (27) to $q^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T$ implies that the minimization over strategies can be restricted to skeleton strategies $\{\phi_t^{\text{D}}\}$, which leads to the lower bound:

$$\min_{\{\phi_t^{\text{D}}\}} \max_{q^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T} \|\varphi_T(\{\phi_t^{\text{D}}\}, q^T)\|_{\infty} \leq L_T^*. \quad (33)$$

For the upper bound, we have for any strategy $\{\phi_t\}$,

$$L_T^* \leq \max_{q^T \in \Delta_m^T} \|\varphi_T(\{\phi_t\}, q^T)\|_{\infty}. \quad (34)$$

We choose the strategy in (34) as the expected skeleton strategy in (30) of some skeleton strategy $\{\phi_t^{\text{D}}\}$, and show that the maximum over $q^T \in \Delta_m^T$ is attained at $q^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T$.

Consider the following upper bound on the rhs of (34) as a function of $q^T \in \Delta_m^T$:

$$\begin{aligned} \|\varphi_T(\{\bar{\phi}_t\}, q^T)\|_{\infty} &\stackrel{(a)}{=} \max_{i \in [m]} \sum_{t=1}^T q_t(i) g_i(\mathbb{E}[\phi_t^{\text{D}}(X^{t-1})]) \\ &\stackrel{(b)}{\leq} \max_{i \in [m]} \sum_{t=1}^T q_t(i) \mathbb{E}[g_i(\phi_t^{\text{D}}(X^{t-1}))] \\ &= \max_{i \in [m]} \sum_{t=1}^T \mathbb{E}[\mathbb{1}\{X_t = \mathbf{e}_i\} \mathbb{E}[g_i(\phi_t^{\text{D}}(X^{t-1}))]] \\ &\stackrel{(c)}{=} \max_{i \in [m]} \mathbb{E} \left[\sum_{t=1}^T \mathbb{1}\{X_t = \mathbf{e}_i\} g_i(\phi_t^{\text{D}}(X^{t-1})) \right] \\ &\stackrel{(d)}{\leq} \max_{x^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T} \max_{i \in [m]} \sum_{t=1}^T \mathbb{1}\{x_t = \mathbf{e}_i\} g_i(\phi_t^{\text{D}}(x^{t-1})) \\ &= \max_{x^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T} \|\varphi_T(\{\phi_t^{\text{D}}\}, x^T)\|_{\infty}, \end{aligned} \quad (35)$$

where (a) follows from the definition of the expected skeleton strategy in (30), (b) follows from Jensen's inequality, and (c) follows from the independence of X_t and X^{t-1} . Step (d) follows from the fact that q^T affects the expected value only via the probabilities allocated to each realization $x^T \in \{\mathbf{e}_1, \dots, \mathbf{e}_m\}^T$. Thus, to maximize the expected value, q^T can be chosen such that $\Pr(X^T = x^T) = 1$ for any tuple x^T that maximizes the argument of the expectation. The proof of the upper bound is completed by recalling that (35) holds for any skeleton strategy $\{\phi_t^{\text{D}}\}$ so we can take a minimum over the latter. ■

B. Optimal Solution via Balancing Trees

In the previous section, we showed that the optimal house loss can be achieved with house skeleton strategies and decisive gamblers. In the case of binary games, decisive gamblers simply choose $q^T \in \{0, 1\}^T$, and the optimal house loss in (13) is simplified to

$$L_T^* = \min_{\{r_t\}} \max_{q^T \in \{0,1\}^T} \max \left\{ \sum_{t=1}^T \frac{1-q_t}{1-r_t(q^{t-1})}, \sum_{t=1}^T \frac{q_t}{r_t(q^{t-1})} \right\}, \quad (36)$$

where $r_t : \{0, 1\}^{t-1} \rightarrow [0, 1]$, $t = 1, \dots, T$ denotes the the house strategy in the binary case. It is convenient to represent the interaction between the house and the gambler as a complete binary tree of depth T , where each node is associated with an odd $r_t \in [0, 1]$, and edges correspond to the gambler's bets, $q_t \in \{0, 1\}$. We illustrate this interaction in Fig. 1 for $T = 2$. The difficulty in solving this optimization problem lies in the fact that the loss function remains unknown to the house and the gambler during the interaction. This is reflected in the inner maximization in (36), which is taken at the end of the interaction. Our solution is based on several simplifying claims that lead to the development of our new methodology, referred to as bi-balancing trees. The first claim concerns the gambler's bet in the final round.

Claim 1: An optimal gambler places its final bet, q_T , on the winning team. As a result, the optimal house loss can be written as

$$L_T^* = \min_{\{r_t\}} \max_{q^T \in \{0,1\}^T} g(\{r_t\}, q^T), \quad (37)$$

where

$$g(\{r_t\}, q^T) = \mathbb{1}\{q_T = 0\} \sum_{t=1}^T \frac{1-q_t}{1-r_t(q^{t-1})} + \mathbb{1}\{q_T = 1\} \sum_{t=1}^T \frac{q_t}{r_t(q^{t-1})}.$$

This claim may seem counterintuitive, as the final bet is placed before the outcome is known, however, both the gambler and the choice of the winning team act as adversaries and can therefore be viewed as a single player.

Proof of Claim 1. For a fixed strategy and bets sequence q^{T-1} , the maximization over q_T and the inner maximization over the sums in (36) can be exchanged. As $q_T = 0$ and $q_T = 1$ maximize the sums $\sum_{t=1}^T \frac{1-q_t}{1-r_t(q^{t-1})}$ and $\sum_{t=1}^T \frac{q_t}{r_t(q^{t-1})}$, respectively, we can write the maximum between the two optimized sums as $\max_{q_T} g(\{r_t\}, q^T)$. ■

We illustrate the structure of the online bookmaking problem and highlight the challenges by solving the special case $T = 2$.

Example 1: The optimization for $T = 2$ is illustrated in Fig. 1. The strategy $\{r_t\}$ consists of three odds: r_1 at the root, and $r_2(0), r_2(1)$ that are chosen based on the branches labeled with $q_1 \in \{0, 1\}$. The challenge is that the house should choose the odds without the knowledge of the loss function. For example, after observing $q_1 = 0$, the house chooses $r_2(0)$, and the incurred loss can be either $\frac{1}{r_2(0)}$ or $\frac{1}{1-r_2(0)}$.

We write the possible losses for the optimization in (37) as follows:

$$\begin{aligned} g(\{r_t\}, q^2 = 00) &= \frac{1}{1-r_1} + \frac{1}{1-r_2(0)} \\ g(\{r_t\}, q^2 = 01) &= \frac{1}{r_2(0)} \\ g(\{r_t\}, q^2 = 10) &= \frac{1}{1-r_2(1)} \\ g(\{r_t\}, q^2 = 11) &= \frac{1}{r_1} + \frac{1}{r_2(1)}. \end{aligned} \quad (38)$$

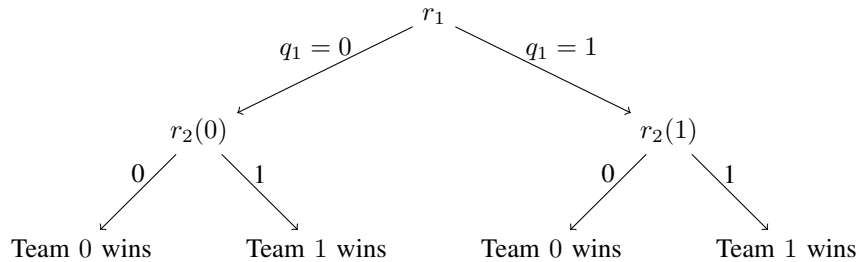


Fig. 1: A binary tree that describes the online bookmaking setup with $T = 2$. First, the house chooses r_1 , followed by the gambler's bet $q_1 \in \{0, 1\}$. The house then chooses the odd $r_2(q_1)$ based on the bet q_1 , and lastly, the bet q_2 is placed. As shown in Claim 1, the last bet q_2 is equal to the winning team.

As the gambler is taking a maximum over q^2 , the best house strategy should be a one that equates the possible losses in (38) (if it is possible). To see this, assume that $r_1, r_2(q_1)$ is an optimal strategy such that the losses are not equal. We have the chain $01 \xrightarrow{r_2(0)} 00 \xrightarrow{r_1} 11 \xrightarrow{r_2(1)} 10$ in which each link corresponds to a single odd that only affects the neighboring losses in a different direction. For instance, the link labeled with $r_2(0)$ implies that $g(\{r_t\}, q^2 = 01)$ is a decreasing function, while $g(\{r_t\}, q^2 = 00)$ is an increasing function of $r_2(0)$. If the strategy is not equalizing there is a chain of links that connects a bets sequence with the maximal loss to a sequence that has a smaller loss. We can then gradually decrease the maximum by perturbing each odd on the links starting from the lower loss towards the maximal loss.

Direct calculation shows that $r_1 = \frac{1}{2}, r_2(0) = 1 - \frac{1}{\sqrt{2}}, r_2(1) = \frac{1}{\sqrt{2}}$ is the only strategy that equates the four losses to $2 + \sqrt{2}$. This provides us with the optimal strategy and the optimal house loss $L_2^* = 2 + \sqrt{2}$. Indeed, the derivation shows that, if the winning team is equal to the final bet of the gambler, the loss is equal to the optimal house loss no matter what is the bets sequence.

From the above example, we deduce another observation.

Claim 2: If there exists a strategy that equalize the losses $g(\{r_t\}, q^T)$ (for all $q^T \in \{0, 1\}^T$), then an optimal strategy is among those that equalize the losses.

The proof of Claim 2 is deferred to the end of the section, as it requires several definitions that will only be given later.

The objective now is to find a strategy $\{r^T\}$ that equates the losses in (37) for all bets sequences q^T (if such a strategy exists). The direct approach to find a strategy that equates all bets' sequences is too involved even for small horizons, e.g., $T = 3, 4$. Our main idea is to study a more general problem which we call the bi-balancing problem: find a strategy that simultaneously equates the accumulated losses for each team. By Claim 1, it reduces to equating the accumulated loss $\sum_{t=1}^T \frac{1-q_t}{1-r_t(q^{t-1})}$ for sequences with $q_T = 0$, and the accumulated loss $\sum_{t=1}^T \frac{q_t}{r_t(q^{t-1})}$ for sequences with $q_T = 1$.

To formalize the bi-balancing problem, we define the individual value function.

Definition 1 (Individual value function): For a fixed strategy $\{r_t\}$, the *individual value function* is the vector $(V^0(q^t), V^1(q^t))$ consisting of

$$\begin{aligned} V^0(q^t) &= \sum_{i=1}^t \frac{1 - q_i}{1 - r_i(q^{i-1})} \\ V^1(q^t) &= \sum_{i=1}^t \frac{q_i}{r_i(q^{i-1})}. \end{aligned} \quad (39)$$

In words, the individual value function vector consists of the two possible future losses assuming that the forthcoming bets sequence is q^t . Note that the individual value function is defined for a tree with an arbitrary depth t that is not necessarily equal to the horizon T . By Claim 1, $g(\{r_t\}, q^T) = V^i(q^T)$ for all sequences that end with $q_T = i$ for $i = 0, 1$. Thus, we partition the sets of binary sequences as $\mathcal{Q}_i^T = \{q^T \in \{0, 1\}^T : q_T = i\}$. We can now define the bi-balancing problem.

Definition 2 (Bi-Balanced Trees): A complete binary tree is *bi-balanced* by the strategy $\{r_t\}$ if the individual value function $(V^0(q^T), V^1(p^T))$ is constant for all $q^T \in \mathcal{Q}_0^T$ and $p^T \in \mathcal{Q}_1^T$. The collection of vectors (V^0, V^1) such that $V^i = V^i(q^T)$ remains constant, for some strategy $\{r_t\}$ and all $q^T \in \mathcal{Q}_i^T$, is referred to as the set of *achievable value functions*.

A solution to the bi-balancing problem specializes to a solution for the online bookmaking problem if we also require $V^0(q^T) = V^1(p^T)$ for some $q^T \in \mathcal{Q}_0^T$ and $p^T \in \mathcal{Q}_1^T$, that is, an identical house loss for all bets sequences. Also, Definition 2 focuses on complete binary trees, while the problem formulation can be trivially generalized to non-binary trees, and even to non-complete trees if some bets sequences are assumed to be not feasible.

Our approach to the bi-balancing problem is to parameterize the set of achievable value functions on any node of the binary tree. Specifically, an explicit formula that relates the vector coordinates of the individual value function will be given using the function

$$f_d(x) = d \frac{x - (d-1)}{x - d}, \quad (40)$$

where d is the tree depth. Note that $f_d(x)$ is an involution, i.e., $f_d(f_d(x)) = x$, a key property that will be used in the sequel.

We first characterize the set of achievable value functions and then focus on deriving the optimal strategy in Section III-C. The following is our main result regarding the bi-balancing problem.

Theorem 3: For a tree of depth d , the set of achievable value functions is characterized by the union of vectors $(f_d(x), x)$ over $x > d$. That is, the individual value function of any bi-balanced tree of depth d has the form $(f_d(x), x)$ for some $x > d$. Moreover, for a fixed $x > d$, there exists a single strategy whose individual value function is $(f_d(x), x)$.

We use Theorem 3 to build a bi-balanced tree in a bottom-up fashion, by connecting two bi-balanced trees of depth d to a single root, ensuring that the new tree, whose depth is $d+1$ is also bi-balanced. Theorem 3 provides a complete characterization

for the relation between the coordinates of the individual value function; for any horizon T where d steps remain, if the future loss for the first team is x , then $f_d(x)$ is the future loss for the second team. The claim is symmetric if we exchange the roles of the first and the second teams; in particular, the set of achievable functions can also be characterized by the union of $(x, f_d(x))$. This can be easily deduced from the fact that $f_d(x)$ is an involution.

The value function of a bi-balanced tree depends on two parameters, the remaining tree depth d and the argument x . It does not depend directly on T , but only via x . Particularly, for a binary tree of height T , there are 2^{T-d} nodes whose remaining depth is d . Clearly, each node at this depth has a different value function, and Theorem 3 does not provide the individual value functions for each tree node (this will be provided later). Still, it is interesting to note that the value function structure is identical, at any fixed depth, no matter what is the overall horizon T and what is the prefix sequence leading to the sub-tree. The set of achievable value functions is illustrated in Fig. 2 where, for improved visibility, the normalized value function $\frac{1}{d}(f_d(x), x)$ is plotted.

As mentioned above, our derivation relies on a recursive construction of a bi-balanced tree from the bottom to the top. Specifically, we employ induction on the tree depth, where at each step two bi-balanced trees are connected by linking their roots to a new common root. This provides a subset of the achievable value functions, as it is based on a particular construction.

Before proceeding to the proof of the theorem, we claim the converse: our construction captures any bi-balanced tree. The left and right sub-trees of a tree with depth d are defined as the two trees of depth $d - 1$ connected to the tree's root with $q_1 = 0$ and $q_1 = 1$, respectively. For a strategy $\{r_t\}_{t=1}^d$, the individual value function of the left sub-tree is a vector that consists

$$\begin{aligned} V_L^0(q_2^d) &= \sum_{t=2}^d \frac{1 - q_t}{1 - r_t(0q_2^{t-1})} \\ V_L^1(q_2^d) &= \sum_{t=2}^d \frac{q_t}{r_t(0q_2^{t-1})}, \end{aligned} \quad (41)$$

and similarly for the right sub-tree, with $q_1 = 0$ replaced by $q_1 = 1$. The optimality of our construction follows from the following claim.

Claim 3: If a tree of depth d is bi-balanced, then its two descendant sub-trees are bi-balanced. Specifically, the values $V_L^0(q_2^d)$, $V_R^0(q_2^d)$, $V_L^1(p_2^d)$, and $V_R^1(p_2^d)$ remain constant for all $q_2^d \in \mathcal{Q}_0^{d-1}$ and $p_2^d \in \mathcal{Q}_1^{d-1}$. This is analogous to the dynamic programming principle, where a problem can be solved optimally by optimally solving its tail subproblems.

Proof of Claim 3. If any of the sub-trees is not bi-balanced, there exist two sequences $q_2^d, \tilde{q}_2^d \in \mathcal{Q}_0^{d-1}$ such that $V_L^0(q_2^d) \neq V_L^0(\tilde{q}_2^d)$ (the choice of $V_L^0(\cdot)$ is without loss of generality). Then, it follows by (39) that $V^0(0q_2^d) \neq V^0(0\tilde{q}_2^d)$. ■

Proof of Theorem 3. We construct a bi-balanced tree whose value function is $(x, f_d(x))$, and then claim the converse: any bi-balanced tree has the particular value function structure constructed by our method.

For the construction, we propose a sequential method that is inspired by Markov decision processes. The main idea is to construct a bi-balanced tree by properly connecting two bi-balanced sub-trees. This allows us to construct a bi-balanced tree using an induction on the depth $d = 1, 2, \dots$. The induction proof:

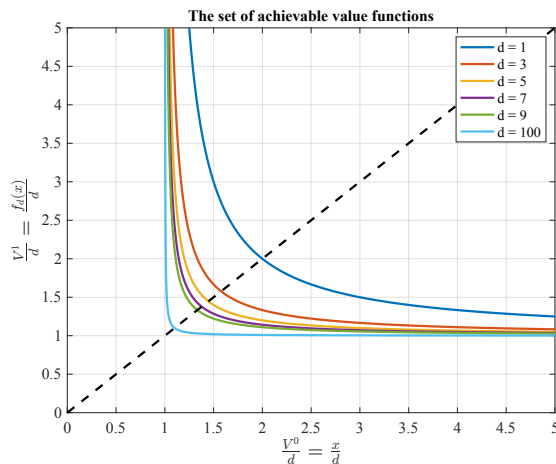


Fig. 2: The set of achievable value functions in Theorem 3 as a function of the remaining depth, normalized by depth. The curves are symmetric around the line $V^1 = V^0$ due to symmetry between the teams. The optimal house loss for an horizon d can be obtained by intersecting its curve with the dashed line $V^0 = V^1$. As d grows, the intersection approaches the lower bound $(V^0, V^1) = (1, 1)$ with a convergence rate of $\frac{1}{\sqrt{d}}$.

- 1) **Base Case:** Consider a tree of depth $d = 1$. The house strategy consists of a single odd $r_1 \in [0, 1]$ at the root, and is followed by two possible bets $q_1 \in \{0, 1\}$, each corresponds to a branch. If the individual value function on the left branch is $V^0(0) = x$, the odd satisfies $\frac{1}{1-r_1} = x$, which implies in turn that $V^1(1) = \frac{1}{r_1} = f_1(x)$ as required. The constraint $r_1 \in (0, 1)$ implies $x > 1$.
- 2) **Induction Hypothesis:** By the induction hypothesis, for any $x > d$, there exists a unique bi-balanced tree of depth d whose value function is $(x, f_d(x))$.
- 3) **Induction Step:** From the induction hypothesis, there exist two bi-balanced trees of depth d with value functions $(f_d(y), y)$ and $(x, f_d(x))$, for some $x, y \geq d$.² The trees with the value function $(f_d(y), y)$ and $(x, f_d(x))$ will be used as the left and right sub-trees, respectively. That is, a new root will be connected via the branch $q_1 = 0$ to the left sub-tree, and via the branch $q_1 = 1$ to the right sub-tree. At the root of the constructed tree, there is a new odd r_1 , to be chosen such that the constructed tree is bi-balanced. The tree construction is illustrated in Fig. 3.

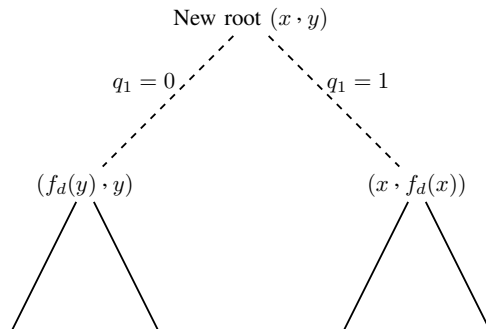


Fig. 3: Illustration of the construction: a new root is connected with two dashed edges to two sub-trees, where each is bi-balanced with value functions $(f_d(y), y)$ and $(x, f_d(x))$. The sub-trees have a depth d and the constructed tree is of depth $d + 1$. The value function at the new root, (x, y) , follows from the fact that $q_1 = 0/1$ do not change the loss if team 1/0 wins. The proof shows that there exists a unique odd at the root, r_1 , such that the tree is bi-balanced and $y = f_{d+1}(x)$.

We proceed to compute the individual value function for the new $d + 1$ -deep tree by utilizing their recursive relation with the individual value functions of the d -deep sub-trees. We first consider sequences of the form $0q^d$ where $q^d \in \mathcal{Q}_0^d$. That is, $q_1 = 0$ leads to the left sub-tree followed by any sequence of length d that ends with $q_d = 0$. The individual value function for such sequences is

$$V^0(0q^d) = \frac{1}{1-r_1} + f_d(y), \quad (42)$$

where we used the facts that the individual value function of the left sub-tree for sequences with $q_d = 0$ is $f_d(y)$, and that the immediate loss is $\frac{1}{1-r_1}$. Similarly, the individual value functions for the other sequences are

$$\begin{aligned} V^1(0q^d) &= y, & \forall q^d \in \mathcal{Q}_1^d \\ V^0(1q^d) &= x, & \forall q^d \in \mathcal{Q}_0^d \\ V^1(1q^d) &= \frac{1}{r_1} + f_d(x), & \forall q^d \in \mathcal{Q}_1^d. \end{aligned} \quad (43)$$

Recall that to bi-balance the constructed tree, we should require $V^i(0q^d) = V^i(1q^d)$ for $i = 0, 1$ (and all $q^d \in \mathcal{Q}_i^d$). That is, we equate the loss of all paths leading to the same winning team. Using (42)-(43), these two equations can be written as

$$\begin{aligned} x &= \frac{1}{1-r_1} + f_d(y) \\ y &= \frac{1}{r_1} + f_d(x), \end{aligned} \quad (44)$$

and can be solved in terms of x as

$$\begin{aligned} y &= \frac{(d+1)(x-d)}{x-(d+1)} \\ &= f_{d+1}(x). \end{aligned} \quad (45)$$

²Without loss of generality, a value functions $(y, f_d(y))$ can be written as $(f_d(y), y)$ since the mapping $f_d(\cdot)$ is an involution, i.e., $f_d(f_d(y)) = y$.

This shows that the value function at the new root is $(x, f_{d+1}(x))$.

It remains to determine the constraints on the value of x such that $r_1 \in (0, 1)$. By (44)-(45),

$$1 - r_1 = \frac{1}{x - f_d(f_{d+1}(x))} \quad (46)$$

$$= \frac{x}{(x - d)^2 + d}, \quad (47)$$

which is greater than 0 only if $x > 0$. We can also arrange the equation as

$$1 - r_1 = \frac{(x - d) + d}{(x - d)^2 + d}, \quad (48)$$

which is smaller than 1 iff $x > d + 1$ or $x - d < 0$. Combined with the induction hypothesis, $x > d$, we conclude that for all $x > d + 1$, the new odd at the root satisfies $1 - r_1 = \frac{1}{x - f_d(f_{d+1}(x))} \in (0, 1)$. We remark that the constructed strategy inherits the strategy at subsequent steps from the sub-trees that are assumed to exist by the induction hypothesis. It can be easily verified that $y = f_{d+1}(x) > d + 1$, and we also showed the constraint $x > d + 1$. These two conditions guarantee that the strategies at the left and right sub-trees are feasible.

Claim 3 shows the converse: any bi-balanced tree of depth $d + 1$ can be constructed by connecting two bi-balanced trees, each of depth d . As the sub-trees used for the construction are arbitrary, our construction characterizes all bi-balanced trees of depth $d + 1$. ■

C. The optimal strategy

In this section we utilize Theorem 3 on the structure of the value function at the root to derive a strategy $\{r_t\}_{t=1}^T$ for the bi-balancing problem. The main idea is that the individual value function at the root can be used to compute the individual value function on any node of the tree in a forward manner. Let $V_{q^j}^i(q_{j+1}^T)$ denote the value function at the root of the sub-tree whose prefix is q^j , and the future bets sequence is q_{j+1}^T . Throughout this section, we assume that trees are bi-balanced so the dependence on future bets is omitted. The following theorem describes a forward sequential algorithm that finds the individual value functions and the optimal odds as a function of the input q_1, q_2, \dots .

Theorem 4: Consider a bi-balanced tree of depth T whose value function at the root is $(V^0, V^1) = (x, f_T(x))$. For any sequence $q_1 q_2 \dots$, the individual value function along this sequence can be computed recursively as

$$(V_{q^t}^0, V_{q^t}^1) = \begin{cases} (f_{T-t}(V_{q^{t-1}}^1), V_{q^{t-1}}^1), & \text{if } q_t = 0 \\ (V_{q^{t-1}}^0, f_{T-t}(V_{q^{t-1}}^0)), & \text{if } q_t = 1 \end{cases} \quad (49)$$

with the initial condition $(V^0, V^1) = (x, f_T(x))$. The subsequent odd is then equal to

$$r_t(q^{t-1}) = \frac{1}{V_{q^{t-1}}^1 - V_{q^{t-1}}^0}. \quad (50)$$

The recursive computation of the individual value functions is the basis for Algorithm 1. Theorem 4 shows that we can update the individual value function in a forward manner so that we only need t computations to execute a strategy up to time t . This is a significant reduction in the complexity since in general MDPs (and vector games), where the value function is updated recursively but in a backward fashion. The reduced complexity is due to the fact that, for the online bookmaking problem, the vector of future losses is changing in one of its coordinates at each step.

Proof of Theorem 4. We use the fact that a new bet $q_t \in \{0, 1\}$ changes the individual value function only in one of its coordinates. In particular,

$$\begin{aligned} V_{q^{t-1}0}^1 &= V_{q^{t-1}}^1 \\ V_{q^{t-1}1}^0 &= V_{q^{t-1}}^0. \end{aligned} \quad (51)$$

The recursive relation in (49) is completed by Theorem 3 showing that the other coordinate of the individual value function can be computed by applying $f_{T-t}(\cdot)$ on the invariant coordinate in (51). The equation for computing the odd in (50) follows from the relation

$$V_{q^{t-1}}^1 = \frac{1}{r_t} + V_{q^{t-1}1}^0. \quad (52)$$

■

D. Proof of Theorem 1 and Claim 2

Proof of Theorem 1. The online bookmaking problem is a special case of the bi-balancing problem with the constraint $V^0(q^{T-1}0) = V^1(q^{T-1}1)$ for some q^{T-1} (which implies that this holds for all $q^{T-1} \in \{0,1\}^{T-1}$ since the tree is bi-balanced). By Theorem 3, this translates to equating the two coordinates of the individual value function at the root, that is,

$$x = f_T(x), \quad (53)$$

which simplifies to $x = T \pm \sqrt{T}$. By the range limit, $x > T$, the optimal house loss is $L_T^* = T + \sqrt{T}$.

We proceed to prove the optimality of Algorithm 1 for decisive gamblers, i.e., $q^T \in \{0,1\}^T$. The algorithm is essentially a memory-efficient implementation of Theorem 4. We interpret the variables a, b, b^+ that are used in Algorithm 1. The state of Algorithm 1, denoted by (a, b) , corresponds to the individual value function $(V_{q^{t-1}}^0, V_{q^{t-1}}^1)$ after observing the sequence q^{t-1} . By (49), the update of the individual value function is only a function of the previous individual value function (at the invariant coordinate) and the new bet q_t . The state update in Algorithm 1 is precisely (49). The variable b^+ in Algorithm 1 corresponds to the individual value function $V_{q^{t-1}}^1$, that is, we hypothetically assume that $q_t = 1$. This is utilized to compute the optimal odd in (52). By (52), first odd is equal to

$$r_1 = \frac{1}{V_\emptyset^1 - V_1^1} = 1 \frac{1}{T + \sqrt{T} - f_{T-1}(T + \sqrt{T})} = 1/2. \quad (54)$$

Finally, by Theorem 2, the house strategy's loss for the expected skeleton strategy (25) with respect to the optimal skeleton strategy in Algorithm 1 is upper bounded by the optimal house loss. ■

Proof of Claim 2. We show that if a strategy $\{r_t\}$ does not induce an equal loss, i.e., $g(\{r_t\}, q^T)$ is the same for all q^T , it is not optimal. The key idea is to perturb this strategy to obtain one that achieves a smaller maximum among the losses. For $\{r_t\}$, let $M = \max_{b^T} g(\{r_t\}, b^T)$, and define $\mathcal{M} = \{b^T \in \{0,1\}^T : g(\{r_t\}, b^T) = M\}$. We partition the sequences in \mathcal{M} by identifying the shortest prefixes such that all subsequent suffixes remain in \mathcal{M} . Specifically, a subset of \mathcal{M} is defined by a prefix b^i and contains $b^i q^{T-i}$ for all q^{T-i} , while the latter property does not hold for the prefix b^{i-1} . We will now present a procedure to decrease the loss of every sequence in such subsets below M . We start by applying the procedure on subsets that have the longest prefix, denoted here by b^i , and then apply the procedure on prefixes with shorter lengths.

The procedure is to *perturb* $r_i := \phi_i(b^{i-1})$ and subsequent strategies along the sequence $b^i 1^{T-i}$, and analyze the effect on the losses (for all binary sequences). In particular, we increase r_i and decrease $r_{i+j} := \phi_{i+j}(b^{i-1} 1^j)$ for $j = 1, \dots, T-i$. The perturbation clearly has no effect on the losses of sequences whose prefix is not b^{i-1} .

The losses for sequences with prefix b^{i-1} can be divided into four cases based on the values of b_i, b_T . Without loss of generality, assume $b_i = 1$. By our partition, all sequences with prefix b^i have a identical maximal loss,

$$\begin{aligned} M &= g(\{r_t\}, q^T = b^{i-1} 1 q_{i+1}^{T-1} 0) = V^0(b^{i-1}) + V_{b^{i-1} 1}^0(q_{i+1}^{T-1} 0), & \forall q_{i+1}^{T-1} \\ M &= g(\{r_t\}, q^T = b^{i-1} 1 q_{i+1}^{T-1} 1) = V^1(b^{i-1}) + \frac{1}{r_i} + V_{b^{i-1} 1}^1(q_{i+1}^{T-1} 1), & \forall q_{i+1}^{T-1}. \end{aligned} \quad (55)$$

Also, recall that $V_{b^i}^i(b_{i+1}^T)$ is the value function starting at the sub-tree whose prefix is b^i , and its future bets sequence is b_{i+1}^T .

Recall our assumption that b^i is the longest prefix among the subsets of \mathcal{M} . Thus, each sequence with prefix $b^{i-1} 0 \neq b^i$ has a loss smaller than M , and can be written as

$$\begin{aligned} M &> g(\{r_t\}, q^T = b^{i-1} 0 q_{i+1}^{T-1} 0) = V^0(b^{i-1}) + \frac{1}{1-r_i} + V_{b^{i-1} 0}^0(q_{i+1}^{T-1} 0), & \forall q_{i+1}^{T-1} \\ M &> g(\{r_t\}, q^T = b^{i-1} 0 q_{i+1}^{T-1} 1) = V^1(b^{i-1}) + V_{b^{i-1} 0}^1(q_{i+1}^{T-1} 1), & \forall q_{i+1}^{T-1}. \end{aligned} \quad (56)$$

Our objective is to guarantee that the losses of all sequences with prefix b^{i-1} , detailed in (55)-(56), is smaller than M . The continuity of the functions $\frac{1}{r}, \frac{1}{1-r}$ implies that r_i can be increased such that the loss of sequences of the form $b^{i-1} 1 q_{i+1}^{T-1} 1$ decreases below M , while the losses of sequences $b^{i-1} 0 q_{i+1}^{T-1} 0$ increase but do not exceed M . We are then left with sequences of the form $b^{i-1} 1 q_{i+1}^{T-1} 0$ whose losses are not effected by r_i and therefore are still equal to M . By decreasing r_{i+j} , for all $j > 0$, the losses of sequences of the form $b^{i-1} 1 q_{i+1}^{T-1} 0$ decrease below M . However, by doing so, we also increase the losses of sequences of the form $b^{i-1} 1 q_{i+1}^{T-1} 1$ (whose losses are smaller than M). Again, by continuity, we can guarantee that the perturbations are small enough such that these losses do not exceed M .

At the end of the perturbation procedure described above, all sequences with prefix b^{i-1} have losses that are smaller than M and, therefore, can be eliminated from \mathcal{M} . We repeat the procedure above for any subset of sequences whose prefix length is i . We then apply the procedure for prefixes of length $i-1$ and so on until all sequences in \mathcal{M} have a loss smaller than M . ■

IV. MONTE-CARLO-BASED APPROXIMATION OF THE OPTIMAL ALGORITHM

In this section, we discuss complexity aspects of implementing the optimal house strategy to calculate r_t and provide an efficient algorithm to closely approximate r_t for general gamblers (i.e. non-decisive gamblers who can potentially choose $q_t \in [0, 1]$).

First, we note that if the gamblers are decisive, Algorithm 1 which computes the optimal strategy has a linear (in T) runtime—this follows since at each time step simple computations are executed to update the state (a, b) and compute r_t as a function of the previous state, t , and the gambler's previous bet q_{t-1} .

For general non-decisive gamblers, computing the odds $\bar{r}_t^{\text{ALG}}(q^{t-1}) = \mathbb{E}[r_t^{\text{ALG}}(X^{t-1})]$, as defined in (25) can take exponential time as discussed in Section II-B. In order to achieve an efficient algorithm to calculate \bar{r}_t^{ALG} in this general case, we propose to evaluate $\mathbb{E}[r_t^{\text{ALG}}(X^{t-1})]$ via a Monte Carlo simulation in lieu of an exact computation. If $X_{(1)}^{t-1}, \dots, X_{(N)}^{t-1}$ are N independent copies of X^{t-1} , then

$$\frac{1}{N} \sum_{j=1}^N r_t^{\text{ALG}}(X_{(j)}^{t-1}) \approx \mathbb{E}[r_t^{\text{ALG}}(X^{t-1})]. \quad (57)$$

The Monte Carlo method based algorithm (Algorithm 2) essentially runs N parallel copies of Algorithm 1, keeping track of N states $(a_{(j)}, b_{(j)})_{j=1}^N$. Upon receiving the gambler's bet $q_t \in [0, 1]$ the algorithm generates $X_{(j)t} \sim \text{Ber}(q_t)$, updates the state $(a_{(j)}, b_{(j)})$ and generates an action for the house as done in Algorithm 1; the N copies of this action are then averaged to approximate the odds $\bar{r}_t^{\text{ALG}}(q^{t-1}) = \mathbb{E}[r_t^{\text{ALG}}(X^{t-1})]$ as in (57). Algorithm 2 has two (nested) loops of size N and T with a constant number of operations in each so that the total time complexity is $\mathcal{O}(NT)$.

Algorithm 2 Monte Carlo Based Efficient Strategy

Inputs: T (Total rounds), q_1, \dots, q_{T-1} (Gambler's bets obtained sequentially)

Output: r_1, \dots, r_t (House strategy, r_t is output after obtaining q_1, \dots, q_{t-1})

Initialization: $a_{(j)} = b_{(j)} = T + \sqrt{T}$, $j \in [N]$

$r_1 \leftarrow \frac{1}{2}$

Output r_1

for $t = 1 : T - 1$ **do**

$r_{t+1} \leftarrow 0$

$d \leftarrow T - t$

for $j = 1 : N$ **do**

$X \sim \text{Ber}(q_t)$

if $X = 0$ **then**

$a_{(j)} \leftarrow d \frac{b_{(j)} - (d-1)}{b_{(j)} - d}$

else if $X = 1$ **then**

$b_{(j)} \leftarrow d \frac{a_{(j)} - (d-1)}{a_{(j)} - d}$

end if

$b_{(j)}^+ \leftarrow (d-1) \frac{a_{(j)} - (d-2)}{a_{(j)} - (d-1)}$

$r_{t+1} \leftarrow r_{t+1} + \frac{1}{N} \cdot \frac{1}{b_{(j)} - b_{(j)}^+}$

end for

 Output r_{t+1}

end for

As N increases, the approximation calculated by Algorithm 2 converges rapidly to the desired odds $\{\bar{r}_t^{\text{ALG}}\}$. This rate of convergence can be quantified as follows.

Proposition 1: Let $\bar{r}_t^{\text{ALG}} = \bar{r}_t^{\text{ALG}}(q^{t-1}) = \mathbb{E}[r_t^{\text{ALG}}(X^{t-1})]$, and $\hat{r}_t^N = \hat{r}_t^N(q^{t-1}) := \frac{1}{N} \sum_{j=1}^N r_t^{\text{ALG}}(X_{(j)}^{t-1})$ be its approximation computed by Algorithm 2. Then for any $\epsilon > 0$, and $q^T \in [0, 1]^T$, we have that $\Pr(|\hat{r}_t^N - \bar{r}_t^{\text{ALG}}| \geq \epsilon) \leq 2 \exp(-2N\epsilon^2)$ and subsequently $\Pr(|\hat{r}_t^N - \bar{r}_t^{\text{ALG}}| \geq \epsilon \text{ for any } t \in [T]) \leq 2T \exp(-2N\epsilon^2)$

Proof. Since \hat{r}_t^N is a sum of N independent and identically distributed random variables $r_t^{\text{ALG}}(X_{(1)}^{t-1}), \dots, r_t^{\text{ALG}}(X_{(N)}^{t-1})$ each bounded in $[0, 1]$ with mean $\bar{r}_t^{\text{ALG}}(q^{t-1})$, the first assertion follows by the Hoeffding inequality. The second assertion follows via the union bound. ■

Proposition 1 implies that in order to achieve $\Pr(|\hat{r}_t^N - \bar{r}_t^{\text{ALG}}| \geq \epsilon \text{ for any } t \in [T]) \leq \delta$, choosing any $N \geq \frac{1}{2\epsilon^2} \log\left(\frac{2T}{\delta}\right)$ suffices. In order to choose an appropriate value of ϵ , we consider the effect of choosing the odds \hat{r}_t^N instead of the desired odds \bar{r}_t^{ALG} . In particular, ϵ should be small enough so that the worst-case loss

$$\hat{L}_T^N := \max_{q^T} \max \left\{ \sum_{t=1}^T \frac{q_t}{\hat{r}_t^N}, \sum_{t=1}^T \frac{1 - q_t}{1 - \hat{r}_t^N} \right\} \quad (58)$$

obtained by playing the (suboptimal) action \widehat{r}_t^N is close to L_T^* . To do this, we utilize the following.

Claim 4: For fixed horizon size T , and any $q^T \in [0, 1]^T$ we have $\frac{1}{T+\sqrt{T}} \leq \bar{r}_t^{\text{ALG}}(q^{t-1}) \leq 1 - \frac{1}{T+\sqrt{T}}$ for all $1 \leq t \leq T$.

Proof. This follows from the fact that the strategy $\{\bar{r}_t^{\text{ALG}}\}$ attains the optimal house loss $L_T^* = T + \sqrt{T}$. Indeed, assume for contradiction that there is some q^{t-1} for which $\bar{r}_t^{\text{ALG}}(q^{t-1}) < \frac{1}{T+\sqrt{T}}$. Then, the individual game loss for any q^T beginning with q^{t-1} is at least $\frac{1}{\bar{r}_t^{\text{ALG}}(q^{t-1})} > T + \sqrt{T} = L_T^*$, in contradiction to the fact that $\{\bar{r}_t^{\text{ALG}}\}$ attains the optimal house loss L_T^* . The upper bound $\bar{r}_t^{\text{ALG}}(q^{t-1}) \leq 1 - \frac{1}{T+\sqrt{T}}$ is proved similarly. ■

Lemma 1: Let $\delta \in (0, 1)$, assume $\epsilon \leq \frac{1}{2(T+\sqrt{T})}$ and $N \geq \frac{1}{2\epsilon^2} \log\left(\frac{2T}{\delta}\right)$. Let $\{\bar{r}^{\text{ALG}}\}$ be the house strategy computed by Algorithm 1 and (25), and $\{\widehat{r}^N\}$ be the house strategy computed by Algorithm 2. Then, with probability at least $1 - \delta$, for any $q^T \in [0, 1]^T$ the corresponding individual game losses satisfy

$$\|\varphi_T(\{\widehat{r}^N\}, q^T)\|_\infty \leq \left(1 + 2\epsilon(T + \sqrt{T})\right) \|\varphi_T(\{\bar{r}^{\text{ALG}}\}, q^T)\|_\infty \leq \left(1 + 2\epsilon(T + \sqrt{T})\right) L_T^*. \quad (59)$$

Proof. Since $N \geq \frac{1}{2\epsilon^2} \log\left(\frac{2T}{\delta}\right)$, Proposition 1 ensures that the event $|\widehat{r}_t^N - \bar{r}_t^{\text{ALG}}| \leq \epsilon$ for all $t \in [T]$ occurs with probability at least $1 - \delta$. If this event occurs, then

$$\frac{q_t}{\widehat{r}_t^N} \leq \frac{q_t}{\bar{r}_t^{\text{ALG}} - \epsilon} \leq \frac{q_t}{\bar{r}_t^{\text{ALG}}} \cdot \frac{1}{1 - \epsilon(T + \sqrt{T})}$$

and similarly,

$$\frac{1 - q_t}{1 - \widehat{r}_t^N} \leq \frac{1 - q_t}{1 - \bar{r}_t^{\text{ALG}}} \cdot \frac{1}{1 - \epsilon(T + \sqrt{T})}.$$

where both of these use Claim 4 to bound \bar{r}_t^{ALG} . Putting these two together implies that for any $q^T \in [0, 1]^T$, with probability $1 - \delta$

$$\|\varphi_T(\{\widehat{r}^N\}, q^T)\|_\infty = \max \left\{ \sum_{t=1}^T \frac{q_t}{\widehat{r}_t^N(q^{t-1})}, \sum_{t=1}^T \frac{1 - q_t}{1 - \widehat{r}_t^N(q^{t-1})} \right\} \quad (60)$$

$$\leq \frac{1}{1 - \epsilon(T + \sqrt{T})} \max \left\{ \sum_{t=1}^T \frac{q_t}{\bar{r}_t^{\text{ALG}}(q^{t-1})}, \sum_{t=1}^T \frac{1 - q_t}{1 - \bar{r}_t^{\text{ALG}}(q^{t-1})} \right\} \quad (61)$$

$$\leq \left(1 + 2\epsilon(T + \sqrt{T})\right) \|\varphi_T(\{\bar{r}^{\text{ALG}}\}, q^T)\|_\infty \quad (62)$$

$$\leq \left(1 + 2\epsilon(T + \sqrt{T})\right) L_T^*, \quad (63)$$

where (62) uses that $\frac{1}{1-x} \leq 1 + 2x$ for $0 \leq x \leq 1/2$, and (63) the fact that $\max_{q^T \in [0, 1]^T} \|\varphi_T(\{\bar{r}^{\text{ALG}}\}, q^T)\|_\infty = L_T^*$ from the fact that the strategy $\{\bar{r}^{\text{ALG}}\}$ attains the optimal house loss. ■

In order to get $1 + 2\epsilon(T + \sqrt{T}) \rightarrow 1$, we must choose N large enough so that the corresponding $\epsilon = o\left(\frac{1}{T}\right)$. For example, $\epsilon = \frac{1}{T^{1+\alpha}}$ for any $\alpha > 0$ suffices. Recalling that the time complexity of Algorithm 2 is $\mathcal{O}(NT)$ and that in order to achieve an ϵ (additive) approximation to \bar{r}_t^{ALG} with probability $1 - \delta$ requires $N \geq \frac{T}{2\epsilon^2} \log\left(\frac{2T}{\delta}\right)$, we get that for choice of $\epsilon = \frac{1}{T^{1+\alpha}}$ the time complexity of Algorithm 2 is $\mathcal{O}\left(T^{3+2\alpha} \log\left(\frac{T}{\delta}\right)\right)$. Even for fairly small values of δ , this is far smaller than the $\mathcal{O}(T^2)$ complexity of computing \bar{r}_t^{ALG} exactly.

V. SUB-OPTIMAL BLACKWELL-INSPIRED ALGORITHM

The purpose of this section is to show that house strategies for the online bookmaking game, whose amortized loss tends to 1 as $T \rightarrow \infty$, can also be derived through standard approaches. However, these strategies are sub-optimal, and do not even attain a $\mathcal{O}(T^{-1/2})$ rate of convergence. We only consider the binary case, but the algorithm below can be extended to $m > 2$ as well.

We follow Blackwell's algorithm [12] (see also [25]), to derive a sub-optimal algorithm for the optimal bookmaking problem. Since our loss functions are unbounded in general, we cannot apply Blackwell's algorithm as-is in order to approach an ℓ_∞ ball, and some adaptation is required. Because of this adaptation, we can only achieve a regret of $\mathcal{O}(T^{-1/4})$, rather than the $\mathcal{O}(T^{-1/2})$ rate of approachability that Blackwell's algorithm provides for bounded loss functions.

While in previous sections it was convenient to think of probability vectors in the simplex Δ_2 as points $r_t, q_t \in [0, 1]$, in this section r_t, q_t will refer to two-dimensional vectors in the simplex. Recall that for two vectors $r = (r(0), r(1))$ and $q = (q(0), q(1))$ in the simplex Δ_2 our loss function is

$$f(r, q) = \left(\frac{q(0)}{r(0)}, \frac{q(1)}{r(1)} \right)^\top. \quad (64)$$

For some $1 < \Delta < 2$ to be specified later, define the vectors

$$\lambda_1 = \left(1 - \frac{1}{\Delta}, \frac{1}{\Delta}\right)^\top, \quad \lambda_1^\perp = (1, 1 - \Delta)^\top \quad (65)$$

$$\lambda_2 = \left(\frac{1}{\Delta}, 1 - \frac{1}{\Delta}\right)^\top, \quad \lambda_2^\perp = (1 - \Delta, 1)^\top. \quad (66)$$

Define the sets $\mathcal{S}, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3 \subset [0, \infty)^2$ as

$$\begin{aligned} \mathcal{S} &= \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_1 \leq 0, (x - (1, 1)^\top)^\top \lambda_2 \leq 0\} \\ \mathcal{A}_1 &= \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_1 > 0, (x - (1, 1)^\top)^\top \lambda_1^\perp \leq 0\} \\ \mathcal{A}_2 &= \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_2 > 0, (x - (1, 1)^\top)^\top \lambda_2^\perp \leq 0\} \\ \mathcal{A}_3 &= \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_1^\perp > 0, (x - (1, 1)^\top)^\top \lambda_2^\perp > 0\} \end{aligned} \quad (67)$$

See Figure 4.

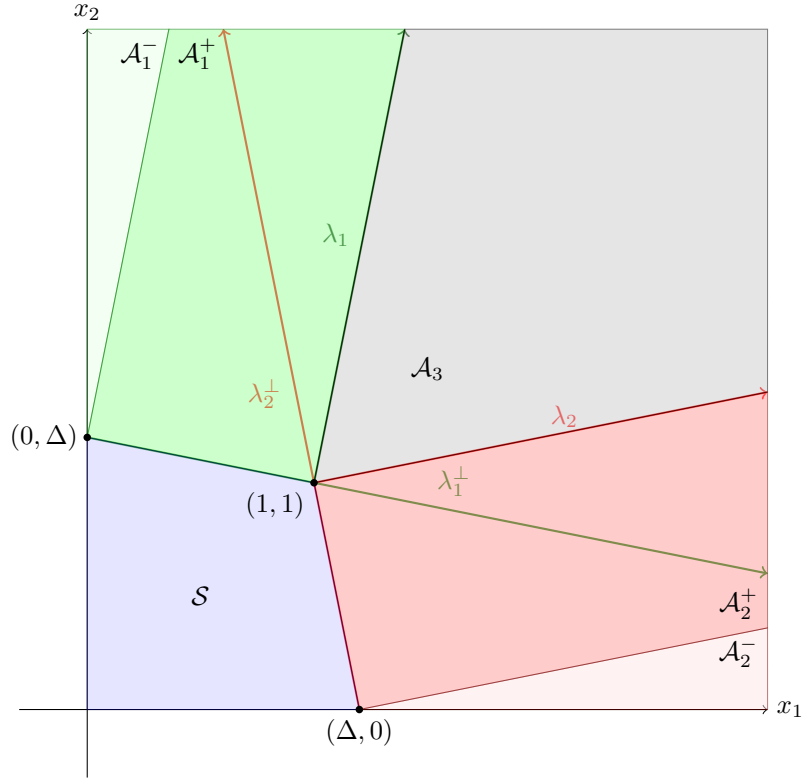


Fig. 4: Illustration of the regions defined in (67) and in (82). Here $\mathcal{A}_1 = \mathcal{A}_1^- \cup \mathcal{A}_1^+$ and $\mathcal{A}_2 = \mathcal{A}_2^- \cup \mathcal{A}_2^+$.

The algorithm is as follows. Let $\bar{\varphi}_0 = (0, 0)^\top$. For any $t = 1, 2, \dots$ let

$$\bar{\varphi}_t = \frac{1}{t} \sum_{i=1}^t f(r_i, q_i), \quad (68)$$

and for any $t = 0, \dots, T - 1$ set

$$r_{t+1} = \begin{cases} \lambda_1 & \bar{\varphi}_t \in \mathcal{A}_1 \\ \lambda_2 & \bar{\varphi}_t \in \mathcal{A}_2 \\ \frac{\bar{\varphi}_t - (1, 1)^\top}{\|\bar{\varphi}_t - (1, 1)^\top\|_1} & \bar{\varphi}_t \in \mathcal{A}_3 \\ \left(\frac{1}{2}, \frac{1}{2}\right)^\top & \bar{\varphi}_t \in \mathcal{S} \end{cases}. \quad (69)$$

Note that $r_{t+1} = \phi_{t+1}(q_1, \dots, q_t)$ is a deterministic function of the previous points in the simplex chosen by the gambler. Thus, it is a valid strategy for the binary online bookmaking game.

Theorem 5: For any $1 < \Delta < 2$, and any $q_1, \dots, q_t \in \Delta_2$, the algorithm specified by equation (69), with $\lambda_1, \lambda_2, \lambda_1^\perp, \lambda_2^\perp, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{S}, \bar{\varphi}_t$, defined by equations (65-68) attains

$$\frac{1}{t} \|\varphi_t(\{\phi_t\}, q^t)\|_\infty = \|\bar{\varphi}_t\|_\infty \leq \Delta + \sqrt{\frac{2}{t}} \frac{\Delta^2}{\Delta - 1}. \quad (70)$$

In particular, if we apply the algorithm for T iterations with $\Delta = 1 + \delta_T$, where

$$\delta_T = \frac{\left(\frac{2}{T}\right)^{1/4}}{1 - \left(\frac{2}{T}\right)^{1/4}}, \quad (71)$$

the algorithm attains

$$\frac{1}{T} \|\varphi_T(\{\phi_t\}, q^T)\|_\infty = \|\bar{\varphi}_T\|_\infty \leq 1 + 2\delta_T = 1 + 2 \frac{\left(\frac{2}{T}\right)^{1/4}}{1 - \left(\frac{2}{T}\right)^{1/4}}. \quad (72)$$

Proof of Theorem 5. For a point $x \in [0, \infty)^2$, and a set $\mathcal{A} \subset \mathbb{R}^2$ let

$$\text{dist}(x, \mathcal{A}) = \min_{a \in \mathcal{A}} \|x - a\| \quad (73)$$

denote the distance between x and \mathcal{A} , and

$$\pi_{\mathcal{A}}(x) = \underset{a \in \mathcal{A}}{\text{argmin}} \|x - a\| \quad (74)$$

denote the projection of x on \mathcal{A} . Note that $\text{dist}(x, \mathcal{A}) = \|x - \pi_{\mathcal{A}}(x)\|$.

Let us define the intervals

$$\mathcal{H}_1 = \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_1 = 0\}, \quad (75)$$

$$\mathcal{H}_1^- = \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_1 = 0, (x - (1, 1)^\top)^\top \lambda_1^\perp \in [-1, 0]\} \quad (76)$$

$$\mathcal{H}_2 = \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_2 = 0\}, \quad (77)$$

$$\mathcal{H}_2^- = \{x \in [0, \infty)^2 : (x - (1, 1)^\top)^\top \lambda_2 = 0, (x - (1, 1)^\top)^\top \lambda_2^\perp \in [-1, 0]\}, \quad (78)$$

such that the boundary of \mathcal{S} is $\mathcal{H}_1^- \cup \mathcal{H}_2^-$. Clearly, for $x \in \mathcal{S}$ we have that $\text{dist}(x, \mathcal{S}) = 0$ and $\pi_{\mathcal{S}}(x) = x$. For any $x \notin \mathcal{S}$ we have that $\pi_{\mathcal{S}}(x)$ lies on the boundary of \mathcal{S} , and therefore

$$\text{dist}(x, \mathcal{S}) = \min\{\text{dist}(x, \mathcal{H}_1^-), \text{dist}(x, \mathcal{H}_2^-)\} \quad (79)$$

$$\pi_{\mathcal{S}}(x) = \underset{\pi_{\mathcal{H}_1^-}(x), \pi_{\mathcal{H}_2^-}(x)}{\text{argmin}} \{\|x - \pi_{\mathcal{H}_1^-}(x)\|, \|x - \pi_{\mathcal{H}_2^-}(x)\|\}. \quad (80)$$

With this, it is easy to see (refer to Figure 4) that

$$\pi_{\mathcal{S}}(x) = \begin{cases} \pi_{\mathcal{H}_1^-}(x) & x \in \mathcal{A}_1 \\ \pi_{\mathcal{H}_2^-}(x) & x \in \mathcal{A}_2 \\ \pi_{\mathcal{H}_1^-}(x) = \pi_{\mathcal{H}_2^-}(x) = (1, 1)^\top & x \in \mathcal{A}_3 \\ x & x \in \mathcal{S} \end{cases}. \quad (81)$$

To write $\pi_{\mathcal{S}}(x)$ more explicitly, we define the partition of the sets \mathcal{A}_1 and \mathcal{A}_2 to

$$\begin{aligned} \mathcal{A}_1^- &= \{x \in \mathcal{A}_1 : (x - (1, 1)^\top)^\top \lambda_1^\perp < -1\}, \quad \mathcal{A}_1^+ = \mathcal{A}_1 \setminus \mathcal{A}_1^-, \\ \mathcal{A}_2^- &= \{x \in \mathcal{A}_2 : (x - (1, 1)^\top)^\top \lambda_2^\perp < -1\}, \quad \mathcal{A}_2^+ = \mathcal{A}_2 \setminus \mathcal{A}_2^-, \end{aligned} \quad (82)$$

such that

$$\pi_{\mathcal{S}}(x) = \begin{cases} (0, \Delta)^\top = (1, 1)^\top - \lambda_1^\perp & x \in \mathcal{A}_1^- \\ (1, 1)^\top + ((x - (1, 1)^\top)^\top \lambda_1^\perp) \cdot \lambda_1^\perp & x \in \mathcal{A}_1^+ \\ (1, 1)^\top + ((x - (1, 1)^\top)^\top \lambda_2^\perp) \cdot \lambda_2^\perp & x \in \mathcal{A}_2^+ \\ (\Delta, 0)^\top = (1, 1)^\top - \lambda_2^\perp & x \in \mathcal{A}_2^- \\ (1, 1)^\top & x \in \mathcal{A}_3 \\ x & x \in \mathcal{S} \end{cases}. \quad (83)$$

Let $\bar{\varphi}_t$ be as defined in (68), and $\eta_t = \pi_{\mathcal{S}}(\bar{\varphi}_t)$. Furthermore, let

$$\rho_t = \min_{s \in \mathcal{S}} \|\bar{\varphi}_t - s\|^2 = \|\bar{\varphi}_t - \pi_{\mathcal{S}}(\bar{\varphi}_t)\|^2 = \|\bar{\varphi}_t - \eta_t\|^2. \quad (84)$$

We have

$$\bar{\varphi}_{t+1} = \frac{1}{t+1} \sum_{i=1}^{t+1} f(r_i, q_i) = \frac{t}{t+1} \bar{\varphi}_t + \frac{1}{t+1} f(r_{t+1}, q_{t+1}). \quad (85)$$

Consequently,

$$\begin{aligned} \rho_{t+1} &= \min_{s \in \mathcal{S}} \|\bar{\varphi}_{t+1} - s\|^2 \\ &\leq \|\bar{\varphi}_{t+1} - \eta_t\|^2 \end{aligned} \quad (86)$$

$$= \left\| \frac{t}{t+1} (\bar{\varphi}_t - \eta_t) + \frac{1}{t+1} (f(r_{t+1}, q_{t+1}) - \eta_t) \right\|^2 \quad (87)$$

$$= \frac{1}{(t+1)^2} (t^2 \rho_t + \|f(r_{t+1}, q_{t+1}) - \eta_t\|^2 + 2t(\bar{\varphi}_t - \eta_t)^\top (f(r_{t+1}, q_{t+1}) - \eta_t)). \quad (88)$$

We now show that $(\bar{\varphi}_t - \eta_t)^\top (f(r_{t+1}, q_{t+1}) - \eta_t) \leq 0$. We first claim that for all $\bar{\varphi}_t \in \mathcal{A}_1$ it holds that $f(r_{t+1}, q_{t+1}) - \eta_t = \beta_1 \lambda_1^\perp$ for some $\beta_1 \in \mathbb{R}$. To see this, note that $\eta_t = \pi_{\mathcal{S}}(\bar{\varphi}_t) = (1, 1)^\top + \gamma \lambda_1^\perp$ for all $\bar{\varphi}_t \in \mathcal{A}_1$ for some $\gamma \in [-1, 0]$. Thus,

$$\lambda_1^\top (f(r_{t+1}, q_{t+1}) - \eta_t) = (\lambda_1(0), \lambda_1(1)) \left[\left(\frac{q_{t+1}(0)}{\lambda_1(0)}, \frac{q_{t+1}(1)}{\lambda_1(1)} \right)^\top - (1, 1)^\top - \gamma \lambda_1^\perp \right] = 0, \quad (89)$$

where the last equality follows since both λ_1 and q_{t+1} are probability vectors. We claim further, that for $\bar{\varphi}_t \in \mathcal{A}_1^-$, we have that $\beta_1 \geq 0$. This follows since both $f(r_{t+1}, q_{t+1})$ and η_t are on the line \mathcal{H}_1 , but $\eta_t = (0, \Delta)^\top$ is the leftmost point on this line. Similarly, for all $\bar{\varphi}_t \in \mathcal{A}_2$ it holds that $f(r_{t+1}, q_{t+1}) - \eta_t = \beta_2 \lambda_2^\perp$, and furthermore, $\beta_2 \geq 0$ if $\eta_t \in \mathcal{A}_2^-$.

Next, note that (recalling (83))

$$\frac{\bar{\varphi}_t - \eta_t}{\|\bar{\varphi}_t - \eta_t\|_1} = \begin{cases} \frac{\bar{\varphi}_t - (0, \Delta)^\top}{\|\bar{\varphi}_t - (0, \Delta)^\top\|_1} = a_{1-} \cdot \lambda_1 + b_{1-} \cdot \lambda_1^\perp & \bar{\varphi}_t \in \mathcal{A}_1^- \\ \lambda_1 & \bar{\varphi}_t \in \mathcal{A}_1^+ \\ \lambda_2 & \bar{\varphi}_t \in \mathcal{A}_2^+ \\ \frac{\bar{\varphi}_t - (\Delta, 0)^\top}{\|\bar{\varphi}_t - (\Delta, 0)^\top\|_1} = a_{2-} \cdot \lambda_2 + b_{2-} \cdot \lambda_2^\perp & \bar{\varphi}_t \in \mathcal{A}_2^- \\ \frac{\bar{\varphi}_t - (1, 1)^\top}{\|\bar{\varphi}_t - (1, 1)^\top\|_1} & \bar{\varphi}_t \in \mathcal{A}_3 \\ (0, 0)^\top & \bar{\varphi}_t \in \mathcal{S} \end{cases}, \quad (90)$$

where $b_{1-}, b_{2-} \leq 0$ (and also $a_{1-}, a_{2-} \geq 0$). Thus,

$$\begin{aligned} \frac{(\bar{\varphi}_t - \eta_t)^\top (f(r_{t+1}, q_{t+1}) - \eta_t)}{\|\bar{\varphi}_t - \eta_t\|_1} &= \begin{cases} (a_{1-} \cdot \lambda_1 + b_{1-} \cdot \lambda_1^\perp)^\top \beta_1 \lambda_1^\perp & \bar{\varphi}_t \in \mathcal{A}_1^- \\ \lambda_1^\top \beta_1 \lambda_1^\perp & \bar{\varphi}_t \in \mathcal{A}_1^+ \\ \lambda_2^\top \beta_2 \lambda_2^\perp & \bar{\varphi}_t \in \mathcal{A}_2^+ \\ (a_{2-} \cdot \lambda_2 + b_{2-} \cdot \lambda_2^\perp)^\top \beta_2 \lambda_2^\perp & \bar{\varphi}_t \in \mathcal{A}_2^- \\ q_{t+1}(1) + q_{t+1}(2) - \frac{\bar{\varphi}_t - (1, 1)^\top}{\|\bar{\varphi}_t - (1, 1)^\top\|_1} (1, 1) & \bar{\varphi}_t \in \mathcal{A}_3 \\ 0 & \bar{\varphi}_t \in \mathcal{S} \end{cases} \\ &= \begin{cases} b_{1-} \cdot \beta_1 \cdot \|\lambda_1^\perp\|^2 & \bar{\varphi}_t \in \mathcal{A}_1^- \\ 0 & \bar{\varphi}_t \in \mathcal{A}_1^+ \\ 0 & \bar{\varphi}_t \in \mathcal{A}_2^+ \\ b_{2-} \cdot \beta_2 \cdot \|\lambda_2^\perp\|^2 & \bar{\varphi}_t \in \mathcal{A}_2^- \\ 0 & \bar{\varphi}_t \in \mathcal{A}_3 \\ 0 & \bar{\varphi}_t \in \mathcal{S}, \end{cases} \end{aligned} \quad (91)$$

and this is non-positive in all regions $\mathcal{A}_1^-, \mathcal{A}_1^+, \mathcal{A}_2^-, \mathcal{A}_2^+, \mathcal{A}_3, \mathcal{S}$.

Consequently, (88) gives

$$\rho_{t+1} \leq \frac{1}{(t+1)^2} (t^2 \rho_t + \|f(r_{t+1}, q_{t+1}) - \eta_t\|^2). \quad (92)$$

Telescoping (92), we obtain that

$$\rho_t \leq \frac{1}{t^2} \sum_{i=1}^t \|f(r_i, q_i) - \eta_{i-1}\|^2, \quad (93)$$

where we defined $\eta_0 = (0, 0)^\top$. We turn to develop an upper bound on $\|f(r_t, q_t) - \eta_{t-1}\|^2$ that holds for all t . We first write

$$\|f(r_t, q_t) - \eta_{t-1}\|^2 \leq 2\|f(r_t, q_t) - \eta_{t-1}\|_\infty^2 \leq 2(\|f(r_t, q_t)\|_\infty + \|\eta_{t-1}\|_\infty)^2. \quad (94)$$

Since $\eta_{t-1} \in \mathcal{S}$, we have that $\|\eta_{t-1}\|_\infty \leq \Delta$. Note that for any q_t in the simplex

$$\|f(r_t, q_t)\|_\infty \leq \max\left\{\frac{1}{r_t(1)}, \frac{1}{r_t(2)}\right\} = \frac{1}{\min\{r_t(1), r_t(2)\}}. \quad (95)$$

For $r_t = \lambda_1$, as well as $r_t = \lambda_2$ we have that $\min\{r_t(1), r_t(2)\} = \frac{\Delta-1}{\Delta}$ (recall that $\Delta \in (1, 2]$). For any $\bar{\varphi}_{t-1} \in \mathcal{A}_3$, the vector $r_t = \frac{\bar{\varphi}_{t-1} - (1, 1)^\top}{\|\bar{\varphi}_{t-1} - (1, 1)^\top\|_1}$ is a convex combination of λ_1 and λ_2 , and consequently $\min\{r_t(1), r_t(2)\} \geq \frac{\Delta-1}{\Delta}$ for all such vectors. For $r_t = (1/2, 1/2)$ we clearly have $\min\{r_t(1), r_t(2)\} = \frac{1}{2} \geq \frac{\Delta-1}{\Delta}$. Thus, by (95), for any r_t chosen by the algorithm (69), and any q_t in the simplex, it holds that

$$\|f(r_t, q_t)\|_\infty \leq \frac{\Delta}{\Delta-1}. \quad (96)$$

Substituting this into (94), we obtain

$$\|f(r_t, q_t) - \eta_{t-1}\|^2 \leq 2\left(\frac{\Delta}{\Delta-1} + \Delta\right)^2 = \frac{2\Delta^4}{(\Delta-1)^2}, \quad (97)$$

so that, by (93),

$$\rho_t \leq \frac{2}{t} \frac{\Delta^4}{(\Delta-1)^2}. \quad (98)$$

Finally,

$$\|\bar{\varphi}_t\|_\infty \leq \|\bar{\varphi}_t - \eta_t\|_\infty + \|\eta_t\|_\infty \leq \sqrt{\rho_t} + \|\eta_t\|_\infty \leq \sqrt{\frac{2}{t} \frac{\Delta^2}{\Delta-1}} + \Delta. \quad (99)$$

■

VI. CONCLUSIONS AND OPEN PROBLEMS

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.

There are many follow-up questions arising from this work. The most obvious one is whether the optimal solution for the general online bookmaking game ($m > 2$) can be obtained? To this end, one has to balance m -ary trees, $m > 2$, and extending our techniques to handle the $m > 2$ case seems to be a highly non-trivial task. Our algorithm is designed for a known horizon T , that is, the house knows how many rounds will be played, and the optimal strategy highly depends on T . Can we design a strategy that does not depend on the horizon T and is near-optimal for all T ? Note that the Blackwell-approachability-inspired algorithm we developed in Section V does not depend on the horizon. However, it attains a loss of $T + \mathcal{O}(T^{3/4})$ for all T , whereas the optimal loss is $L_T^* = T + T^{1/2}$. Finally, the tree bi-balancing technique we developed here may be applicable to other problems and loss functions. The key property our construction used is that the pair of values we can get for a tree of depth one is parametrized as $(x, f(x))$ where the function $f(x)$ is an involution. We believe that other important problems should have this property.

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] D. P. Foster and R. V. Vohra, "Calibrated learning and correlated equilibrium," *Games and Economic Behavior*, vol. 21, no. 1-2, pp. 40–55, 1997.
- [14] —, "Asymptotic calibration," *Biometrika*, vol. 85, no. 2, pp. 379–390, 1998.
- [15] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge university press, 2006.
- [16] D. P. Foster, "A proof of calibration via blackwell's approachability theorem," *Games and Economic Behavior*, vol. 29, no. 1-2, pp. 73–78, 1999.
- [17] S. Mannor and G. Stoltz, "A geometric proof of calibration," *Mathematics of Operations Research*, vol. 35, no. 4, pp. 721–727, 2010.
- [18] J. Abernethy, P. L. Bartlett, and E. Hazan, "Blackwell approachability and no-regret learning are equivalent," in *Proceedings of the 24th Annual Conference on Learning Theory*. JMLR Workshop and Conference Proceedings, 2011, pp. 27–46.
- [19] V. Vovk, A. Takemura, and G. Shafer, "Defensive forecasting," in *International Workshop on Artificial Intelligence and Statistics*. PMLR, 2005, pp. 365–372.
- [20] S. M. Kakade and D. P. Foster, "Deterministic calibration and nash equilibrium," in *International Conference on Computational Learning Theory*. Springer, 2004, pp. 33–48.
- [21] D. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 2001, vol. 1 and 2.
- [22] N. Merhav and M. Feder, "Universal prediction," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2124–2147, 1998.
- [23] R. Krichevsky and V. Trofimov, "The performance of universal encoding," *IEEE Transactions on Information Theory*, vol. 27, no. 2, pp. 199–207, 1981.
- [24] Y. Polyanskiy and Y. Wu, "Information theory: From coding to learning," *Book draft*, 2022.
- [25] G. Farina, "Blackwell approachability and regret minimization on simplex domains," *Lecture Notes for CMU 15-888: Computational Game Solving*, 2021. [Online]. Available: https://www.mit.edu/~gfarina/2021/15888f21_L04_blackwell_rm/L04_blackwell_rm.pdf