Log-Optimal Portfolio Construction for Binary Options with Combinatorial Constraints

A Preprint

Jeff Decary
Department of Operations and Information Management
University of Connecticut Storrs, CT 06268
Jeff.Sylvestre-Decary@uconn.edu

David Bergman
Department of Operations and Information Management
University of Connecticut Storrs, CT 06268
David.Bergman@uconn.edu

Bin Zou
Department of Mathematics
University of Connecticut Storrs, CT 06268
Bin.Zou@uconn.edu

September 20, 2025

Abstract

We study the problem of optimal wealth allocation across independent binary options with known payouts, aiming to maximize log utility under practical constraints. This general framework arises in settings such as prediction markets (e.g., *Kalshi* and *Polymarket*), financial event contracts (e.g., *Nadex*), and sports betting (e.g., *Draftkings* and *FanDuel*). The Kelly Criterion provides a classical solution for the bet sizing of a single binary option, and numerous papers have explored extensions to multiple binary options. Our work expands on this body of research by investigating how to incorporate combinatorial constraints into the model, including limits on the number of binary options to select in a portfolio, a requirement that arises in many settings. These constraints considerably increase the computational complexity of the problem, thereby necessitating advanced solution methodologies. To address this challenge, we develop a logic-based Benders decomposition algorithm that provides a scalable and computationally efficient solution framework. Although broadly applicable, we focus on sports betting due to its market scale and unique inclusion of *parlay options*. We also study how sportsbooks can make slight modifications to parlay pricing so that optimal allocations do not include parlay options, even though such options may remain attractive to bettors.

1 Introduction

Optimal investment across uncertain binary outcomes with known payoffs is a core problem in decision theory, operations research, and financial optimization. From elections to geopolitical events, economic reports, or sports, individuals increasingly wager capital on binary outcomes. These are examples of *binary options*, offering fixed payouts for correct predictions and loss of stake otherwise. Given a finite bankroll and many such opportunities, how should a trader—or more broadly, a decision-maker—allocate wealth to maximize long-term returns?

This problem naturally arises in prediction markets and event-based exchanges like *Kalshi* and *Polymarket*, where users bet on outcomes ranging from inflation to elections (see Figure 1). A related platform is *Nadex*, a regulated exchange offering binary contracts that pay \$0 or \$100 depending on the outcome. Contracts trade between 0 and 100, reflecting market-implied probabilities.

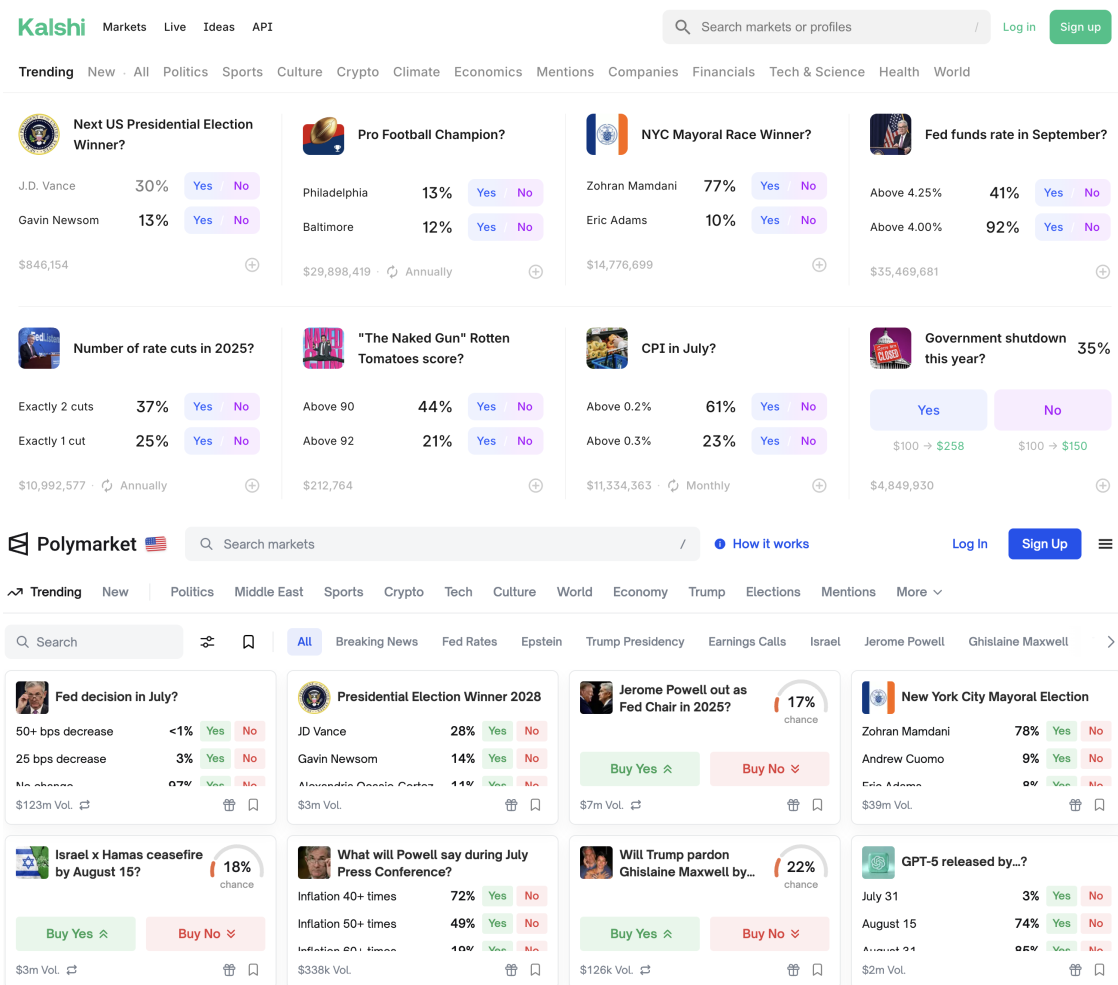


Figure 1: Example prediction markets for binary-outcome wagering offered on Kalshi and Polymarket offered at 6:30AM on July 30, 2025.

Another rapidly growing application is *sports betting*, a rich domain due to its large market size, high-quality data, and variety of wagering options. Platforms such as *FanDuel* and *DraftKings* offer thousands of binary options daily, including point spreads, moneylines, and player props. Sports betting provides a particularly interesting setting for inquiry, study, and investigation for at least three reasons: market inefficiency, parlay options, and combinatorial requirements.

1. *Market Inefficiency*: Sports betting markets remain partially inefficient (Birge et al. 2021, Simon 2024), allowing informed bettors to exploit mispriced bets with positive expected return on investment (ROI). A key reason is that sportsbooks do not price bets solely to reflect true event probabilities; instead, they manage risk and maximize profit by anticipating bettor behavior. For example, lines may be shaded to exploit public bias (e.g., toward popular teams) or to balance exposure across outcomes, not eliminate arbitrage. Odds also move in response to betting volume rather than new information, creating transient mispricings. These dynamics—along with latency in line updates and variation across sportsbooks—lead to persistent deviations from efficient pricing. See Levitt (2004), Durand et al. (2021), Moskowitz (2021), Simon (2024) for related evidence.

2. *Parlay Options*: A key feature distinguishing sports betting from other binary-option settings is the widespread use of *parlays*—bundles of multiple wagers that pay only if all legs win. While appealing to bettors for their high payouts, parlays pose major modeling challenges due to exponentially many combinations and interdependent outcomes.

3. *Combinatorial Requirements*: Although bettors face thousands of opportunities, they are often limited in how many bets they can place. Parlay bets further increase complexity by introducing exponentially many combinations. Additionally, most platforms lack API access, forcing manual and expedient bet placement. Identifying high-quality portfolios under cardinality constraints is thus a practical and computational challenge.

In 2024, the total handle (amount wagered) on sports outcomes was estimated at $148.7B (Docsports.com 2025), reflecting how sports betting has growth as a mainstream financial activity (Baker et al. 2024). This surge has fueled academic interest in several elements related to sports betting. One key focus is, of course, prediction estimation, which comes with a vast body of work (see, e.g., Horvat and Job (2020) for a survey). The other major focus is on optimal strategies in various formats, including, for example, *March Madness* (Kaplan and Garstka 2001, Decary et al. 2024) and *Daily Fantasy Sports* (Hunter et al. 2016, Haugh and Singal 2021, Bergman et al. 2023). Estimates for the total handle for these two sports betting formats for 2025 are $3.1B (American Gaming Association 2025) and $14.29B (Market Research Future 2025). With parlays accounting for approximately 25% of all wagers (Sayre and Simonetti 2025) (other figures have that number even higher, for example, Lutz (2024) estimates that parlays make up roughly 70% of all NFL and NBA bets on *FanDuel*) the parlay betting handle alone can be approximated at $37.175B annually. These figures indicate that studying optimal betting strategies for binary options in sports betting, with a focus on parlays, is more broadly of interest than other games, including *March Madness* and *Daily Fantasy Sports*, in addition to other styles of competition with lower levels of participation, for example in *NFL Survivor Pools* (Bergman and Imbrogno 2017, Imbrogno and Bergman 2022).

Selecting the optimal size of a wager involves tradeoffs between expected growth and risk of ruin. The literature on binary option bet sizing is vast. The *Kelly Criterion* (Kelly Jr. 1956) prescribes the exact wealth fraction to wager on a single binary option to maximize long-term growth, and spawned the entire research area. Despite the simplicity of each individual binary option, the portfolio allocation problem becomes complex when multiple wagers are available simultaneously.

In the absence of parlays, work on optimizing in the presence of multiple single options began in Breiman (1961), with numerous modeling and algorithmic contributions and extensions (see for example (Samuelson 1971, Miller 1975, Algoet and Cover 1988, Bell and Cover 1988, MacLean et al. 1992, Whitrow 2007, Thorp 2008, Baker and McHale 2013, Busseti et al. 2016, Rujeerapaiboon et al. 2016, Sun and Boyd 2018, Han et al. 2019, Mercurio et al. 2020)). As an example, Mercurio et al. (2020) study log optimal betting with objective function penalties leading to diversification, with application both to binary options in *Nadex* and on sports gambling. What this entire body of work does not allow for is the imposition of hard combinatorial constraints, which is essential in practice when bettors are unable or unwilling to allocate capital across all available opportunities.

There is a small but important body of work that has explored optimal bet sizing in the presence of parlays (Grant et al. 2008, 2019). These studies derive closed form expressions for log optimal betting strategies, but these formulations require simultaneously placing all possible parlay combinations. For instance, if a bettor identifies 100 profitable individual bets, the optimal strategy would involve placing $2^{100} - 1$ distinct wagers, an approach that is computationally and operationally infeasible. Moreover, their analysis assumes that all bets are mutually independent. In practice, sportsbooks may allow parlays involving dependent bets (e.g., multiple bets within the same game), but the odds in such cases are not multiplicative and often vary widely. As a result, a more tractable and realistic approach is to extract high-quality portfolios only allowing a choice for which the collection of bets in the portfolio are all mutually independent.

Our main contributions are as follows. We provide a formal study and analysis of log optimal bet sizing on binary options in sportsbooks in the presence of parlays and practical constraints. We formulate the problem as a mixed-integer non-linear optimization model. Our model incorporates combinatorial constraints to ensure practical feasibility, and allows for dependent binary options bets to be present for consideration, while enforcing that selected portfolios consist only of mutually independent bets. To solve our model at scale, we design a logic-based Benders decomposition algorithm that performs efficiently in large instances, and we show how additional constraints—such as limiting downside risk—can be seamlessly integrated into the framework. We evaluate our approach on both synthetic and real-world data, using outcome probabilities derived from a proprietary sports betting model.

Although our work is centered on sports betting, the underlying framework is broadly applicable to other binary-outcome wagering settings. In particular, our model can easily be adapted to exclude parlay options entirely, thereby reducing the problem to a setting equivalent to that of other references listed above (e.g., Mercurio et al. (2020)). The core insight that fuels the algorithm is a provably optimal bound on the expected logarithmic growth of wealth for any allocation on any collection of independent binary options, regardless of whether parlay betting is available. Our approach contributes even to this simplified setting, in that we introduce the ability to impose combinatorial constraints.

Finally, our analysis also includes an economic examination of sportsbook pricing behavior. We show how minor reductions in parlay payouts can shift optimal bettor behavior toward single-option wagers, insights that may inform policy and platform design.

The remainder of the paper is organized as follows. Section 2 introduces the notation and problem scope. Section 3 formally defines the problem we study. Section 4 presents benchmark methods. Section 5 discusses structural results. Section 6 introduces our LBBD algorithm. Section 7 presents computational experiments, and Section 8 analyzes sportsbook parlay pricing. Section 9 concludes.

2 Preliminaries

Sportsbooks (or other exchanges) offer to bettors (or more generally participants) a set of n_s *single-option bets*, $\mathcal{S} = \{1, 2, \dots, n_s\}$. Each binary option $s \in \mathcal{S}$ is associated with a future real-world sport event outcome (or any real-world outcome) E_s and has *decimal odds* (which we will refer to simply as *odds* or alternatively *price*) $d_s > 1$. We assume that E_s is a random binary event that is either `true` or `false`, which encompasses a large portion of the betting market. A bettor plays against the sportsbook by selecting an option s and betting x_s on that outcome. If $E_s = \texttt{true}$, the bet is said to be successful, and the bettor receives $d_s x_s$ return for a profit of $d_s x_s - x_s$. If $E_s = \texttt{false}$, the bet is said to be unsuccessful, and the bettor receives 0 return for a profit of $-x_s$. We denote by p_s the probability that $E_s = \texttt{true}$, with $0 < p_s < 1$ to exclude trivial cases. We let $\tilde{\xi}_s$ be the Bernoulli random variable indicating if bet s succeeds, i.e., indicating when E_s is `true`. We can then let $\tilde{\xi} = \left(\tilde{\xi}_1, \dots, \tilde{\xi}_{n_s}\right)$ be the associated random vector.

A bettor can also construct a *parlay*, which is any non-empty subset of single-option bets. Let $\mathcal{O}(\mathcal{S}) := 2^{\mathcal{S}} \setminus \{\emptyset\}$ be the collection of all non-empty subsets of $\mathcal{S}$. For any $O \in \mathcal{O}(\mathcal{S})$, each $s \in O$ is called a *leg* of the parlay. A parlay is successful if all of its legs are successful, and it is unsuccessful otherwise. A parlay consisting of only one single-option bet is equivalent to that single-option bet.

Sportsbooks use algorithms to price constructed parlays. Pricing becomes particularly challenging when the legs of a parlay are derived from events within the same game—such parlays are referred to as *same-game parlays*. Bettors are typically unaware of the specific pricing mechanism employed by the sportsbook for these types of parlays. Conversely, when the events originate from different games, they are (approximately) independent, and parlays with independent legs are often priced as the product of the individual odds.

Let $\mathcal{H}$ be the set of games indexed by h, and let $\mathcal{S}(h)$ denote the set of single-option bets available in game h (a slight abuse of notation, as $\mathcal{S}$ refers to all single options, while $\mathcal{S}(h)$ returns those for a specific game). Let $h(s) \in \mathcal{H}$ denote the game associated with single-option s. We define the set of independent parlays as:

$$\mathcal{O}^I(\mathcal{S}) \coloneqq \{O \in \mathcal{O}(\mathcal{S}) : \forall s', s'' \in O, s' \neq s'' \to h(s') \neq h(s'')\}.$$

For every parlay $O \in \mathcal{O}^I(\mathcal{S})$, the odds d_O of this parlay is given by $d_O = \prod_{s \in O} d_s$. Note that if the number of the legs of a parlay grows very large, a sportbook may deviate from this prescriptive price assignment, but we will assume for this paper that the prices are assigned as defined above. As with single-option bets, a wager of x_O on independent parlay O results in return $d_O x_O$ if the option is successful and 0 otherwise, i.e., profit $d_O x_O - x_O$ if the option is successful and $-x_O$ otherwise.

Let p_O be the probability that parlay O is successful, i.e., $p_O \coloneqq \mathbb{P}\left(\bigcap_{s \in O} E_s\right)$. For independent parlays, $p_O = \prod_{s \in O} p_s$. For parlays that are not independent, the probability p_O depends on the correlation between the events associated with each leg. The expected profit of betting x_O on O is $p_O \left(d_O x_O - x_O\right) + (1 - p_O)\left(-x_O\right) = x_O \left(p_O d_O - 1\right)$. Therefore, for independent parlay O, we denote the *expected net profit* (NP) of an independent parlay as $\pi_O \coloneqq p_O d_O - 1$ and the *expected return on investment* (ROI) as $\rho_O \coloneqq 100\left(p_O d_O - 1\right)$. A parlay is said to be *profitable* if $\pi_O > 0$. Due to independence, O is profitable if each of its legs is profitable (the converse is not necessarily true, i.e., there can be options that are profitable with some legs being unprofitable). For the remainder of the paper, we assume that all single-option bets are profitable (see Table 1 for real-world examples).

Table 1 displays two *DraftKings* single-option bets taken on September 10th, 2024. These probabilities are taken from a proprietary machine learning model owned by one of the authors. E_s indicates the event of the wager and reads as *Player Name (Team) Player Prop-Over/Under Count* where Player prop represents a statistic associated with the respective player, and over/under represent the binary condition defining a successful single-option bet. The first single-option bet from Table 1 features *Bryan Ramos* of the Chicago

s	E_s	d_s	p_s	ρ_s (%)
1	Bryan Ramos (Chicago WS) batter strikeouts-Under 0.5	2.3	0.48	11
2	Joey Gallo (Washington) Batter singles-Over 0.5	4.2	0.27	15

Table 1: *DraftKings single-option bets* - 09-10-2024

White Sox facing the Cleveland Guardians. To win the first single-option bet ($s = 1$), *Bryan Ramos* must record fewer than 0.5 strike outs. Similarly, the second single-option bet involves *Joey Gallo* of the Washington Nationals, who are playing against the Atlanta Braves. A bettor wins the second single-option bet ($s = 2$) if *Joey Gallo* records at least one single. Since these single-option bets are from different games, they are independent.

Let O denote the set of single-option bets in a given parlay, with d_O, p_O, and ρ_O representing its decimal odds, success probability, and ROI, respectively. For example, combining the two bets in Table 1 yields a parlay $O = \{1, 2\}$ with $d_O = 4.2 \cdot 2.3 = 9.6$, $p_O = 0.13$, and $\rho_O = 27$. A parlay succeeds only if all legs win, making its success probability the product of individual probabilities. While combining bets can increase ROI, it probability of hitting is diminished.

3 Problem Description

A bettor seeking to maximize long-term wealth (under the assumptions listed below, which are commonly accepted in the literature) does so by maximizing the geometric mean growth rate, which is equivalent to maximizing the expected logarithmic growth of wealth. This approach accounts for the compounding effect of multiple bets and balances risk and reward over time, ensuring the optimal balance between aggressive betting (for growth) and conservative betting (to avoid ruin). The logarithmic function is particularly useful because it reflects the diminishing marginal utility of wealth—large losses are more harmful than equivalent gains are beneficial, which is key to preserving and growing wealth in the long run.

Expected logarithmic growth rate is generally adopted in the literature because, under the following assumptions, it translates to maximizing long-term growth rate (e.g., Maclean et al. (2010)):

- **Repeated bets**: The same bet can be made over and over per decision period;
- **Known probabilities**: Accurate knowledge of the probability of winning each bet is available;
- **Reinvestment**: Winnings are reinvested each period; and
- **Finite bankroll**: The bettor has a limited initial wealth.

We assume that a bettor's initial wealth is \$1. In this way, we can recast seeking x_O (in dollar amount) for every option into finding $f_O \geq 0$, the *fraction* of wealth to wager on every option, with the constraint that the sum of the fractions are less than or equal to 1.

We call $\mathbf{f} := (f_O)_{O \in \mathcal{O}^I(\mathcal{S})}$ a (betting) strategy and denote the set of all admissible strategies by $\mathcal{F}(\mathcal{S}) := \{\mathbf{f} = (f_O)_{O \in \mathcal{O}^I(\mathcal{S})} \mid f_O \geq 0$ for all $O \in \mathcal{O}^I(\mathcal{S})$ and $\sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O \leq 1 - \epsilon\}$ where $\epsilon > 0$ is a small constant. Note that the non-negative constraint on each f_O means that the bettor can wager on a single event or parlay, but cannot "short sell" it; the inequity constraint on the sum of fractions implies that the bettor does not borrow money to make bets.

Let $W(\mathbf{f}; \tilde{\xi})$ be the random variable for the bettor's wealth under an admissible strategy $\mathbf{f}$. For brievety, we denote it as $W(\mathbf{f})$. Since the bettor aims to optimize the expected logarithmic utility of her terminal wealth, the bettor's optimization problem is

$$\max_{\mathbf{f} \in \mathcal{F}(\mathcal{S})} \mathbb{E}\left[\ln\left(W(\mathbf{f})\right)\right]. \tag{LG}$$

Dependencies arise not only between single-option bets in the same game, but also across parlays, which may share legs or reference events from the same game. As a result, even when wagers are placed only on independent options, expressing LG in terms of probabilities and odds requires conditioning on all possible outcomes.

To this end, we introduce additional notation to represent the probability that a collection of single-option bets $S' \subseteq \mathcal{S}$ are successful and the collection of the other single-option bets $\mathcal{S} \setminus S'$ are unsuccessful. We denote these probabilities in the following way: for all $S' \subseteq \mathcal{S}$,

$$q\left(\mathcal{S}, S'\right) := \mathbb{P}\left(\bigcap_{s \in S'} E_s \bigcap_{s' \in \mathcal{S} \setminus S'} E_{s'}^c\right). \tag{1}$$

Therefore, we rewrite the problem LG as

$$\max_{\mathbf{f} \geq 0} \sum_{S' \subseteq \mathcal{S}} q\left(\mathcal{S}, S'\right) \ln\left(1 - \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O + \sum_{O \in \mathcal{O}^I(S')} f_O d_O\right) \text{ subject to } \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O \leq 1 - \epsilon. \tag{LG'}$$

The sum expression in the objective considers all possible outcomes of options, where we sum over all subsets S' of $\mathcal{S}$; note that by the definition of $q(\cdot)$ in (1), all single-option bets in S' are successful, and all the remaining single-option bets in $\mathcal{S} \setminus S'$ are unsuccessful. We then take the log of the wealth for that outcome, where the term $1 - \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O$ represents the initial wealth less the amount wagered, and $\sum_{O \subseteq \mathcal{O}^I(S')} f_O d_O$ represents the return from the successful bets that are placed. (Indeed, all single-option bets in S' are successful, so are parlays with legs in S'.) The constraints limit the amount wagered to be less than 1-ϵ (where $\epsilon > 0$ is a small safety margin) and each individual bet to be greater than or equal to 0. In practice we set $\epsilon = 0.01$ to avoid numerical issues.

A well-known classical result (Kelly Jr. 1956) is that if a bettor is considering just one single-option bet $s \in \mathcal{S}$, the so-called Kelly Criterion

$$k_s := \frac{p_s d_s - 1}{d_s - 1} \tag{2}$$

identifies the optimal bet fraction. By recalling $d_s > 1$ and $p_s d_s - 1 > 0$, we have $0 < k_s < 1$.

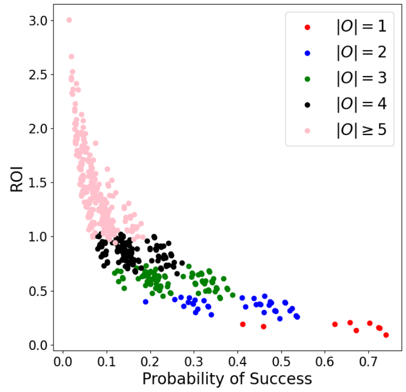
From our real-world data set on MLB betting, which we describe in full detail in Section 7.1, consider Figures 2a and 2b, which illustrate a wide-range of parlay options as scatter plots. In particular, we take nine of the most profitable bets, and calculate the probability of each single bet and derived parlay, as well as the ROI and the expected logarithmic growth rate of each. In these plots, $|O|$ is the size of the parlay, and we group large parlays to ease of visualization.

These plots reveal that the relationship between ROI and log utility casts the attractiveness of a bet in very different lights. In terms of ROI, the interaction between probabilities and decimal odds can produce deceptively high returns, especially for parlays. However, when evaluated through the lens of logarithmic growth rate, parlays with only a few legs may actually be more appealing than larger, riskier ones. The challenge intensifies when selecting among parlays that share individual legs, creating overlapping dependencies. Addressing this selection problem under practical constraints is the focus of our study, where we build on results from Grant et al. (2008).

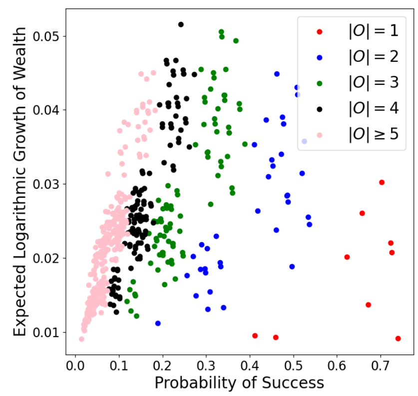
If all bets are from unique games—i.e., $\mathcal{O}^I(\mathcal{S}) = \mathcal{O}(\mathcal{S})$—and the bettor can place a wager on every possible parlay, Grant et al. (2008) shows that the optimal allocation is given by the vector $(g_O)_{O \in \mathcal{O}^I(\mathcal{S})}$, where each component is defined as

$$g_O = \prod_{s \in O} k_s \prod_{s' \in \mathcal{S} : s \notin O} \left(1 - k_{s'}\right). \tag{3}$$

When faced with two independent single-option bets, a bettor has the flexibility to wager any amount on either option, provided the total allocation does not exceed her total wealth. One straightforward strategy is to treat these bets separately, allocating an amount based on the Kelly Criterion, as long as the total



(a) ROI over probability of success



(b) Expected logarithmic bankroll growth over probability of success

Figure 2: Real-world baseball bet on 2024-08-07

remains below 1. Alternatively, the bettor could explore a parlay option and adjust her wager according to (3). Table 2 illustrates the outcomes of betting based on (2) versus (3), using the single-option and the parlay bets detailed in Tables 1. Both betting allocation strategies perform similarly, but (3) offers a slightly better expected logarithmic growth rate with a smaller allocation, as theory indicates.

Allocation Strategy	$f_{\{1\}}$	$f_{\{2\}}$	$f_{\{1,2\}}$	Total Allocation	Expected logarithmic growth of wealth
Kelly Jr. (1956)	0.084	0.046	0	0.129	0.00773
Grant et al. (2008)	0.08	0.042	0.004	0.126	0.00776

Table 2: Kelly vs Grant allocation - expected growth of wealth analysis

There are two fundamental challenges with the direct use of this result from Grant et al. (2008):

- **Exponential Growth**: the number of options is exponential; and
- **Independence**: the set of profitable bets are not generally independent.

We therefore introduce practical constraints which result in a model whose solution is practical but resolution is challenging. In particular, we define three arguments that a bettor can specify:

- $\bar{s}$: the maximum number of single-option bets the bettor will base parlays on;
- $\bar{f}$: the maximum number of bets a bettor will place; and
- $\bar{o}$: the largest size of any parlay the bettor will place.

$\bar{s}$ limits attention to a subset of single-option bets, $\bar{f}$ caps the total number of bets placed, and $\bar{o}$ restricts parlay size, reflecting that sportsbooks may reduce odds or disallow large parlays.

We thereby modify LG' to incorporate these parameters in the optimization model and formulate the *constrained logarithmic growth of wealth problem* as:

$$\begin{aligned}
\max \quad & \sum_{S' \subseteq \mathcal{S}} q(\mathcal{S}, S') \ln \left(1 - \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O + \sum_{O \in \mathcal{O}^I(S')} f_O d_O\right) \\
\text{subject to} \quad & \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O \leq 1 - \epsilon, \\
& \sum_{s \in \mathcal{S}} z_s \leq \bar{s}, \quad \sum_{O \in \mathcal{O}(\mathcal{S})} y_O \leq \bar{f} \\
& \sum_{O \in \mathcal{O}^I(\mathcal{S}): s \in O} y_O \leq z_s, && \forall s \in \mathcal{S}, \\
& y_O = 0, && \forall O \in \mathcal{O}(\mathcal{S}) : |O| > \bar{o}, \\
& z_s \in \{0, 1\}, && \forall s \in \mathcal{S}, \\
& y_O \in \{0, 1\}, 0 \leq f_O \leq y_O, && \forall O \in \mathcal{O}^I(\mathcal{S}).
\end{aligned} \tag{CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$)}$$

The objective remains identical. We add two new sets of binary variables: z_s will indicate if single-option bet s is considered for any parlay, and y_O will indicate if a wager is placed on O.

We now describe the constraints in sequence. We still limit the total allocation to be less than $1 - \epsilon$. The number of single-option bets that are considered is limited to $\bar{s}$ and the number of total bets is restricted to $\bar{f}$. For all single-option bets, any parlay containing it is required to have 0 wager if we do not select the single-option bet to consider through variable z_s. If the parlay size is beyond the limit of $\bar{o}$, we do not allow that parlay to be considered. Binary variables z_s and y_O are specified above, and constraints enforcing the connection between the wager fractions f_O and the selection variables y_O completes the model.

The focus of this paper is on understanding problem CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$). We study structural results, present a scalable optimization algorithm that utilizes these results, and run experimental results to evaluate the algorithm's performance on synthetic instances and real-world betting data.

4 Adapting Benchmark Algorithms

In this section, we adapt existing algorithms to serve as benchmarks for evaluating our proposed approach. We first tailor the *Fractional Kelly* strategy to Problem LG, considering only single-option bets, and then introduce a methodology based on the *Sample Average Approximation* framework.

4.1 Fractional Kelly

Given a set of profitable single-option bets, one approach is to exclude parlays ($\bar{o} > 1$) and apply the Kelly Criterion independently to each bet until the budget is exhausted. However, when the number of bets exceeds available capital, allocations become interdependent—even for independent bets—since each wager reduces the wealth available for others.

To address this, we build on the *Fractional Kelly* (MacLean et al. 1992) and *Shrinkage Kelly* (Baker and McHale 2013) approaches, which adjust single-option Kelly bets by scaling allocations. For a bet s, the bettor selects $\beta \in [0, 1]$ and allocates βk_s. This accounts for parameter uncertainty and promotes a more conservative risk posture. To adapt this method to our setting with multiple single-option bets, and to provide a heuristic benchmark, we propose the following procedure. We first precompute k_s for each single-option bet, and then solve for β via the following:

$$\begin{aligned}
\max_{0 \leq \beta \leq 1} \quad & \sum_{S' \subseteq \mathcal{S}} q(\mathcal{S}, S') \ln \left(1 + \beta \left(- \sum_{s \in \mathcal{S}} k_s + \sum_{s \in S'} k_s d_s\right)\right) \\
\text{subject to} \quad & \beta \sum_{s \in \mathcal{S}} k_s \leq 1 - \epsilon,
\end{aligned} \tag{FK}$$

This model has only one decision variable, β, which makes it attractive from a computational standpoint. It also reduces symmetry in the allocation when multiple bets have identical parameters. While the objective still involves an exponential number of terms the model is simple and can be solved efficiently for moderate-sized problems.

4.2 Sample Average Approximation Methodology

As an alternative benchmark, we consider the sample-average approximation (SAA), where we generate random outcomes and optimize for the mean logarithmic growth rate over those simulations. In particular, we draw n_c independent realizations $\boldsymbol{\Xi} := (\xi^1, \ldots, \xi^{n_c})$ of ξ, where $\xi^c = (\xi^c_s)_{s \in S} \in \{0,1\}^{\mathcal{S}}$, and $\xi^c_s = 1$ (resp. 0) if bet s wins (resp. loses) in scenario c. Then for each scenario c the deterministic terminal wealth is denoted by $W(\mathbf{f}; \xi^c)$. The SAA model uses the same feasibility constraints as model CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$), but replaces the objective function denoting the true expected log-wealth $\mathbb{E}\big[\ln\left(W(\mathbf{f})\right)\big]$ with its sample average approximation:

$$\begin{aligned} \max_{\mathbf{f} \in \mathcal{F}(\mathcal{S})} \widehat{\mathbb{E}\big[\ln\left(W(\mathbf{f})\right)\big]} &= \max_{\mathbf{f} \in \mathcal{F}(\mathcal{S})} \frac{1}{n_c} \sum_{c=1}^{n_c} \ln\Big(W(\mathbf{f}; \xi^c)\Big) \\ &= \max_{\mathbf{f} \in \mathcal{F}(\mathcal{S})} \frac{1}{n_c} \sum_{c=1}^{n_c} \ln\Big(1 - \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O + \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O d_O \prod_{s \in O} \xi^c_s\Big). \end{aligned} \tag{4}$$

We refer to this optimization model as (SAA($\mathcal{S}$, Ξ, $\bar{s}$, $\bar{f}$, $\bar{o}$)).

5 Theoretical Insights

While closed-form solutions exist for (a) a single-option bet and (b) for multiple bets together with all possible derived parlay options, a general closed-form solution for the optimal allocation given two independent, profitable single-option bets (without the parlay option) has yet to be determined. Assuming, for $i = 1, 2$, that $0 < p_i < 1 < d_i$ and $p_i d_i > 1$, the problem can be formulated as follows:

$$\begin{aligned} \max_{\mathbf{f}} \; & p_1 p_2 \ln\Big(1 + f_1(d_1 - 1) + f_2(d_2 - 1)\Big) + p_1(1 - p_2) \ln\Big(1 + f_1(d_1 - 1) - f_2\Big) + \\ & (1 - p_1) p_2 \ln\Big(1 - f_1 + f_2(d_2 - 1)\Big) + (1 - p_1)(1 - p_2) \ln\Big(1 - f_1 - f_2\Big). \end{aligned}$$

Little is known even for this small case. When $d_1 = d_2 = 2$, the optimal solution (Thorp 2008) is

$$f_1^* = \frac{\pi_1\left(1 - \pi_2^2\right)}{1 - \pi_1^2 \pi_2^2} \qquad \text{and} \qquad f_2^* = \frac{\pi_2\left(1 - \pi_1^2\right)}{1 - \pi_1^2 \pi_2^2}.$$

Expanding on this, we consider the scenario of two profitable identically parametrized bets.

Proposition 5.1. *Given two independent and profitable identically parametrized single-option bets with probability p and odds $d \neq 2$, the optimal allocation to each bet is*

$$f^* = \frac{\left(\pi^2 - 2(d-2)\,\pi + (d-1)\right) - \sqrt{\Delta}}{-4\,(d-1)(d-2)},$$

where $\Delta = (\pi^2 - 2(d-2)\pi + (d-1))^2 + 8(d-1)(d-2)\pi$.

A detailed proof of Proposition 5.1 is provided in Appendix A. Closed-form solutions for more general cases are unknown. As such, we proceed by establishing an upper bound for Problem CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$), which holds for any values of p and d.

5.1 Upper Bound for CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$)

We start by defining, for all $s \in \mathcal{S}$, w_s^* as the optimal expected logarithmic growth rate of wealth for just that single-option bet. This is achieved when k_s is wagered on s:

$$\begin{aligned} w_s^* &= p_s \ln\Big(1 + k_s(d_s - 1)\Big) + (1 - p_s) \ln\Big(1 - k_s\Big) \\ &= p_s \ln\Big(d_s p_s\Big) + (1 - p_s) \ln\Big(1 - k_s\Big), \end{aligned} \tag{5}$$

where k_s is given by (2).

We now state the main theoretical result of the paper, which is critical for our LBBD algorithm.

Theorem 5.2. *For any subset of single-option bets $\mathcal{S}' \subseteq \mathcal{S}$ and any given parameters $\bar{s}$, $\bar{f}$, and $\bar{o}$, the optimal value to CLG$(\mathcal{S}', \bar{s}, \bar{f}, \bar{o})$ is bounded above by $\sum_{s \in \mathcal{S}'} w_s^*$.*

Theorem 5.2 relies on a key lemma that computes the optimal value of LG in linear time, avoiding an exponential sum. This allows efficient evaluation and bounding of the maximum logarithmic growth rate over constrained subsets. We now show that the expected growth rate under allocation (3) equals the sum of optimal single-option growth rates.

Lemma 5.3. *The optimal objective value for LG is $\sum_{s \in \mathcal{S}} w_s^*$; i.e.,*

$$\max_{\mathbf{f} \in \mathcal{F}(\mathcal{S})} \mathbb{E}\left[\ln\left(W(\mathbf{f})\right)\right] = \sum_{s \in \mathcal{S}} w_s^*,$$

where w_s^* is given by (5).

Proof. Consider n_s independent single-option bets occurring sequentially. Theorem 3 in Grant et al. (2008) proves that the terminal wealth obtained by employing the Grant allocation strategy (3) is identical to the terminal wealth achieved by sequentially placing bets according to the Kelly Criterion (2). More formally, given n_s independent single-option bets, the terminal wealth under both the Grant allocation and the Kelly allocation are:

$$\left(1 - \sum_{O \in \mathcal{O}^I(\mathcal{S})} g_O + \sum_{O \in \mathcal{O}^I(\mathcal{S})} g_O(d_O x_O - 1)\right) = \prod_{s \in \mathcal{S}} \Big(1 + k_s(d_s x_s - 1)\Big).$$

This equality implies that their corresponding expected logarithmic growth rates must also be identical, i.e.,

$$\mathbb{E}\left[\ln\left(1 - \sum_{O \in \mathcal{O}^I(\mathcal{S})} g_O + \sum_{O \in \mathcal{O}^I(\mathcal{S})} g_O(d_O x_O - 1)\right)\right] = \mathbb{E}\left[\ln \prod_{s \in \mathcal{S}} \Big(1 + k_s(d_s x_s - 1)\Big)\right].$$

Furthermore, as demonstrated in Section 8 of Grant et al. (2019), one can show the following:

$$\begin{aligned}
\mathbb{E}\left[\ln \prod_{s \in \mathcal{S}} \left(1 + k_s(d_s x_s - 1)\right)\right] &= \mathbb{E}\left[\sum_{s \in \mathcal{S}} \ln\left(1 + k_s(d_s x_s - 1)\right)\right] \\
&= \sum_{s \in \mathcal{S}} \mathbb{E}\left[\ln\left(1 + k_s(d_s x_s - 1)\right)\right] \\
&= \sum_{s \in \mathcal{S}} w_s^*.
\end{aligned}$$

By combining these two results, we deduce that for any collection of simultaneous single-option bets $\mathcal{S}$, the maximum expected logarithmic wealth under Grant allocation is:

$$\mathbb{E}\left[\ln\left(W(\mathbf{f})\right)\right] = \sum_{s \in \mathcal{S}} w_s^*.$$

□

A self-contained proof is provided in Appendix B.

Corollary 5.4. *Suppose bets are ordered so that $w_1^* > w_2^* > \ldots > w_{n_s}^*$. The optimal objective value for CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$) with bounded $\bar{s}$ and no limit on $\bar{f}$ and $\bar{o}$ is $\sum_{s=1}^{\bar{s}} w_s^*$.*

Proof. For any set of single-option bets $\mathcal{S}'$ and any possible parlay of those options, Lemma 5.3 tell us that the optimal value of LG is $\sum_{s\in\mathcal{S}'} w_s^*$. The subset of size at most $\bar{s}$ that maximizes expected logarithmic growth is therefore $\{1, \ldots, \bar{s}\}$. □

This immediately proves Theorem 5.2, the main bounding mechanism in our LBBD algorithm.

5.2 Dominance

We later exploit the following simple yet effective dominance relation:

Definition 5.5. Given two profitable single-option bets, we say that single-option bet s' **dominates** single-option bet s'' if $p_{s'} > p_{s''}$ and $d_{s'} > d_{s''}$.

Proposition 5.6. *Suppose that s' dominates s''. If $(f_O^*)_{O\in\mathcal{O}(\mathcal{S})}$ is an optimal solution for LG',*

$$\sum_{O\in\mathcal{O}(\mathcal{S}):s''\in O} f_O^* > 0 \implies \sum_{O\in\mathcal{O}(\mathcal{S}):s'\in O} f_O^* > 0. \tag{6}$$

The proof of Proposition 5.6 is provided in Appendix C.

Remark 5.7. The result of Proposition 5.6 makes intuitive sense, although its proof is rather technical and lengthy (see Appendix C). The relation in (6) implies that if the optimal solution uses a single-option bet s'', then all single-option bets s' that dominate s'' will be used in the optimal solution. We can prove a weaker version of Proposition 5.6: Suppose we are given two games, h_1 and h_2, each with sets of single-option bets $\mathcal{S}(h_1)$ and $\mathcal{S}(h_2)$, respectively, both of which are in $\mathcal{S}$. Furthermore, assume that there exists an $s' \in \mathcal{S}(h_1)$ that dominates $s'' \in \mathcal{S}(h_2)$. For any $\bar{s}$, $\bar{f}$ and $\bar{o}$, if $(f_O^*)_{O\in\mathcal{O}(\mathcal{S})}$ is an optimal solution to CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$), then the following holds:

$$\sum_{O\in\mathcal{O}(\mathcal{S}):s''\in O} f_O^* > 0 \implies \exists s \in \mathcal{S}(h_1) \text{ such that } \sum_{O\in\mathcal{O}(\mathcal{S}):s\in O} f_O^* > 0.$$

6 Logic-Based Benders Decomposition

We now present an optimization algorithm that leverages the bounding mechanism from Theorem 5.2 and the dominance relation in Proposition 5.6 to address the exponential complexity of the objective in CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$). Specifically, if a solution allocates positive wagers only to options with legs in $\mathcal{T} \subseteq \mathcal{S}$, the objective can be evaluated by summing only over outcomes involving subsets of $\mathcal{T}$.

More formally, suppose a bettor is given a set of single-option bets $\mathcal{S}$, ordered so that $w_1^* > w_2^* > \ldots > w_{n_s}^*$. Let $\mathcal{T}$ be a subset of $\mathcal{S}$ and introduce a more precise notation on allocations (strategies) by $\mathbf{f}(\mathcal{T}) = \big(f_O(\mathcal{T})\big)_{O\in\mathcal{O}(\mathcal{T})}$, emphasizing that the allocation $\mathbf{f}(\mathcal{T})$ only uses single-option bets in $\mathcal{T}$ but not $\mathcal{S}\setminus\mathcal{T}$. Consider two strategies $\mathbf{f}(\mathcal{T})$ and $\mathbf{f}'(\mathcal{S})$. Suppose we have the following: $\forall O \in O(\mathcal{S})$, we have: (a) $O \not\subseteq \mathcal{T} \to f'(O) = 0$ and (b) $O \subseteq \mathcal{T} \to f(O) = f'(O)$ (i.e., the strategies are identical on options contained in $\mathcal{T}$ and strategy f' does not allocate anything outside of $\mathcal{T}$). Then $\mathbb{E}\left[\ln\left(W(\mathbf{f}'(\mathcal{S}))\right)\right] = \mathbb{E}\left[\ln\left(W(\mathbf{f}(\mathcal{T}))\right)\right]$.

$$\mathbb{E}\left[\ln\left(W\left(\mathcal{F}(\mathcal{S})\right)\right)\right] = \mathbb{E}\left[\ln\left(W\left(\mathcal{F}(\mathcal{T})\right)\right)\right].$$

We can consider the constrained logarithmic growth of wealth problem as a problem parametrized by $\mathcal{T}, \bar{s}, \bar{f}$, and $\bar{o}$ and write CLG $\left(\mathcal{T}, \bar{s}, \bar{f}, \bar{o}\right)$. This problem returns a pair $(\mathbf{f}^*(\mathcal{T}), v(\mathcal{T}))$ for every $\mathcal{T} \subseteq \mathcal{S}$, with $\mathbf{f}^*(\mathcal{T})$ being the *optimal* allocation and $v(\mathcal{T})$ the value of that solution. We can thereby recast CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$) as

$$\max_{\mathcal{T}\subseteq\mathcal{S}:|\mathcal{T}|=\bar{s}} \text{CLG}\left(\mathcal{T}, \bar{s}, \bar{f}, \bar{o}\right). \tag{CLG-B}$$

By Theorem 5.2, the optimal value for CLG-B is bounded above by $\sum_{s\in\{1,\ldots,\bar{s}\}} w_s^*$. The set $\mathcal{T}$ considered in CLG-B can be reduced by eliminating those that contain dependent single-option bets and applying the dominance rule. More specifically, we need not consider any $\mathcal{T}$ with a pair of single-option bets that are dependent, since we restrict to only independent parlays. Also, by Proposition 5.6 and Remark 5.7, we can reduce the set of $\mathcal{T}$ we consider for any parameters $\bar{s}$, $\bar{f}$, and $\bar{o}$ by employing a dominance rule: If a single-option bet $s \in \mathcal{S}(h)$ in game $h \in \mathcal{H}$ is dominated by a single-option bet $s' \in \mathcal{S}(h')$ in game $h' \in \mathcal{H}$ with $h \neq h'$, then if $s \in \mathcal{T}$, there must also be some other single-option bet $s'' \in \mathcal{S}(h')$ that is included in $\mathcal{T}$. For this purpose, for all $s \in \mathcal{S}$, let $\mathcal{D}(s)$ be the set of games for which there exists a bet that dominates s.

The LBBD algorithm is as follows:

LBBD $\left(\mathcal{S}, \mathcal{H}, \bar{s}, \bar{f}, \bar{o}, \mathcal{D}\right)$

- **Initialize**:
  - UB $:= \sum_{s\in\{1,\ldots,\bar{s}\}} w_s^*$; LB $:= 0$; $\mathcal{Q} = \emptyset$; $\mathcal{T}^* = \emptyset$
- **While** UB $>$ LB:
  - $$\mathcal{T}' =_{\mathcal{T}\subseteq\mathcal{S}} \left\{ \sum_{s\in\mathcal{T}} w_s^* \text{ s.t. } |\mathcal{T}| = \bar{s},\ \mathcal{T} \notin \mathcal{Q}, \\ \forall h \in \mathcal{H}, \forall s \in \mathcal{S}(h) \text{ if } s \in \mathcal{T} \implies \forall s' \in \mathcal{S}(h) \setminus \{s\}, s' \notin \mathcal{T}, \quad \text{(MP)} \\ \forall s \in \mathcal{S}, \text{if } s \in \mathcal{T} \Rightarrow \forall h \in \mathcal{D}(s), \exists s' \in \mathcal{S}(h) \text{ s.t. } s' \in \mathcal{T} \right\}$$
  - UB $= \sum_{s\in\mathcal{T}'} w_s^*$
  - $(\mathbf{f}^*(\mathcal{T}'), v(\mathcal{T}')) = \text{CLG}\left(\mathcal{T}', \bar{s}, \bar{f}, \bar{o}\right)$ (SP)
  - **If** $v(\mathcal{T}') >$ LB:
    - **Update Incumbant:** $\mathbf{f}^* = \mathbf{f}^*(\mathcal{T}')$ and $\mathcal{T}^* = \mathcal{T}'$
    - LB $= v(\mathcal{T}')$
  - $\mathcal{Q} := \mathcal{Q} \cup \{\mathcal{T}'\}$
- **Return** $(\mathcal{T}^*, \mathbf{f}^*)$

The algorithm is given inputs $\mathcal{S}, \mathcal{H}, \bar{s}, \bar{f}, \bar{o}$, and $\mathcal{D}$, and we recall their meanings as follows. $\mathcal{S}$ is an initial set of single option bets and $\mathcal{H}$ is the set of all sports games; for every $s \in \mathcal{S}$, $\mathcal{D}(s)$ denotes the set of games in $\mathcal{H}$ that have at least one single option that dominates s, and for every game $h \in \mathcal{H}$, $\mathcal{S}(h)$ is the set of all single options from h; the parameters, $\bar{s}$, $\bar{f}$, and $\bar{o}$, represent the bettor's practical constraints as introduced in CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$). The algorithm maintains an upper bound UB (initialized to the sum of the $\bar{s}$ highest w_s^*), a lower bound LB (initialized to 0), a family of subsets $\mathcal{Q}$ of $\mathcal{S}$ that have been examined (initialized to the empty set), and a best-known subset $\mathcal{T}^*$ (initialized to the empty set). While the upper bound exceeds the lower bound, a candidate subset $\mathcal{T}'$ is identified by solving the *master problem*, which selects the subset of size $\bar{s}$ from $\mathcal{S}$ with the highest possible upper bound (according to Theorem 5.2) among all subsets that have not yet been examined and satisfy the independent and dominance rule constraints. The upper bound is replaced by the upper bound related to set $\mathcal{T}$, and a smaller CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$) instance is solved returning $\mathbf{f}(\mathcal{T}')$ as the candidate solution with value $v(\mathcal{T}')$. If $v_{\mathcal{T}'}$ is better than the value of the best found solution, the global lower bound LB is updated and the incumbent is replaced by $\mathbf{f}(\mathcal{T}')$, also recording $\mathcal{T}'$ as the best subset. Additionally, $\mathcal{T}'$ is added to $\mathcal{Q}$.

The identification of a candidate set $\mathcal{T}'$ with maximum upper bound from Theorem 5.2 requires solving an optimization problem problem which we model as a binary optimization problem. The optimization model is as follows:

$$
\begin{aligned}
\max \quad & \sum_{s\in\mathcal{S}} w_s^* \chi_s \\
\text{s.t.} \quad & \sum_{s\in\mathcal{S}} \chi_s = \bar{s}, \\
& \sum_{s\in\mathcal{T}} \chi_s \le \bar{s} - 1, && \forall \mathcal{T} \in \mathcal{Q} \\
& \sum_{s\in\mathcal{S}(h)} \chi_s \le 1, && \forall h \in \mathcal{H} \\
& \sum_{s'\in\mathcal{S}(h)} \chi_{s'} - \chi_s \ge 0, && \forall s \in \mathcal{S}, \forall h \in \mathcal{D}(s) \\
& \chi_s \in \{0,1\}, && \forall s \in \mathcal{S}.
\end{aligned}
\tag{MP}
$$

In MP, the variables select a subset of $\mathcal{S}$ with the maximum upper bound from Theorem 5.2. The first constraint guarantees that exactly $\bar{s}$ single-option bets are selected. The second set of constraints forces the identification of a different $\mathcal{T}$ in each iteration. The third set of constraints requires that an independent collection of single-option bets is chosen. The fourth set of constraints verifies that the dominance property (i.e., Proposition 5.6 and Remark 5.7) is respected. Finally, binary constraints on χ_s complete the master problem.

Due to the recast of CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$) as CLG-B, the prohibition of cycling in the master problem, the upper bound proven in Theorem 5.2, and the dominance relationship proven in Proposition 5.6, we have the following result:

Proposition 6.1. **LBBD** $\left(\mathcal{S}, \mathcal{H}, \bar{s}, \bar{f}, \bar{o}, \mathcal{S}, \mathcal{D}\right)$ *finds an optimal solution in a finite number of steps.*

The algorithm is computationally efficient for several reasons. First, the objective is exponential in $\bar{s}$, not in n_s, so limiting $\bar{s}$ keeps subproblems tractable. Second, the gap between the best solution for a given $\mathcal{T}$ and the upper bound from Theorem 5.2 is often small, thereby requiring few iterations.

7 Computational Analysis

We now provide a detailed computational evaluation. All experiments were run on AMD Ryzen 5 3600X CPU at 3.80GHz, limited to a single thread. Optimization code was run with `Python` and all models were constructed in `Pyomo`. Non-linear models were solved with `BARON` setting `CPLEX` as the linear solver. All linear models were solved with `Gurobi`. The specific versions are as follows: `Python 3.9.18`, `Pyomo 6.7.0`, `BARON 2024.5.8`, `CPLEX 22.1.1`, and `Gurobi 10.0.0`.

The computational section is organized as follows. We begin by describing the synthetic and real-world datasets used in our analysis, along with the algorithms under consideration. We then present a comparative evaluation of wealth growth across these algorithms on real-world instances, including a detailed analysis of hyperparameter settings and the impact of a well-established risk constraint that controls drawdown. This is followed by experiments on synthetic data, which highlight the computational advantages of the proposed algorithms in terms of solution time and optimality gap. Finally, we examine the limitations of SAA, supported by empirical results.

7.1 Instances and Algorithms

We have two benchmark datasets to explore the scalability and efficacy of different algorithms. All instances will be made available upon reasonable request.

- **Synthetic instances**: We generate instances for varying n_s. Each single-option bet $s \in \mathcal{S}$ is randomly generated with $p_s \in \mathrm{U}(0.2, 0.8)$ and $\rho_s \in \mathrm{U}(5\%, 20\%)$ ($\mathrm{U}(a, b)$ denotes a uniform distribution over (a, b)). Therefore, each option is individually profitable and independent of the other single-option bets, with $d_s = \dfrac{(1+\rho_s)}{p_s}$.
- **Real-World Instances**: Real-world betting data used in this study was sourced from a proprietary betting model developed by one of the authors, focusing on prop bets for Major League Baseball

games played on August 3, 5, 6, 7, 8, and 11, 2024. The dataset was filtered to include only bets placed on *DraftKings* that met the following criteria: an estimated success probability between 0.2 and 0.8, an estimated ROI between 5% and 20%, the game had not yet started, and the lineups were confirmed. This resulted in 406 bets. To align the probability forecasts with actual outcomes, we adjusted the probability estimates for each day so that the expected ROI matched the actual ROI. Table 3 describes the best daily available bets in each of our six days of instance.

Date	E_s	p_s (Adjusted)	ρ_s (%)	k_s	w_s^*	Expected/actual wealth (% of initial) with k_s allocation
08-03-2024	Bo Naylor (Cleveland) Batter Strikeouts - Over 0.5	0.76	17	0.31	0.03	105 / 117
08-05-2024	Bo Naylor (Cleveland) Batter Hits - Under 0.5	0.59	21	0.20	0.02	104 / 80
08-06-2024	Matt Vierling (Detroit) Batter Strikeouts - Under 1.5	0.78	13	0.30	0.02	104 / 113
08-07-2024	Michael A. Taylor (Pittsburgh) Batter Strikeouts - Under 1.5	0.70	20	0.29	0.03	106 / 120
08-08-2024	Jakson Reetz (San Francisco) Batter Hits - Under 0.5	0.57	16	0.15	0.01	102 / 85
08-11-2024	Coby Mayo (Baltimore) Batter Hits - Under 0.5	0.52	9	0.09	0.00	101 / 109
					Cumulative	125 / 119

Table 3: Summary of the Best Available Bets Across Six Days of Instances

In the computational analysis, we consider five betting strategies and compare their performance in terms of both expected and actual bankroll growth. Recall that $\bar{s}$ denotes the maximum number of single-option bets used in the bettor's strategy. We denote each algorithm as follows:

- **FK**: The bettor wagers on the top $\bar{s}$ single-option bets (each day), with the allocations obtained by solving the fractional Kelly problem in (FK) using `BARON`.
- **SAA**: The bettor wagers single-option bets (each day), with the allocation determined by solving SAA($\mathcal{S}$, Ξ, $\bar{s}$, $\bar{f}$, $\bar{o}$) (we show later that **SAA** does not scale when parlays are included).
- **G**: The bettor wagers on *all* $2^{\bar{s}} - 1$ possible choices, with the allocations determined by (3).
- **B**: The bettor wagers on a *subset* of the $2^{\bar{s}} - 1$ possible choices, and the allocations are obtained by solving Problem CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$) using `BARON` and `CPLEX`.
- **L**: The bettor also wagers on a *subset* of the $2^{\bar{s}} - 1$ parlay bets, and allocations by solving the LBBD algorithm, with `Gurobi` solving the master problem (MP), and `BARON` and `CPLEX` solving the subproblems in (CLG-B).

The first three strategies—**SAA**, **FK**, and **G**—serve serve as benchmarks on, with **SAA** and **FK** focusing only on single-option bets and excluding parlays. Strategy **B** represents the direct numerical solution to Problem CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$), while **L** is our proposed solution to Problem CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$) using **LBBD** $(\mathcal{S}, \mathcal{H}, \bar{s}, \bar{f}, \bar{o}, \mathcal{S}, \mathcal{D})$. **B** is presented only on synthetic instances to provide a comparison with **L** in terms of scalability and efficacy.

7.2 Computational Results on Real-World Instances

In this section, we report on the results we obtain on real-world betting instances.

7.2.1 Main Results

Tables 4 and 5 present, respectively, the cumulative geometric mean growth of wealth $\left(e^{\mathbb{E}[\ln(W(\mathbf{f}))]}\right)$ percent, and the observed terminal wealth achieved, as a percent of initial investment, over six-day of real-world instances obtained through our tested algorithms. We imposed a limit of a 2 minute runtime on each day of data, for all algorithms. The **FK** method is evaluated based on the top $\bar{s}$ daily bets, as defined by w_s^*. For **SAA**, we set $n_c = 1,000$ and limit $\bar{o} = 1$ due to scalability limitations, which is discuss in Section 7.5. **L** is evaluated for $\bar{o} \in \{1, 2, 3\}$. We report results for all three algorithms with $\bar{s} \in \{2, 3, 4, 5, 6, 7, 8, 9\}$, and $\bar{f} = 10$. Note that we do not show results for **G** because it is not applicable to constrained settings, however we do compare against the solution quality in Section 7.2.4.

	$\bar{o}=1$			$\bar{o}=2$	$\bar{o}=3$
$\bar{s}$	**FK**	**SAA**	**L**	**L**	
2	**123**	116	**123**	**123**	**123**
3	131	117	131	**133**	**133**
4	139	123	139	**142**	**142**
5	146	132	146	**150**	**150**
6	151	138	151	**156**	**156**
7	155	144	155	**161**	**161**
8	158	148	158	164	**165**
9	160	148	161	**166**	165

Table 4: Geometric mean growth $\left(e^{\mathbb{E}[\ln(W(\mathbf{f}))]}\right)$ %

	$\bar{o}=1$			$\bar{o}=2$	$\bar{o}=3$
$\bar{s}$	**FK**	**SAA**	**L**	**L**	
2	227	97	227	**235**	**235**
3	248	150	248	**262**	261
4	212	166	214	**223**	216
5	239	107	245	260	**280**
6	312	75	323	342	**415**
7	293	167	324	**363**	358
8	338	81	362	**513**	308
9	336	118	404	**483**	394

Table 5: Observed terminal wealth, % of initial

We also evaluate the out-of-sample performance of our solutions. To do so, we fix our decisions for the six days, and simulate, 1 million times, the outcomes of all single options, and record the terminal wealth achieved. Table 6 summarizes the results, reporting the empirical terminal wealth (mean and standard deviation) as a percentage of initial investment over the set of scenario.

	$\bar{o}=1$			$\bar{o}=2$	$\bar{o}=3$
$\bar{s}$	**FK**	**SAA**	**L**	**L**	
2	227 ± 189	251 ± 231	227 ± 189	**252** ± 233	**252** ± 233
3	278 ± 279	288 ± 334	278 ± 279	358 ± 432	**364** ± 445
4	313 ± 356	325 ± 407	314 ± 356	475 ± 703	**482** ± 721
5	336 ± 418	371 ± 532	338 ± 418	617 ± 1088	**619** ± 1116
6	352 ± 475	426 ± 690	354 ± 477	646 ± 1197	**662** ± 1345
7	349 ± 483	406 ± 652	359 ± 504	641 ± 1206	**646** ± 1306
8	340 ± 474	391 ± 627	355 ± 508	621 ± 1245	**629** ± 1250
9	334 ± 468	369 ± 592	352 ± 506	**640** ± 1353	560 ± 1086

Table 6: Empirical terminal wealth (mean ± std), % of initial.

7.2.2 Discussion on $\bar{o} = 1$.

L consistently achieves the best results. **FK**, a strategy derived from approaches commonly used by casual bettors, also delivers strong performance. The geometric mean growth for the obtained solutions through **FK** are nearly identical to those obtained through **L**, indicating competitive long-run and risk-adjusted performance. In contrast, **SAA** is substantially worse than the other methods tested, both in terms of expected growth rate and the ending wealth. We have a thorough discussion of the limitation of **SAA** in this context in Section 7.5.

7.2.3 Discussion on $\bar{o} > 1$.

Introducing parlay options leads to substantial improvements. In terms of geometric mean growth, the solutions obtained by **L** further increases when extending to $\bar{o} = 2$ with a geometric mean growth reaching 164% when $\bar{s} = 8$ comparatively to 158% with $\bar{o} = 1$. A jump from 158% to 164% geometric mean growth corresponds to a $\frac{164}{158} - 1 \approx 3.8\%$ gain in long-run compounded growth. This results is also reflected in the real-world instances when $\bar{s} = 8$ with an observed terminal wealth of 513 × initial wealth when $\bar{o} = 2$ in comparison to a factor of 362× when $\bar{o} = 1$.

While **L** should improve monotonically with increasing $\bar{o}$ and $\bar{s}$, `BARON` fails to close the optimality gap for large instances within reasonable time, so this trend is not always observed. As discussed in Appendix D, increasing $\bar{s}$ and $\bar{o}$ greatly raises problem complexity—explaining why, for $\bar{s} = 9$, **L** with $\bar{o} = 2$ outperforms $\bar{o} = 3$.

Finally, Table 6 shows that **L** with $\bar{o} = 2$ attains a higher growth than **L** with $\bar{o} = 1$, but with significantly larger standard deviations of terminal wealth. This highlights a trade-off between growth and volatility: while parlay options enhance the geometric mean growth, they also increase the likelihood of extreme outcomes—both positive and negative. We extend this discussion in Section 7.2.5.

7.2.4 $\bar{f}$ Analysis Using Real-world Data

The previous analysis provided a comparison of all algorithms by evaluating their performance under a fixed maximum number of bets, $\bar{f} = 10$. However, as $\bar{s}$ increases, the number of available parlay combinations grows exponentially. Figure 3 compares the geometric mean growth achieved by **L** (with $\bar{o} = 2$) against **G**, as $\bar{f}$ increases. Recall that **G** selects the $\bar{s}$ single-option bets with the highest individual wealth measures w_s^* and requires wagering on all corresponding parlay combinations (i.e., $2^{\bar{s}} - 1$ bets) according to (3). Note, however, that the constraint given by $\bar{f}$ does *not* apply to method **G**, as shown by the dashed horizontal line. In contrast, the red triangle markers represent the values achieved by **L**. As expected, Figure 3 demonstrates

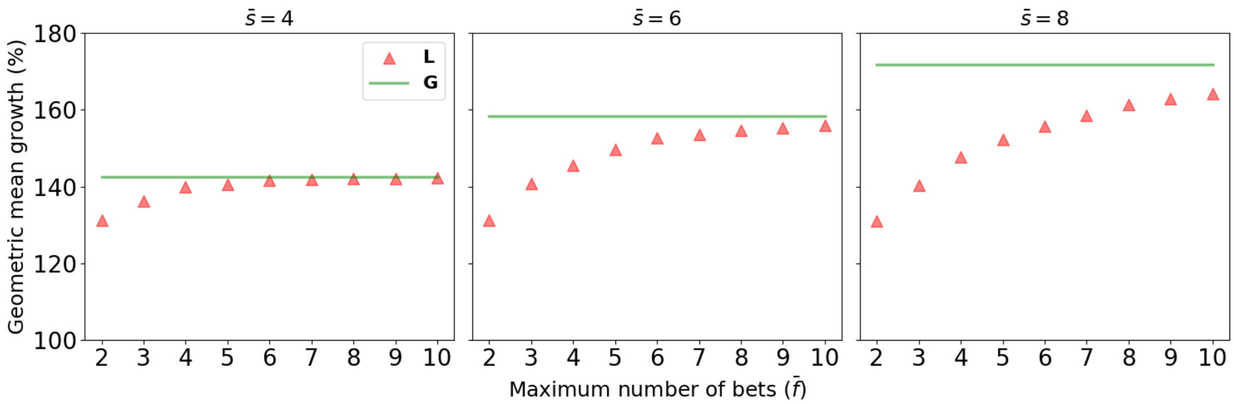


Figure 3: Geometric mean growth %: **L** vs. **G**.

that the gap in geometric mean growth of wealth between **G** and **L** decreases as $\bar{f}$ increases. Amazingly, the solutions obtained through **L** are reasonably close to **G** with substantially fewer bets. In particular, even for $\bar{s} = 8$, where betting on all $2^8 - 1 = 255$ achieves a maximum of approximately 175, we can obtain a solution with value over 130 even with limiting $\bar{f} = 2$. If we allow $\bar{f} = 10$ we achieve nearly equivalent objective value with an order of magnitude fewer allocations.

7.2.5 Drawdown and Tail Risk Analysis

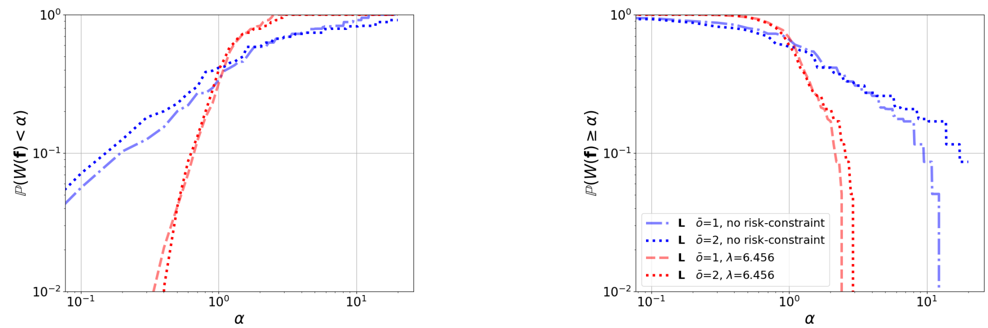
A significant concern with the log-utility objective is the potential for large losses. To mitigate this risk, Busseti et al. (2016) propose a drawdown constraint

$$\ln\left(\sum_{S' \subseteq \mathcal{S}} \exp\left[\ln\left(q\left(\mathcal{S}, S'\right)\right) - \lambda \ln\left(1 - \sum_{O \in \mathcal{O}^I(\mathcal{S})} f_O + \sum_{O \in \mathcal{O}^I(S')} f_O d_O\right)\right]\right) \leq 0,$$

parametrized by $\lambda > 0$, which controls the level of risk aversion. The purpose is to limit the likelihood that terminal wealth $W(\mathbf{f})$ falls drastically. As noted in their study, setting $\lambda = 6.456$ limits the probability of a 50% drawdown to approximately 1.1%. The authors in Busseti et al. (2016) show that this constraint is convex and thereby computationally attractive.

Given its convexity, the drawdown constraint can be naturally integrated into optimization problem (CLG($\mathcal{S}$, $\bar{s}$, $\bar{f}$, $\bar{o}$))—and by extension, into the LBBD framework—to explicitly manage downside risk, potentially at the cost of reduced expected wealth growth. Using the real-world instance spanning six days, we compare the terminal wealth of **L** with and without the drawdown constraint over 1,000,000 simulations. Figures 4a and 4b respectively illustrate the probabilities of drawdown (i.e., $\mathbb{P}(W(\mathbf{f}) < \alpha)$) and of maintaining or exceeding the wealth threshold α (i.e., $\mathbb{P}(W(\mathbf{f}) \geq \alpha)$). The x-axis and the y-axis in both plots are presented on a logarithmic scale. Results are shown for $\bar{s} = 6$, $\bar{o} \in \{1, 2\}$, and $\lambda = 6.456$ with $\alpha \in (0, 20)$.

Figure 4a shows that **L** with $\bar{o} = 2$ has a 23% probability of experiencing a drawdown exceeding 51%. In contrast, incorporating the risk constraint (with $\lambda = 6.456$) significantly mitigates downside risk, reducing the probability of a 51% drawdown to just 4%—a highly favorable outcome for risk-averse bettors. However, this reduction in risk comes at the cost of diminished upside potential, resulting in a less positively skewed wealth distribution. Indeed, Figure 4b illustrates that **L** with $\bar{o} = 2$ yields a 50% probability of achieving at least a 50% increase in wealth, compared to only 27% under the risk-constrained formulation. Under the risk-averse setting ($\lambda = 6.456$), the allocations for $\bar{o} = 1$ and $\bar{o} = 2$ are nearly identical; however, observe that $\bar{o} = 2$ is slightly less risky (Figure 4a) and also achieves a higher probability of exceeding α (Figure 4b).



(a) Drawdown probability (CDF) **(b)** Probability of exceeding α (Inverse CDF)

Figure 4: Impact of Risk Constraints on Drawdown and Upside Probabilities

Finally, Table 4 indicates that the standard deviation of outcomes increases with $\bar{o}$. Figure 4b further illustrates that **L** with $\bar{o} = 2$ yields a more positively skewed wealth distribution compared to **L** with $\bar{o} = 1$, a desirable attribute for risk-neutral bettors.

7.3 Comparative Analysis of B and L

The performance of **L** relies on `BARON`'s ability to solve **B**, which is comprehensively analyzed in Appendix D across various configurations of n_s, $\bar{s}$, $\bar{o}$, and $\bar{f}$. Section 7.2 presents a detailed examination of **L** on large-scale real-world instances. Note that `BARON` cannot solve such instances exactly, due to the large number of scenarios and number of parlay options.

In this section, we compare the performance of **L** and **B** in terms of solving time and optimality gap on small instances. We randomly generated 100 synthetic instances as described in Section 7.1, and recorded both the solving time and the optimality gap for **L** and **B**, given a 120-second time limit. Figure 5 provides a comparison of **B** (depicted as colored circles) and **L** (depicted as colored triangles) under the parameter configuration $n_s = 10$, $\bar{f} = 10$, $\bar{o} = 2$, and $\bar{s} \in \{4, 5, 6\}$. The figure shows the cumulative number of solved instances over time (in seconds) and the cumulative number of unsolved instances as a function of the optimality gap (in %).

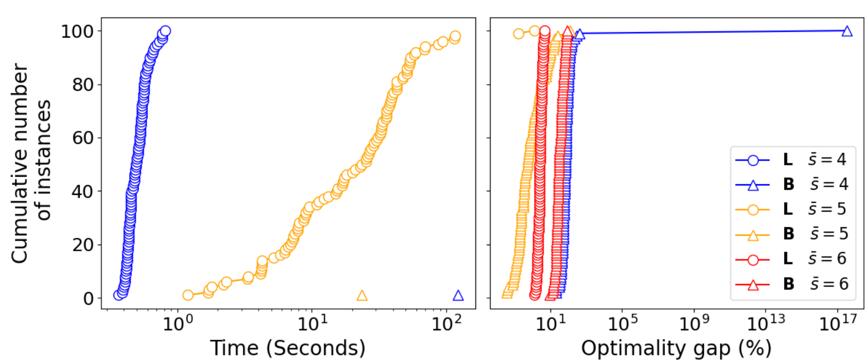


Figure 5: Comparative analysis between **L** and **B** with parameter setting $n_s = 10$, $\bar{o} = 2$, $\bar{f} = 10$, and varying $\bar{s}$

Figure 5 shows that **L** outperforms **B**. In particular, for $\bar{s} = 4$, nearly all 100 instances are solved to provable optimality by **L** but almost only one instance is solved to optimality by **B**. As $\bar{s}$ grows, the relative dominance shrinks. The reason for this is because **BARON** is used to solve the subproblems in **L**, thereby becoming the bottleneck and disallowing multiple rounds of Benders cuts.

7.4 Benders cut convergence

To better understand the performance gains that **L** provides, we analyze the number of cuts required to prove optimality. In these experiments we fix the maximum number of single-option bets to $\bar{s} = 5$ and the total number of bets to $\bar{f} = 10$, impose a 120-second time limit on each solve, and analyze the dataset collected on August 6, 2024. Figure 6 shows, for $\bar{o} \in \{1, 2, 3\}$, the progression of the lower bound (LB, blue triangles) and upper bound (UB, orange circles)—both reported as geometric mean growth (%)—over successive Logic-Based Benders iterations.

- $\bar{o} = 1$**:** A total of 29 cuts are generated before LB and UB meet. The UB begins at 105.02% while the LB starts near 104.81%. With each iteration the UB steadily falls, until both bounds coincide at iteration 29.
- $\bar{o} = 2$**:** By the second iteration LB and UB already coincide, illustrating how permitting two-leg parlays drastically shrinks the master problem's search space.
- $\bar{o} = 3$**:** Under the 120-second time limit, `BARON` fails to solve the initial instance, leaving the case $\bar{o} = 3$ unresolved. Appendix D discusses these limitations as $\bar{o}$ increases.

As $\bar{o}$ goes from single-option to two-leg parlay options, **L** requires dramatically fewer iterations. Allowing parlays of sizes two reduce the gap between the lower bound and the upper bound, yielding near optimal solutions quickly, and often even a tight optimality certificate in the first iteration.

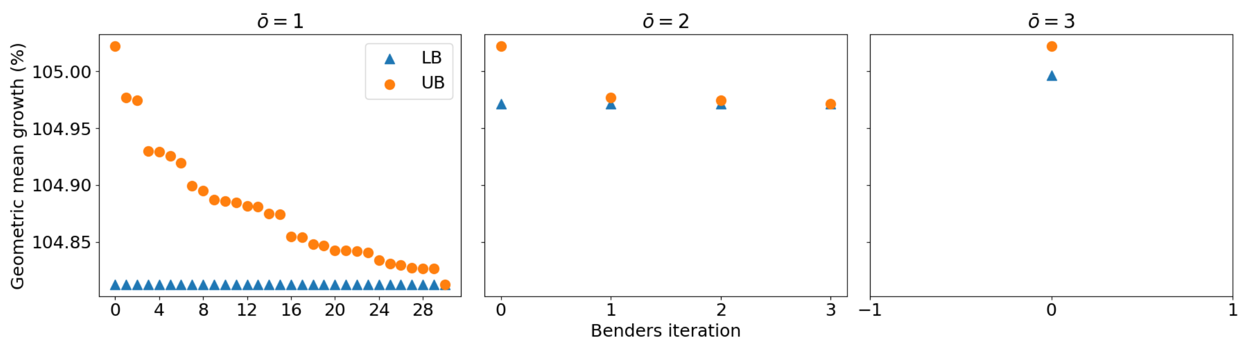


Figure 6: Benders Analysis for **L** with $\bar{s} = 5$, $\bar{f} = 10$, and Varying $\bar{o}$ (Data: 08-06-2024)

7.5 Challenges with SAA

For $\bar{o} = 1$, Tables 4–6 show that while **SAA** lags **FK** and **L** in geometric mean growth, it finds solutions with high expected terminal wealth—highlighting the trade-off between ROI and long-term growth. This is because **SAA** is highly sensitive to small sampling errors, leading to larger, riskier allocations.

We illustrate the challenge with applying **SAA**, which stems from sampling error, using a simple example with n_s identical profitable single-option bets (i.e., $p_s = p$ and $d_s = d$ for all $s \in [n_s]$). Suppose the bettor is restricted to selecting at most one bet (i.e., $\bar{s} = 1$ and $\bar{o} = 1$) and aims to maximize the expected logarithmic growth of wealth. Since all bets are identical, any single option s^* can be selected and the Kelly allocation f^* is given by $f^*_{s^*} = k_{s^*}$ and all other weights zero.

Contrastingly, for an implementation of **SAA**, the bettor generates n_c independent simulations and estimates the win probability for each option s as $\hat{p}_s := \frac{1}{n_c} \sum_{c=1}^{n_c} \xi_{c,s}$. The option $\hat{s}$ with the highest $\hat{p}_s$ would then be selected and the allocation $\hat{f}$ prescribed by **SAA** would be $\hat{f}_{\hat{s}} = \max\left\{\frac{\hat{p}_{\hat{s}} d - 1}{d-1}, 0\right\}$, with no weight on other bets.

The difference in expected terminal wealth achieved by $\hat{f}$ and f^* can be expressed as:

$$\begin{aligned}\mathbb{E}\left[W(\hat{f})\right] - \mathbb{E}\left[W(f^*)\right] &= \left[p(1 + \hat{f}_{\hat{s}}(d-1)) + (1-p)(1 - \hat{f}_{\hat{s}})\right] \\ &\quad - \left[p(1 + f^*_{s^*}(d-1)) + (1-p)(1 - f^*_{s^*})\right] \\ &= (dp - 1)(\hat{f}_{\hat{s}} - f^*_{s^*}).\end{aligned}$$

Therefore, if $\hat{p}_{\hat{s}} > p$ (which it generally would even if the number of possible options is small), then $\hat{f}_{\hat{s}} > f^*_{s^*}$ and $\mathbb{E}[W(\hat{f})] > \mathbb{E}[W(f^*)]$. However, this is too high for long-run wealth.

The tendency for **SAA** to over-allocate stems from the fact that the maximum of n_s empirical success rates is often much higher than the success rates of each single-option bets. Again for the simple case of identical bets, for each $s \in [n_s]$, let the number of successes be defined as $\tilde{Y}_s := \sum_{c=1}^{n_c} \tilde{\xi}_s^c$, which will follow a Binomial distribution, i.e., $\tilde{Y}_s \sim \text{Binomial}(n_c, p)$. Let $\tilde{p}_s = \frac{\tilde{Y}_s}{n_c}$. Then, the expected maximum empirical success rate is:

$$\mathbb{E}\left[\max_{s=1,\dots,n_s} \tilde{p}_s\right] = \frac{1}{n_c}\sum_{c=0}^{n_c} c \cdot \left[F(c)^{n_s} - F(c-1)^{n_s}\right],$$

where $F(c) = \mathbb{P}(X \leq c)$ is the CDF of the Binomial distribution with parameters n_c and p, and $F(-1)^{n_s} := 0$. Assume $n_s = 50$ single-option bets with $p = 0.5$ and $d = 2.5$, and $n_c = 1{,}000$ sampled simulations. The expected maximum empirical success rate is 0.536. Under this estimate, **SAA** selects the highest empirical probability bet, yielding a geometric mean growth of 101.8, compared to 102.1 for the optimal Kelly allocation. This example shows that even a small sampling error (3.6%) can lead **SAA** to a suboptimal choice—typically via over-allocation—even when selecting a single bet ($\bar{s} = 1$). This effect can be more pronounced in generalized settings.

While **SAA** yields asymptotically optimal allocations as n_c increases, its runtime grows sharply. An analysis of **SAA**'s scalability across various parameters appears in Appendix E.

8 Economic Analysis of Parlay Option's Pricing

In this section, we examine the pricing of parlay options and their impact on the expected logarithmic growth of wealth for a sports bettor. As shown in Section 7, in our data set, a bettor can boost expected long-run compounded growth by 3.1% by placing parlay bets of size two rather than restricting wagers to single-option bets. From the sportbook perspective, parlay option can be seen as risky asset when exploited by sophisticated bettors.

To analyze this effect in detail, we consider two profitable single-option bets, denoted by $\mathcal{S} = \{1, 2\}$. We assume that the decimal odds for the parlay option $O = \{1, 2\}$ satisfy $d_O \leq d_1\, d_2$, reflecting that a sportsbook has no incentive to offer overly favorable odds. Given that both single-option bets are profitable, the parlay option O remains profitable as long as $d_O > \frac{1}{p_1 p_2}$. For illustrative purposes, we set the event probabilities to $p_i = \frac{1+\rho_i}{d_i}$, with $\rho_i = 15\%$.

Our analysis considers the scenario where $d_1 = d_2 = d$ and $p_1 = p_2 = p$. Proposition 5.1 provides the optimal allocation when $\bar{o} = 1$. Each subplot in Figure 7 illustrates the expected logarithmic growth of wealth as the decimal odds for the parlay options vary. Figure 7 depicts the optimal allocation for the two single-option bets (pink square), the Grant allocation (assuming $d_O = d_1\, d_2$; green line), and the optimal allocation including the parlay for $\bar{o} = 2$ (blue scatter points).

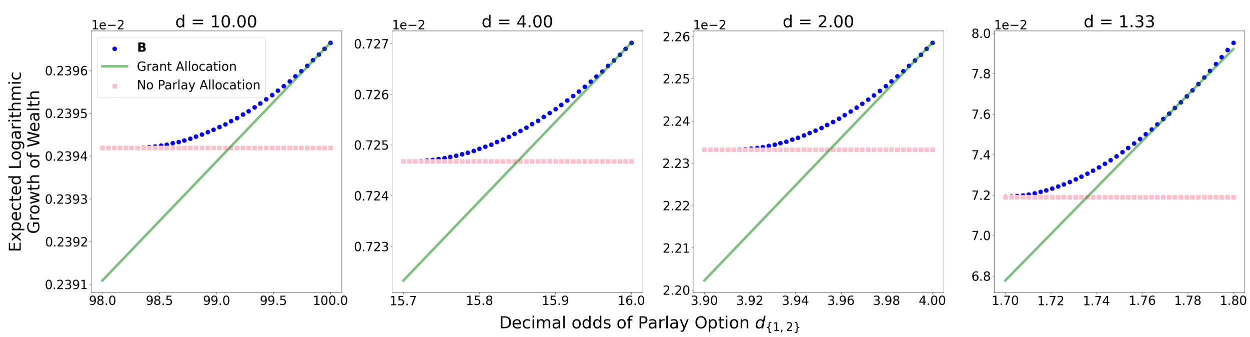


Figure 7: Expected logarithmic growth of wealth with respect to the decimal odd of the parlay option - $d_1 = d_2 = D$ and $\rho_1 = \rho_2 = 15\%$

Figure 7 demonstrates that as d decreases (with the corresponding increase in p), the incremental improvement in expected logarithmic growth of wealth from incorporating parlay options grows exponentially. Notably, when $d = 2$, a reduction in the parlay option's decimal odds from 4.00 to 3.92 eliminates any benefit from parlaying bets; the parlay offers no extra expected logarithmic growth over two single-option bets. Any further reduction yields suboptimal betting.

This offers valuable insights for sportsbooks. Specifically, if a sportsbook wants to limit its exposure to sophisticated bettors exploiting parlays, even a slight reduction in parlay pricing can make them unattractive in an optimal betting allocation. Casual bettors may still find value and continue placing these bets, so that the sportsbook also benefits from paying less if the parlay hits.

9 Conclusion

This work tackles the problem of optimally allocating wealth across binary options under real-world constraints, with a focus on sports betting. Specifically, we study the problem of maximizing logarithmic growth of wealth under combinatorial constraints and developed a scalable logic-based Benders decomposition algorithm to identify optimal allocation. Our economic analysis further reveals how sportsbooks influence a bettor's behavior through parlay pricing, showing that small payout adjustments can neutralize log-optimal strategies. While grounded in sports betting, our framework generalizes to binary prediction markets and other financial domains, with future directions including modeling option dependencies and incorporating estimation uncertainty.

References

Paul H. Algoet and Thomas M. Cover. Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. *The Annals of Probability*, 16(2):876 – 898, 1988. doi:10.1214/aop/1176991793. URL https://doi.org/10.1214/aop/1176991793.

American Gaming Association. Americans to legally wager estimated $3.1 billion on march madness. Press release, March 2025. URL https://www.americangaming.org/americans-to-legally-wager-estimated-3-1-billion-on-march-madness/. Accessed 2025-08-13.

Rose D. Baker and Ian G. McHale. Optimal betting under parameter uncertainty: Improving the kelly criterion. *Decision Analysis*, 10(3):189–199, 2013. doi:10.1287/deca.2013.0271. URL https://doi.org/10.1287/deca.2013.0271.

Scott R Baker, Justin Balthrop, Mark J Johnson, Jason D Kotter, and Kevin Pisciotta. Gambling away stability: Sports betting's impact on vulnerable households. Technical report, National Bureau of Economic Research, 2024. URL https://www.nber.org/papers/w33108.

Robert Bell and Thomas M. Cover. Game-theoretic optimal portfolios. *Management Science*, 34(6):724–733, 1988. doi:10.1287/mnsc.34.6.724. URL https://doi.org/10.1287/mnsc.34.6.724.

David Bergman and Jason Imbrogno. Surviving a national football league survivor pool. *Operations Research*, 65(5): 1343–1354, October 2017. doi:10.1287/opre.2017.1633. URL https://doi.org/10.1287/opre.2017.1633.

David Bergman, Carlos Cardonha, Jason Imbrogno, and Leonardo Lozano. Optimizing the expected maximum of two linear functions defined on a multivariate gaussian distribution. *INFORMS Journal on Computing*, 35(2): 304–317, 2023. doi:10.1287/ijoc.2022.1259. URL https://doi.org/10.1287/ijoc.2022.1259.

John R. Birge, Yifan Feng, N. Bora Keskin, and Adam Schultz. Dynamic learning and market making in spread betting markets with informed bettors. *Operations Research*, 69(6):1746–1766, 2021. doi:10.1287/opre.2021.2109. URL https://doi.org/10.1287/opre.2021.2109.

Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

L. Breiman. Optimal gambling systems for favorable games. In *Berkeley Symposium on Mathematical Statistics and Probability*, volume 4.1, pages 65–78. 1961.

Enzo Busseti, Ernest K. Ryu, and Stephen Boyd. Risk-constrained kelly gambling, 2016. URL https://arxiv.org/abs/1603.06183.

Jeff Decary, David Bergman, Carlos Cardonha, Jason Imbrogno, and Andrea Lodi. The madness of multiple entries in march madness, 2024. URL https://arxiv.org/abs/2407.13438.

Docsports.com. 50 sports betting statistics for 2025: Growth, trends, and market insights, June 2025. URL https://www.docsports.com/sports-betting-statistics.html. Accessed: 2025-08-13.

Robert B. Durand, Fernando M. Patterson, and Corey A. Shank. Behavioral biases in the nfl gambling market: Overreaction to news and the recency bias. *Journal of Behavioral and Experimental Finance*, 31:100522, 2021. ISSN 2214-6350. doi:https://doi.org/10.1016/j.jbef.2021.100522. URL https://www.sciencedirect.com/science/article/pii/S2214635021000666.

Andrew Grant, David Johnstone, and Oh Kang Kwon. Optimal betting strategies for simultaneous games. *Decision Analysis*, 5(1):10–18, 2008. doi:10.1287/deca.1080.0106. URL https://doi.org/10.1287/deca.1080.0106.

Andrew Grant, David Johnstone, and Oh Kang Kwon. A probability scoring rule for simultaneous events. *Decision Analysis*, 16(4):301–313, 2019. doi:10.1287/deca.2019.0393. URL https://doi.org/10.1287/deca.2019.0393.

Yongli Han, Philip Leung Ho Yu, and Thomas Mathew. Shrinkage estimation of Kelly portfolios. *Quantitative Finance*, 19(2):277–287, February 2019. URL https://doi.org/10.1080/14697688.2018.1483583.

Martin B. Haugh and Raghav Singal. How to play fantasy sports strategically (and win). *Management Science*, 67(1): 72–92, 2021. doi:10.1287/mnsc.2019.3528. URL https://doi.org/10.1287/mnsc.2019.3528.

Tomislav Horvat and Josip Job. The use of machine learning in sport outcome prediction: A review. *WIREs Data Mining and Knowledge Discovery*, 10(5):e1380, 2020. doi:https://doi.org/10.1002/widm.1380. URL https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1380.

David Scott Hunter, Juan Pablo Vielma, and Tauhid Zaman. Picking winners in daily fantasy sports using integer programming. *arXiv preprint arXiv:1604.01455*, 2016.

Jason Imbrogno and David Bergman. Computing the number of winning nfl survivor pool entries. *The College Mathematics Journal*, 53(4):282–291, 2022. doi:10.1080/07468342.2022.2099704. URL https://doi.org/10.1080/07468342.2022.2099704.

Edward H. Kaplan and Stanley J. Garstka. March madness and the office pool. *Management Science*, 47(3):369–382, 2001. doi:10.1287/mnsc.47.3.369.9769. URL https://doi.org/10.1287/mnsc.47.3.369.9769.

J. L. Kelly Jr. A new interpretation of information rate. *Bell System Technical Journal*, 35(4):917–926, 1956. doi:https://doi.org/10.1002/j.1538-7305.1956.tb03809.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1956.tb03809.x.

Steven Levitt. Why are gambling markets organized so differently from financial markets? *Economic Journal*, 114: 223–246, 02 2004. doi:10.1111/j.1468-0297.2004.00207.x.

Zach Lutz. Is america addicted to parlays? rapid rise in parlay bets. *Birches Health*, 2024. URL https://bircheshealth.com/resources/addicted-to-parlays. accessed: December 6, 2024.

L. C. MacLean, W. T. Ziemba, and G. Blazenko. Growth versus security in dynamic investment analysis. *Management Science*, 38(11):1562–1585, 1992. doi:10.1287/mnsc.38.11.1562. URL https://doi.org/10.1287/mnsc.38.11.1562.

Leonard C Maclean, Edward O Thorp, and William T. Ziemba. Long-term capital growth: the good and bad properties of the kelly and fractional kelly capital growth criteria. *Quantitative Finance*, 10(7):681–687, 2010. doi:10.1080/14697688.2010.506108. URL https://doi.org/10.1080/14697688.2010.506108.

Market Research Future. Daily fantasy sports market size, industry growth 2034. https://www.marketresearchfuture.com/reports/daily-fantasy-sports-market-22607, 2025. Accessed: 2025-08-13.

Peter Joseph Mercurio, Yuehua Wu, and Hong Xie. Option portfolio selection with generalized entropic portfolio optimization. *Entropy*, 22(8), 2020. ISSN 1099-4300. doi:10.3390/e22080805. URL https://www.mdpi.com/1099-4300/22/8/805.

Bruce L. Miller. Optimal portfolio decision making where the horizon is infinite. *Management Science*, 22(2):220–225, 1975. doi:10.1287/mnsc.22.2.220. URL https://doi.org/10.1287/mnsc.22.2.220.

Tobias J. Moskowitz. Asset pricing and sports betting. *The Journal of Finance*, 76(6):3153–3209, 2021. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/45396178.

Napat Rujeerapaiboon, Daniel Kuhn, and Wolfram Wiesemann. Robust growth-optimal portfolios. *Management Science*, 62(7):2090–2109, 2016. doi:10.1287/mnsc.2015.2228. URL https://doi.org/10.1287/mnsc.2015.2228.

Paul A. Samuelson. The "fallacy" of maximizing the geometric mean in long sequences of investing or gambling. *Proceedings of the National Academy of Sciences*, 68(10):2493–2496, 1971. doi:10.1073/pnas.68.10.2493. URL https://www.pnas.org/doi/abs/10.1073/pnas.68.10.2493.

Katherine Sayre and Isabella Simonetti. America has fallen in love with long-shot sports bets. *The Wall Street Journal*, 2025. URL https://www.wsj.com/business/what-are-parlays-sports-betting-gambling-881bdaee.

Jay Simon. Inefficient forecasts at the sportsbook: An analysis of real-time betting line movement. *Management Science*, 70(12):8583–8611, 2024. doi:10.1287/mnsc.2022.00456. URL https://doi.org/10.1287/mnsc.2022.00456.

Qingyun Sun and Stephen Boyd. Distributional robust kelly gambling. *arXiv preprint arXiv:1812.10371*, 2018.

Edward O Thorp. The kelly criterion in blackjack sports betting, and the stock market. In *Handbook of asset and liability management*, pages 385–428. 2008.

Chris Whitrow. Algorithms for optimal allocation of bets on many simultaneous events. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 56(5):607–623, 2007. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9876.2007.00594.x.

A Closed-form solution for symmetric two-option case with $d \neq 2$ and $p > \frac{1}{d}$

In this section, we provide a self-contained proof of Proposition 5.1.

Proof. Consider the symmetric two single-option bets with

$$p_1 = p_2 = p, \quad d_1 = d_2 = d,$$

and assume

$$\frac{1}{d} < p < 1, \quad dp \neq 2, \quad \pi := dp - 1 > 0,$$

where π denotes the expected net profit. We introduce the new decision variable A as the total fraction of the portfolio wagered on both single-option bets, and $u_i \in [0,1]$ as the proportion of A wagered on s_i. Given two single-option bets, we therefore have $f_1 = Au_1$ and $f_2 = Au_2$ with $u_1 = 1 - u_2$. As shown by Boyd and Vandenberghe (2004), an affine transformation of f_1 and f_2 preserves the convexity of the problem. Specifically, the objective function can be rewritten as:

$$\max_{u_1, A} p^2 \ln\Big(1 + Au_1(d-1) + A(1-u_1)(d-1)\Big) + p(1-p)\ln\Big(1 + Au_1(d-1) - A(1-u_1)\Big) + \qquad (LG_{2^*})$$
$$p(1-p)\ln\Big(1 - Au_1 + A(1-u_1)(d-1)\Big) + (1-p)^2 \ln\Big(1 - A\Big).$$

Because each log-term is strictly concave in (u_1, A) and we maximize over a convex set, (LG_{2^*}) is a strictly concave maximization. Hence any point satisfying the first-order (KKT) conditions is the unique global optimum.

Using the chain rule of derivatives, the first-order condition are

$$\begin{aligned}
\frac{dLG_{2^*}}{du_1} &= \frac{dLG_2}{df_1}\frac{df_1}{du_1} + \frac{dLG_2}{df_2}\frac{df_2}{du_1} \\
&= A(\frac{dLG_2}{df_1} - \frac{dLG_2}{df_2}) \\
\frac{dLG_{2^*}}{dA} &= \frac{dLG_2}{df_1}\frac{df_1}{dA} + \frac{dLG_2}{df_2}\frac{df_2}{dA} \\
&= u_1\frac{dLG_2}{df_1} + (1-u_1)\frac{dLG_2}{df_2},
\end{aligned}$$

which reduce to

$$\frac{dLG_{2^*}}{du_1} = p(1-p)\left(\frac{Ad}{1 + A(u_1 d - 1)} - \frac{Ad}{1 + A(d(1-u_1) - 1)}\right) \qquad (7)$$

$$\frac{dLG_{2^*}}{dA} = \frac{p^2(d-1)}{1 + A(d-1)} + p(1-p)\left(\frac{u_1 d - 1}{1 + A(u_1 d - 1)} + \frac{(1-u_1)d - 1}{1 + A((1-u_1)d - 1)}\right) - \frac{(1-p)^2}{1-A} \qquad (8)$$

Solving Equation 7 for u_1 gives $u_1 = \frac{1}{2}$. By replacing $u_1 = \frac{1}{2}$ in Equation 8, we can rewrite the derivative as

$$\frac{dLG_{2^*}}{dA} = \frac{A^2(-2(d-1)(d-2)) - A((dp)^2 - 2d(dp) + 2(dp + 3d - 4) + 2(dp-1)}{2(1 + 2A(d-1))(1 + A(d-2))(1-A)} \qquad (9)$$

After some algebraic manipulation, the first derivative of LG_{2^*} with respect to A can be written as

$$\frac{dLG_{2^*}}{dA} = -2(d-1)(d-2)\, f^2 - (\pi^2 - 2(d-2)\pi + d - 1)\, f + \pi.$$

Setting this derivative to zero yields the two critical points

$$A = \frac{-1}{2(d-2)(d-1)}(\pi^2 - 2(d-2)\pi + d - 1 \pm$$
$$\sqrt{(dp)^4 - 4d(dp)^3 + 4d^2(dp)^2 + 4(dp)^3 - 2d(dp)^2 - 4d^2(dp) - 4(dp)^2 + 4d(dp) + d^2}).$$

Let Δ denote the discriminant of this quadratic in f. A direct calculation shows

$$\begin{aligned}\Delta &= (dp)^4 - 4d(dp)^3 + 4d^2(dp)^2 + 4(dp)^3 - 2d(dp)^2 - 4d^2(dp) - 4(dp)^2 + 4d(dp) + d^2\\ &= (\pi^2 - 2(d-2)\pi + (d-1))^2 + 8(d-1)(d-2)\pi.\end{aligned}$$

To ensure feasibility, we next verify that $\Delta > 0$. By construction, we have $\pi > 0$ and $d \neq 2$. For $d > 2$, it is simple to see that $\Delta > 0$ since the first term is squared and the second term is strictly positive. For $1 < d < 2$, a deeper analysis is required. First, a direct computation shows that for all real π,

$$\frac{d^2\Delta}{d\pi^2} = 2\big[2\pi - 2(d-2)\big]^2 + 4\big[\pi^2 - 2(d-2)\pi + (d-1)\big] > 0,$$

so Δ is strictly convex for all π. Here $p \in \big[1/d,\, 1\big]$, so

$$\pi = dp - 1 \implies \pi \in \big[d\cdot(1/d) - 1,\ d\cdot 1 - 1\big] = \big[0,\ d-1\big].$$

A strictly convex function on the closed interval attains its (unique) minimum at one of the two endpoints. Hence the endpoints in π correspond to

$$\pi = 0 \implies \Delta = (d-1)^2 > 0, \qquad \pi = d-1 \implies \Delta = d^2(d-1)^2 > 0.$$

By convexity (under the hypothesis $1 < d < 2$), it follows that $\Delta(\pi) > 0$ for all $\pi \in [0, d-1]$.

Since by symmetry $u_1 = u_2 = \frac{1}{2}$ we have

$$f_1 = f_2 = f = \frac{A}{2},$$

so $A = 2f$. The two stationary candidates are

$$\hat{f}_\pm = \frac{\big(\pi^2 - 2(d-2)\pi + d - 1\big) \pm \sqrt{\Delta}}{-4\,(d-1)(d-2)}.$$

Since LG_{2^*} is strictly concave, there exists a unique global optimal solution. We always require $f \in [0, 1/2]$ so that $\ln(1-2f)$ remains well-defined. Therefore, we now analyze both $\hat{f}_+$ and $\hat{f}_-$ for all values of d and π. By construction, we have that $p \in [\frac{1}{d}, 1]$ and therefore $\pi \in [0, d-1]$. For $d > 2$ and $1 < d < 2$, $\hat{f}_\pm$ is continuous on both domains, and strictly increasing in π. Therefore, we next verify which of $\hat{f}_+$ and $\hat{f}_-$ is the global optimal solution for the end points of the domains.

Case 1: $d > 2$.

- Denominator $-4(d-1)(d-2) < 0$.
- At $p = 1/d$ $(\pi = 0)$:

$$\sqrt{\Delta} = d - 1 \implies \hat{f}_+ = \frac{(d-1)+(d-1)}{-4(d-1)(d-2)} < 0, \quad \hat{f}_- = \frac{(d-1)-(d-1)}{-4(d-1)(d-2)} = 0.$$

- At $p = 1$ $(\pi = d-1)$:

$$\sqrt{\Delta} = d(d-1) \implies \hat{f}_+ = \frac{(d-1)(d-1-2(d-2)+1+d)}{-4(d-1)(d-2)} < 0, \quad \hat{f}_- = \frac{1}{2}.$$

Hence as p runs from $1/d$ to 1, $\hat{f}_-$ increases from 0 to $\frac{1}{2}$ (staying inside $[0, \frac{1}{2}]$), while $\hat{f}_+$ lies strictly outside $[0, 1/2]$.

Case 2: $1 < d < 2$.

- Denominator $-4(d-1)(d-2) > 0$.
- At $p = 1/d$ $(\pi = 0)$:

$$\sqrt{\Delta} = d - 1 \implies \hat{f}_+ = \frac{(d-1)+(d-1)}{-4(d-1)(d-2)} < 0, \quad \hat{f}_- = \frac{(d-1)-(d-1)}{-4(d-1)(d-2)} = 0.$$

- At $p = 1$ $(\pi = d - 1)$:

$$\sqrt{\Delta} = d(d-1) \implies \hat{f}_+ = \frac{(d-1)(d-1-2(d-2)+1+d)}{-4(d-1)(d-2)} = \frac{1}{2-d} > \frac{1}{2}, \quad \hat{f}_- = \frac{1}{2}.$$

Again $\hat{f}_-$ runs from 0 up to 1/2, while $\hat{f}_+$ falls outside $[0, 1/2]$.

By strict concavity of the objective, the unique global maximizer valid for all $d > 1$, $d \neq 2$ is therefore always

$$f^* \;=\; \hat{f}_- = \frac{\left(\pi^2 - 2(d-2)\,\pi + (d-1)\right) - \sqrt{\Delta}}{-4\,(d-1)(d-2)},$$

where $\Delta = (\pi^2 - 2(d-2)\pi + (d-1))^2 + 8(d-1)(d-2)\pi$.

□

B Proof of Lemma 5.3: Simplified Optimal Logarithmic Growth Rate

In this section, we provide a self-contained proof of Lemma 5.3.

Proof. Grant et al. (2008) proves that $\mathbf{f}^* := (g_O)_{O \in \mathcal{O}(\mathcal{S})}$ in (3) is an optimal solution to LG. Furthermore, in Theorem 3 of Grant et al. (2008), the following statement (adapted to our notation) is presented: for all $S' \in \mathcal{S}$,

$$1 - \sum_{O \in \mathcal{O}(\mathcal{S})} g_O + \sum_{O \in \mathcal{O}(S')} g_O d_O = \prod_{s \in \mathcal{S}: s \in S'} \Big(1 - k_s + d_s k_s\Big) \prod_{s \in \mathcal{S}: s \notin S'} \Big(1 - k_s\Big). \tag{10}$$

Note that we purposely write the indices of the product terms as $s \in \mathcal{S} : s \in S'$ and $s \in \mathcal{S} : s \notin S'$ instead of $s \in S'$ and $s \notin S'$ for clarity later.

Using this identify, we can now write:

$$\begin{aligned} 1 - \sum_{O \in \mathcal{O}(\mathcal{S})} g_O + \sum_{O \in \mathcal{O}(S')} g_O d_O &= \prod_{s \in \mathcal{S}: s \in S'} \left(1 + (d_s - 1)\frac{p_s d_s - 1}{d_s - 1}\right) \prod_{s \in \mathcal{S}: s \notin S'} \Big(1 - k_s\Big) \\ &= \prod_{s \in \mathcal{S}: s \in S'} \Big(p_s d_s\Big) \prod_{s \in \mathcal{S}: s \notin S'} (1 - k_s) \end{aligned}$$

Therefore,

$$
\begin{aligned}
\mathbb{E}\left[\ln\left(W(\mathbf{f}^*)\right)\right] &= \sum_{S'\subseteq\mathcal{S}} q\left(\mathcal{S},S'\right)\ \ln\left(1-\sum_{O\in\mathcal{O}(\mathcal{S})} g_O + \sum_{O\in\mathcal{O}(S')} g_O d_O\right) &(11)\\
&= \sum_{S'\subseteq\mathcal{S}} q\left(\mathcal{S},S'\right)\ \ln\left(\prod_{s\in\mathcal{S}:s\in S'}\left(p_s d_s\right)\prod_{s\in\mathcal{S}:s\notin S'}\left(1-k_s\right)\right) &(12)\\
&= \sum_{S'\subseteq\mathcal{S}} q\left(\mathcal{S},S'\right)\left(\sum_{s\in\mathcal{S}:s\in S'}\ln(p_s d_s)+\sum_{s\in\mathcal{S}:s\notin S'}\ln\left(1-k_s\right)\right) &(13)\\
&= \sum_{S'\subseteq\mathcal{S}}\left(\sum_{s\in\mathcal{S}:s\in S'} q\left(\mathcal{S},S'\right)\ln(p_s d_s)+\sum_{s\in\mathcal{S}:s\notin S'} q\left(\mathcal{S},S'\right)\ln\left(1-k_s\right)\right) &(14)\\
&= \sum_{S'\subseteq\mathcal{S}}\sum_{s\in\mathcal{S}:s\in S'} q\left(\mathcal{S},S'\right)\ln(p_s d_s)+\sum_{S'\subseteq\mathcal{S}}\sum_{s'\in\mathcal{S}:s'\notin S'} q\left(\mathcal{S},S'\right)\ln\left(1-k_s\right) &(15)\\
&= \sum_{s\in\mathcal{S}}\sum_{S'\subseteq\mathcal{S}:s\in S'} q\left(\mathcal{S},S'\right)\ln(p_s d_s)+\sum_{s'\in\mathcal{S}}\sum_{S'\subseteq\mathcal{S}:s'\notin S'} q\left(\mathcal{S},S'\right)\ln\left(1-k_s\right) &(16)\\
&= \sum_{s\in\mathcal{S}}\left(\sum_{S'\subseteq\mathcal{S}:s\in S'} q\left(\mathcal{S},S'\right)\ln(p_s d_s)+\sum_{S'\subseteq\mathcal{S}:s\notin S'} q\left(\mathcal{S},S'\right)\ln(1-k_s)\right) &(17)\\
&= \sum_{s\in\mathcal{S}}\left(\ln(d_s p_s)\sum_{S'\subseteq\mathcal{S}:s\in S'} q\left(\mathcal{S},S'\right)+\ln(1-k_s)\sum_{S'\subseteq\mathcal{S}:s\notin S'} q\left(\mathcal{S},S'\right)\right), &(18)
\end{aligned}
$$

To sequence of arithmetic equivalences is as follows. Line (11) is by the optimality of $(g_O)_{O\in\mathcal{O}(\mathcal{S})}$. Line (12) is from the identity above. Line (13) uses the log identify. Line (14) distributes the probability of outcome within the inner summations. Line (15) distributes the outer summation. Line (16) follows because in the first summation there is one and only one term for every pair of s and S' containing s, and in the second summation there is one and only one term for every pair of s and S' not containing s. Line (17) factors the sum over all $s\in\mathcal{S}$. Finally, line (18) factors the log terms.

For any $s\in\mathcal{S}$, we can write the set of outcomes for which s is successful as $\{S'\subseteq\mathcal{S}: s\in S'\}$, so that

$$
p_s = \sum_{S'\subseteq\mathcal{S}:s\in S'} q\left(\mathcal{S},S'\right)
$$

and

$$
1-p_s = \sum_{S'\subseteq\mathcal{S}:s\notin S'} q\left(\mathcal{S},S'\right).
$$

Therefore,

$$
\begin{aligned}
\mathbb{E}\left[\ln\left(W(\mathbf{f}^*)\right)\right] &= \sum_{s\in\mathcal{S}} p_s\ln\left(d_s p_s\right)+\left(1-p_s\right)\ln\left(1-k_s\right)\\
&= \sum_{s\in\mathcal{S}} w_s^*.
\end{aligned}
$$

□

C Proof of Proposition 5.6:

Proof. By way of contradiction to (6), suppose there exists an optimal allocation $\mathbf{f}'' = (f''_O)_{O\in\mathcal{O}(\mathcal{S})}$ with $\sum_{O\in\mathcal{O}(\mathcal{S}):s'\in O} f''_O = 0$ and $\sum_{O\in\mathcal{O}(\mathcal{S}):s''\in O} f''_O > 0$. Note that this implies that for any O with $s',s''\in O$, $f''_O = 0$. Additionally, we define $\mathcal{O}''$ as the set of all parlay options O'' such that $f''_{O''} > 0$, where $s''\in O''$ but $s'\notin O''$. By construction, $f''_{O'} = 0$. Furthermore, let $\mathcal{O}'$ denote the set obtained by replacing s'' with s' in each option $O''\in\mathcal{O}''$, formally given by:

$$
\mathcal{O}' := \left\{O' = (O''\setminus\{s''\})\cup\{s'\} : O''\in\mathcal{O}''\right\}.
$$

We define a new solution $\mathbf{f}' = (f'_O)_{O\in\mathcal{O}(\mathcal{S})}$, which is identical to $\mathbf{f}''$ except that $f'_{O''} = 0$ and $f'_{O'} = f''_{O''}$. The total allocation does not change, i.e., $\sum_{O\in\mathcal{O}} f'_O = \sum_{O\in\mathcal{O}} f''_O$. Notice that for any pair of (O', O''), we have $d_{O''} < d_{O'}$ because $d_{s''} < d_{s'}$.

Next, we argue that $\mathbf{f}'$ constructed above is more preferred to $\mathbf{f}''$, i.e., $\mathbb{E}\left[\ln\left(W(\mathbf{f}'')\right)\right] - \mathbb{E}\left[\ln\left(W(\mathbf{f}')\right)\right] < 0$, contradicting the optimality of $\mathbf{f}''$ and thereby establishing (6). To that end, we divide the full sample space into four subspaces $\mathcal{S}_i \subseteq \mathcal{S}$, $i = 1, 2, 3, 4$ by

1. $\mathcal{S}_1 = \{S' \subseteq \mathcal{S} : s'' \notin S' \text{ and } s' \notin S'\}$ (s'' and s' are not a winning bet);
2. $\mathcal{S}_2 = \{S' \subseteq \mathcal{S} : s'' \in S' \text{ and } s' \notin S'\}$ (s'' is winning bet and s' is not);
3. $\mathcal{S}_3 = \{S' \subseteq \mathcal{S} : s' \notin S' \text{ and } s'' \notin S'\}$ (s' and s'' are not winning bet);
4. $\mathcal{S}_4 = \{S' \subseteq \mathcal{S} : s' \in S' \text{ and } s'' \notin S'\}$ (s' is winning bet and s'' is not);

Note that $\mathcal{S}_1 \cup \mathcal{S}_2$ and $\mathcal{S}_3 \cup \mathcal{S}_4$ respectively exclude s' and s''. The idea is that $\mathcal{S}_1 \cup \mathcal{S}_2$ (resp. $\mathcal{S}_3 \cup \mathcal{S}_4$) is constructed for $\mathbf{f}''$ (resp. $\mathbf{f}'$), where by construction $\mathbf{f}''$ (resp. $\mathbf{f}'$) allocates 0 to all parlay options $O' \in \mathcal{O}'$ (resp. $O'' \in \mathcal{O}''$). By construction, $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$, and $\mathcal{S}_4$ have the same cardinality and both $\mathcal{S}_1 \cup \mathcal{S}_2$ and $\mathcal{S}_3 \cup \mathcal{S}_4$ form exhaustive subspaces of $\mathcal{S} \setminus \{s'\}$ and $\mathcal{S} \setminus \{s''\}$, respectively. This allows us to rewrite

$$
\begin{aligned}
\mathbb{E}\left[\ln\left(W(\mathbf{f}'')\right)\right] - \mathbb{E}\left[\ln\left(W(\mathbf{f}')\right)\right] &= \sum_{S'\subseteq\mathcal{S}} q\left(\mathcal{S}, S'\right) \ln\left(W(\mathbf{f}'')(S')\right) - \sum_{S'\subseteq\mathcal{S}} q\left(\mathcal{S}, S'\right) \ln\left(W(\mathbf{f}')(S')\right) \\
&= \sum_{S_1\in\mathcal{S}_1} q\left(\mathcal{S}_1, S_1\right)\left(1 - p_{s''}\right) \ln\left(W(\mathbf{f}'')(S_1)\right) \\
&\quad + \sum_{S_2\in\mathcal{S}_2} q\left(\mathcal{S}_2 \setminus \{s''\}, S_2 \setminus \{s''\}\right) p_{s''} \ln\left(W(\mathbf{f}'')(S_2)\right) \\
&\quad - \sum_{S_3\in\mathcal{S}_3} q\left(\mathcal{S}_3, S_3\right)\left(1 - p_{s'}\right) \ln\left(W(\mathbf{f}')(S_3)\right) \\
&\quad - \sum_{S_4\in\mathcal{S}_4} q\left(\mathcal{S}_4 \setminus \{s'\}, S_4 \setminus \{s'\}\right) p_{s'} \ln\left(W(\mathbf{f}')(S_4)\right)
\end{aligned}
$$

Observe that $\forall S_1 \in \mathcal{S}_1$, there exists exactly one scenario $S_2 \in \mathcal{S}_2$, $S_3 \in \mathcal{S}_3$, and $S_4 \in \mathcal{S}_4$ such that $S_1 = S_2 \setminus \{s''\} = S_3 = S_4 \setminus \{s'\}$. Given that both $S_1 \cup S_2$ and $S_3 \cup S_4$ are exhaustive subspaces of $\mathcal{S}$ with identical cardinality, we next show that the logarithmic bankroll growth for $\mathbf{f}''$ over (S_1, S_2) is always less than that for $\mathbf{f}'$ over (S_3, S_4).

By construction, the two allocations, $\mathbf{f}'$ and $\mathbf{f}''$, are identical except for bets in $\mathcal{O}'$ and $\mathcal{O}''$. As such, the bettor's wealth, excluding the possible payout from $O' \in \mathcal{O}'$ in $\mathbf{f}'$ or from $O'' \in \mathcal{O}''$ in $\mathbf{f}''$, is always identical for every $\bar{S} \subseteq \mathcal{S} \setminus \{s', s''\}$, which we denote by $\widehat{W}$. For simplicity, we use vector notation to represent the bankroll growth associated with $\mathbf{f}''$ (resp. $\mathbf{f}'$) over all parlay options $\forall O'' \in \mathcal{O}''$ (resp. $\forall O' \in \mathcal{O}'$) as $\mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}$ (resp. $\mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}$).

Using the unique pair of scenarios (S_1, S_2, S_3, S_4) as defined above, the difference in logarithmic bankroll growth between $\mathbf{f}'$ and $\mathbf{f}''$ is given by:

$$
\begin{aligned}
& q\left(\mathcal{S}_1, S_1\right)\left(1 - p_{s''}\right) \ln(\widehat{W}) + q\left(\mathcal{S}_2 \setminus \{s''\}, S_2 \setminus \{s''\}\right) p_{s''} \ln\left(\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}\right) \\
&\qquad - q\left(\mathcal{S}_3, S_3\right)\left(1 - p_{s'}\right) \ln(\widehat{W}) - q\left(\mathcal{S}_4 \setminus \{s'\}, S_4 \setminus \{s'\}\right) p_{s'} \ln\left(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}\right) \\
&= q\left(\mathcal{S}_1, S_1\right) \Bigg(\left(1 - p_{s''}\right) \ln(\widehat{W}) + p_{s''} \ln\left(\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}\right) \\
&\qquad - \left(1 - p_{s'}\right) \ln(\widehat{W}) - p_{s'} \ln\left(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}\right) \Bigg) \qquad (19) \\
&\propto \left(1 - p_{s''}\right) \ln(\widehat{W}) + p_{s''} \ln\left(\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}\right) - \left(1 - p_{s'}\right) \ln(\widehat{W}) - p_{s'} \ln\left(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}\right) \qquad (20)
\end{aligned}
$$

The probability of observing any of S_1, S_2, S_3, or S_4, excluding s' and s'', is identical, i.e.,

$$q(\mathcal{S}_1, S_1) = q(\mathcal{S}_2 \setminus \{s''\}, S_2 \setminus \{s''\}) = q(\mathcal{S}_3, S_3) = q(\mathcal{S}_4 \setminus \{s'\}, S_4 \setminus \{s'\}).$$

This allows us to factor out $q(\mathcal{S}_1, S_1)$ in Line (19). Since $q(\mathcal{S}_1, S_1) > 0$ for all pairs of scenarios (S_1, S_2, S_3, S_4), we can omit this positive term and proceed by analyzing only the proportional function given in Line (20). Specifically, we will show that the expression in Line (20) is strictly negative for all such pairs, which directly implies that the overall difference in bankroll growth is negative. The proof is as follows:

$$(1 - p_{s''}) \ln(\widehat{W}) + p_{s''} \ln(\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}) - (1 - p_{s'}) \ln(\widehat{W}) - p_{s'} \ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}) \tag{21}$$

$$= p_{s''} \ln(\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}) - p_{s'} \ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}) + (1 - p_{s''}) \ln(\widehat{W}) - (1 - p_{s'}) \ln(\widehat{W}) \tag{22}$$

$$= p_{s''} \ln(\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}) - p_{s'} \ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}) - \ln(\widehat{W})(p_{s''} - p_{s'}) \tag{23}$$

$$< p_{s''} \ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}) - p_{s'} \ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}) - \ln(\widehat{W})(p_{s''} - p_{s'}) \tag{24}$$

$$= (p_{s''} - p_{s'}) \ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}) - \ln(\widehat{W})(p_{s''} - p_{s'}) \tag{25}$$

$$= (p_{s''} - p_{s'}) \ln\left(\frac{\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'}}{\widehat{W}}\right) < 0. \tag{26}$$

The sequence of arithmetic steps is as follows. Line (21) represents the difference between the growth rate of $\mathbf{f}''$ over (S_1, S_2) and that of $\mathbf{f}'$ over (S_3, S_4). Line (22) reorders the terms for simplicity. Line (23) factors out $\ln(\widehat{W})$ and reduces the probabilities as $(1 - p_{s''}) - (1 - p_{s'}) = -(p_{s''} - p_{s'})$. Line (24) substitutes $\ln(\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''})$ with $\ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'})$, which is strictly greater by construction. Line (25) factors out $\ln(\widehat{W} + \mathbf{f}'^T \mathbf{d}_{\mathcal{O}'})$. Line (26) factors out $p_{s''} - p_{s'}$. Finally, since $\ln\left(\frac{\widehat{W} + \mathbf{f}''^T \mathbf{d}_{\mathcal{O}''}}{\widehat{W}}\right) > 0$ and $p_{s''} - p_{s'} < 0$, the result of Line (26) is strictly less than 0. This implies that the logarithmic bankroll growth rate $\mathbf{f}''$ achieves over (S_1, S_2) is strictly less than the one $\mathbf{f}'$ achieves over (S_3, S_4).

The previous results hold for $\forall S_1 \in \mathcal{S}_1, S_2 \in \mathcal{S}_2, S_3 \in \mathcal{S}_3, S_4 \in \mathcal{S}_4 : S_1 = S_2 \setminus \{s''\} = S_3 = S_4 \setminus \{s'\}$. By summing over all pairs of scenarios, we have that $\mathbf{f}'$ generate a greater expected logarithmic growth rate than $\mathbf{f}''$, which is a contradiction with the initial assumption. Therefore, $\mathbf{f}'$ is strictly a better allocation than $\mathbf{f}''$, and it must be the case that

$$\sum_{O \in \mathcal{O}(\mathcal{S}): s'' \in O} f^*_O > 0 \implies \sum_{O \in \mathcal{O}(\mathcal{S}): s' \in O} f^*_O > 0$$

if s' dominates s''.

□

D Scalability Analysis of B

This section provides an extensive analysis of the scalability of **B** across various parameter settings and synthetic instances. We selected all possible combinations of the following parameters:

- $n_s = \bar{s} \in \{4, 6, 8\}$
- $\bar{f} \in \{2, 5, 10\}$
- $\bar{o} \in \{1, 2, 3\}$.

For each value of n_s, we generated 10 instances of single-option bets. Table 7 presents the average solving time (in seconds) and the average optimality gap given by `BARON` within a 120-second time limit across all instances. In the Solving Time and Optimality Gap columns, we include, in parentheses, the count of solved and unsolved instances out of the 10 generated. A "-" in these columns indicates that either all instances were unsolved (for Solving Time) or all instances were solved (for Optimality Gap).

Table 7 reveals that the efficacy of **B** in solving instances is significantly influenced by n_s and $\bar{o}$, while the impact of $\bar{f}$ is comparatively minimal. **B** can optimally solve instances of size up to $n_s = 6$ when $\bar{o} \leq 2$. This limitation suggests that **B** struggles to find optimal allocations for cases involving more than 6 single-option bets.

This paper is interested in the case where $n_s >> \bar{s}$. Next, we analyze the impact of $\bar{s}$ compared to n_s. Identically to Table 7, we selected all possible combinations of the following parameters:

		Solving time (s)			Optimality gap (%)		
n_s	$\bar{f}$ \ $\bar{o}$	1	2	3	1	2	3
4	2	0.18	0.27	0.40	–	–	–
	5	0.13	0.39	0.44	–	–	–
	10	0.13	0.12	0.40	–	–	–
6	2	0.40	4.67	20.06	–	–	–
	5	0.29	58.51	–	–	–	9.01 (8/10)
	10	0.17	36.71 (8/10)	–	–	2.37 (2/10)	2.40
8	2	3.47	82.34 (6/10)	–	–	–	15.51 (4/10)
	5	4.79	–	–	–	17.26	21.01
	10	0.58	–	–	–	7.20	6.82

Table 7: Performance of **B** under varying n_s, $\bar{o}$, and $\bar{f}$.

- $n_s \in \{8, 10, 12\}$
- $\bar{s} \in \{4, 6, 8\}$
- $\bar{f} = 10$
- $\bar{o} = 2$.

For each value of n_s, we generate 10 instances of single-option bets. Table 8 presents the average solving time (in seconds) and the average optimality gap achieved within a 120-second time limit across all instances.

	Solving time (s)			Optimality gap (%)		
n_s \ $\bar{s}$	4	6	8	4	6	8
8	35.42 (9/10)	–	–	13.61 (1/10)	6.74	7.19
10	–	–	–	15.64	12.04	13.49
12	–	–	–	96.68	100	100

Table 8: Performance of **B** under varying n_s and $\bar{s}$ (with $\bar{f} = 10$ and $\bar{o} = 2$). Dashes indicate instances solved to optimality within the time limit.

Table 7 shows that **B** efficiently solves all instances with $n_s = \bar{s} = 4$ and $\bar{o} = 2$, requiring less than 0.39 seconds on average. In contrast, Table 8 reveals that **B** takes significantly longer—an average of 35.42 seconds—to solve instances with $n_s = 8$, $\bar{s} = 4$, and $\bar{o} = 2$. Moreover, Table 7 demonstrates that **B** can handle most instances when $n_s = \bar{s} = 6$. However, as n_s increases, the number of scenarios and decision variables grows exponentially, which drastically impacts **B**'s ability to solve instances, even with $\bar{s} \leq 6$. This limitation highlights scenarios where **L** is likely to outperform **B**.

E Scalability Analysis of SAA

This section provides an extensive analysis of the scalability of **SAA** across various parameter settings and synthetic instances.

SAA delivers asymptotically optimal allocations as the number of scenarios n_c grows, but at the expense of sharply increasing run-time. Here, we quantify the trade-off between solution quality and computational effort when solving **SAA** with `BARON`. We generate 10 problem instances (see Section 7.1) for $n_s \in \{15, 50\}$. Each instance is then solved via **SAA** with $\bar{s} = 6, \bar{f} = 10,\ \bar{o} \in \{1, 2, 3\},$ and $n_c \in \{500, 1,000,\ 2,500,\ 5,000\}$. All optimal allocations are then post-evaluated exactly to obtain both the expected logarithmic growth of wealth and its standard deviation.

Table 9 reports, for each configuration $(n_s, n_c, \bar{o})$, the average value of $\mathbb{E}(\ln(W(\mathbf{f})))$ across the 10 instances; values in parentheses indicate in how many instances the log-return was strictly positive. A "—" denotes that BARON failed to find any profitable allocation.

n_s	n_c \ $\bar{o}$	1	2	3
15	500	0.061	0.063	0.057 (7/10)
	1,000	0.0693	0.068	0.031 (2/10)
	2,500	0.075	0.066 (9/10)	—
	5,000	0.070	—	—
50	500	0.071	—	—
	1,000	0.096	—	—
	2,500	0.104 (8/10)	—	—
	5,000	0.091(8/10)	—	—

Table 9: Performance metrics for **SAA** with n_s single-options bets across varying n_c and $\bar{o}$.

Two key patterns emerge from Table 9:

1. **Solver breakdown for large models:** When $\bar{o} \geq 2$, most instances fail to solve due to the combinatorial explosion in parlay options.
2. **Diminishing returns as n_c grows:** Even with $\bar{o} = 1$, the average log-wealth growth slightly declines as n_c increases—reflecting BARON's struggles on very large scenario sets (it fails on two instances at $n_c = 2,500$ and $n_s = 50$).

Since every real-world instance contains at least 49 single-option bets, the results in Table 9 justify our focus on $\bar{o} = 1$ and $n_c = 1{,}000$ for **SAA** in Section 7.2.