Tight Concentrations and Confidence Sequences From the Regret of Universal Portfolio

Francesco Orabona, Member, IEEE, and Kwang-Sung Jun

Abstract—A classic problem in statistics is the estimation of the expectation of random variables from samples. This gives rise to the tightly connected problems of deriving concentration inequalities and confidence sequences, i.e., confidence intervals that hold uniformly over time. Previous studies have shown that it is possible to convert the regret guarantee of an online learning algorithm into concentration inequalities, but these concentration results were not tight. In this paper, we show regret guarantees of universal portfolio algorithms applied to the online learning problem of betting give rise to new implicit time-uniform concentration inequalities for bounded random variables. The key feature of our concentration results is that they are centered around the maximum log wealth of the best fixed betting strategy in hindsight. We propose numerical methods to solve these implicit inequalities, which results in confidence sequences that enjoy the empirical Bernstein rate with the optimal asymptotic implicit inequalities, which results in confidence sequences that enjoy the empirical Bernstein rate with the optimal asymptotic implicit inequalities, which results in confidence sequences that enjoy the empirical Bernstein rate with the optimal asymptotic.

---

I. INTRODUCTION

Consider the problem of constructing valid confidence intervals for bounded random variables. We are interested in estimating the conditional mean \( \mu \) of a sequence of random variables \( X_1, X_2, \ldots \), such that for any \( i \) we have \( 0 \leq X_i \leq 1 \) almost surely and

\[
\mathbb{E}[X_i | X_1, \ldots, X_{i-1}] = \mu.
\]

A simple example of the condition above is obtained when \( X_i \)'s are independent random variables with their mean equal to \( \mu \). Of particular interest are discrete-time, time-uniform concentration inequalities that upper bound the probability that the empirical mean deviates from the expectation \( \mu \). A related concept is the one of producing \( (1 - \delta) \)-level confidence sequence [1], recently revived by [2], which is a sequence of confidence sets \( I_t = I_t(X_1, \ldots, X_t) \) such that

\[
P\{ \mu \in I_t, \forall t \geq 1 \} \geq 1 - \delta.
\]

We are interested in \( I_t \) being intervals, i.e., \( I_t = [\ell_t, u_t] \) for some \( \ell_t < u_t \). The time-uniform property makes these estimates more useful in the real world because they are “safer” to be used, allowing one to monitor the average of samples being collected and then stop gathering new samples at any time point in an adaptive way (e.g., when the required precision has been reached). This safety aspect was recently argued in much more general terms by [3].

In addition to conventional applications of testing in clinical trials and social sciences, confidence sequences have recently found numerous applications in statistics and machine learning. For example, A/B testing, or A/B/N testing more generally, is a form of randomized experimentation where two or more versions of a variable (e.g., layouts, contents, recommended products in web pages or mobile applications) are shown to different segments of website visitors at the same time to determine which version leaves the maximum impact and drives business metrics [4]. Time-uniform confidence sequences play an important role here as they allow one to monitor the data being collected (e.g., clicks or business metrics) to determine when to stop and conclude which alternative is more attractive. Given the limited amount of A/B testing that can be run in reality, it is crucial that the confidence sequence is not only order-wise tight but also numerically tight. Another application is the one of off-policy evaluation for recommendation systems [5], [6], [7], where it is critical to use a tight confidence sequence to be able to find a policy that has provably better performance than the currently-deployed policy with the desired confidence level [7].

For bounded distributions, the classic Bernstein inequality provides a concentration property of the sample mean that depends on the variance of the distribution. However, this is not useful for deriving a confidence sequence because it requires knowledge of the variance or an upper bound of it. Instead, empirical Bernstein inequalities [8], [9] allow us to leverage the empirical variance in place of the true variance, leading to a valid and computable confidence bound. However, it is commonly observed that empirical Bernstein bounds are not numerically tighter than Bernoulli-type confidence...
bounds [10, Eq. (3.4)] (commonly referred to as KL bounds) in the small-to-medium sample size regime, see, e.g., [11]. KL bounds, however, do not adapt to the variance, so they are asymptotically suboptimal compared to empirical Bernstein bounds. While one can compute both bounds and take the intersection of the bounds to achieve the best of both worlds, the only known way to do this is to use a union bound and split the target failure probability $\delta$ into two to be used for each bound, resulting in inflated widths. This is not desirable in the aforementioned applications where the tightness of the confidence width is critical.

In this work, we obtain tight time-uniform confidence sequences that have the empirical Bernstein rate, enjoy the optimal asymptotic behavior, and are also provably tighter than KL bounds at all sample sizes. The numerical tightness of our confidence sequences even allows us to have a nonvacuous confidence width (i.e., strictly less than 1) even with a single sample. To derive such a bound, we build on a recent line of work that uses online gambling algorithms to construct confidence sequences [12], [13], [14], [15]. In particular, we build on top of [14] that first showed a reduction of no-regret online betting algorithms to time-uniform confidence sequences. We extend their method showing how we can obtain non-vacuous confidence sequences from the regret of a class of portfolio selection algorithms. Our contributions are as follows:

- We derive a new implicit concentration inequality for bounded random variables satisfying (1).
- We show how to symbolically invert it, showing that our concentration inequality is, at the same time, a time-uniform empirical Bernstein-type inequality and a KL bound.
- We numerically invert our implicit concentration, obtaining empirical results that match or outperform prior algorithms. In particular, our algorithm is the first one to give a non-vacuous confidence sequence for any number of samples (Algorithm 1: PRECiSE-CO96).
- We also show how to obtain a fast numerical inversion with $O(1)$ computational complexity per sample while enjoying the same asymptotic performance of the exact inversion (Algorithm 2: PRECiSE-A-CO96).
- Finally, we show how to obtain state-of-the-art confidence sequences that satisfy the law of the iterated logarithm (Algorithm 3: PRECiSE-R70).

The rest of the paper is organized as follows: In Section II, we precisely define some concepts of online learning. Our main results are presented in Section III and IV, with their proofs deferred to Section V. In Section VI we present our numerical evaluation. Finally, in Section VII we discuss prior work and the history of the connection between betting and concentration inequalities and we conclude in Section VIII with a discussion.

II. SETTING AND DEFINITIONS

In this section, we explain the setting and definitions of online learning and betting, which are the basic tools we use to derive our results.

A. Online Learning

The online learning framework is a learning scenario that proceeds in rounds, without probabilistic assumptions. Specifically, in each round $t \in \{1, \ldots, T\}$, the learner outputs a prediction $x_t$ from a feasible set $V$, and then the adversary reveals a loss function $\ell_t : V \to \mathbb{R}$ from a fixed set of functions, and the algorithm pays $\ell_t(x_t)$. The aim of the learner is to minimize its regret, defined as the difference between its cumulative loss over $T$ rounds and a fixed arbitrary point $u \in V$:

$$\text{Regret}_T(u) := \sum_{t=1}^{T} \ell_t(x_t) - \sum_{t=1}^{T} \ell_t(u).$$

Note that we do not assume anything about how the adversary chooses the functions $\ell_t$ nor any knowledge of the future. We say that an algorithm is no-regret for a certain class of functions if its regret is $o(T)$ as $T$ goes to infinity for any $u \in V$. In the following, the keyword online will always denote an instantiation of the framework above. We refer the readers to [16] and [17] for complete introductions to this topic.

B. Online Betting on a Continuous Coin

Consider a gambler who makes repeated bets on the outcomes of adversarial coin flips. The gambler starts with initial money of $1. In each round $t$, he bets on the outcome of a coin flip $c_t \in \{-1, 1\}$ where $+1$ denotes heads and $-1$ denotes tails. Here, we do not make any assumption on how $c_t$ is generated. The gambler can bet any amount on either heads or tails. However, he is not allowed to borrow any additional money and thus cannot bet more than what he currently has. If he loses, he loses the betted amount; if he wins, he gets the betted amount back and, in addition to that, gets the same amount as a reward. We encode the gambler’s bet in round $t$ by a single number $\beta$ that is such that the sign of $\beta$ encodes whether we bet on heads or tails. $\beta$ The absolute value of $\beta$ is the amount of the current wealth to bet and its sign encodes whether we bet on heads or tails. The constraint that the gambler cannot borrow money implies that $\beta_t \in [-1, 1]$. We also slightly generalize the problem by allowing the outcome of the coin flip $c_t$ to be any real number in $[-a, b]$ where $a, b > 0$. Following [18], we call this continuous coin; the definition of wealth remains the same.

In fact, we can view the problem above as an online learning game where the algorithm predicts the signed fraction $\beta_t \in [-\frac{1}{b}, \frac{1}{a}]$ and the adversary reveals the convex loss function $\ell_t(\beta) = -\ln(1 + \beta c_t)$ where $c_t \in [-a, b]$. In this case, $\sum_{t=1}^{T} \ell_t(\beta_t)$ is the negative log wealth of the algorithm and
the regret is the logarithm of the ratio between the wealth of a constant betting fraction \( u \) and the wealth of the algorithm.

C. Online Portfolio Selection

We now describe the problem of sequential investment in a market with 2 stocks. The behavior of the market is specified by non-negative market gains vectors \( w_1, \ldots, w_T \in [0, +\infty)^2 \). The coordinates of a market gain vector \( w_t \) represent the ratio of the closing price to the opening price for the 2 stocks. An investment strategy is specified by a vector \( b_t \in B := \{[b_1, 1-b_1] \in \mathbb{R}^2 : 0 \leq b_1 \leq 1\} \), and its elements specify the fraction of the wealth invested on each stock at round \( t \). The wealth of the algorithm at the end of round \( t \) is given by 

\[
\text{Wealth}_{t}^{\text{portfolio}} := \prod_{i=1}^{t} w_i^T b_t
\]

We set \( \text{Wealth}_{0}^{\text{portfolio}} = 1 \).

As in the coin-betting problem, we can define the regret for this problem as the difference between the log wealth of the best constant rebalanced portfolio minus the log wealth of the algorithm. Denote by \( \text{Wealth}_{t}^{\text{portfolio}}(b) \) the wealth of the constant rebalanced portfolio with allocation \( b \), we have

\[
\text{Regret}_T = \max_{b \in B} \ln \text{Wealth}_{T}^{\text{portfolio}}(b) - \ln \text{Wealth}_{1}^{\text{portfolio}}.
\]

We will say that a portfolio algorithm is universal if the regret against any sequence of market gain vectors is sublinear in the number of rounds.

In the following, we focus on the \( F \)-weighted portfolio algorithms [19] that output at each step

\[
b_t = \frac{\int_B b \text{Wealth}_{t-1}^{\text{portfolio}}(b) \; dF(b)}{\int_B \text{Wealth}_{t-1}^{\text{portfolio}}(b) \; dF(b)},
\]

where \( F \) is a probability distribution over the two stocks.\(^1\)

In words, we predict a weighted-averaged betting where each weight for \( b \) is computed based on both \( \text{Wealth}_{t-1}^{\text{portfolio}}(b) \) and the prior \( F \). It is easy to see that for \( F \)-weighted portfolio algorithms the wealth of the algorithm can be written as

\[
\text{Wealth}_{t}^{\text{portfolio}} = \int_B \text{Wealth}_{t-1}^{\text{portfolio}}(b) \; dF(b).
\]

The coin-betting problem we described before can be reduced to the problem of online portfolio selection by setting the market gains as a function of the coin outcomes. The following lemma (proof in Section V-A) shows such a specialization for the case where the coin outcomes satisfy \( c \in [-m, 1-m] \) for some \( m \in [0, 1] \).

**Lemma 1.** For a continuous coin \( c \in [-m, 1-m] \), set two market gains as \( w_1 = 1 + \frac{c}{m} \) and \( w_2 = 1 - \frac{c}{m} \). Note that \( w_1, w_2 \geq 0 \) so they are valid market gains. Define \((b, 1-b)\) to be the play of a 2-stocks portfolio algorithm, where \( 0 \leq b \leq 1 \). Then, by taking

\[
\beta = -\frac{1}{1-m} + \left(\frac{1}{1-m} + \frac{1}{m}\right) b \in \left[-\frac{1}{1-m}, \frac{1}{m}\right]
\]

as the signed betting fraction, a continuous-coin-betting algorithm on \( c \) ensures that the gain in the coin betting problem coincides with the gain in the portfolio selection problem.

We can also write \( b \) in terms of \( \beta \) by \( b = (1-m)m\beta + m \). Note that \( \beta = \frac{1}{1-m} \) corresponds to \( b = 0 \), \( \beta = 0 \) corresponds to \( b = m \), and \( \beta = \frac{1}{m} \) corresponds to \( b = 1 \). Thanks to the reduction of Lemma 1, one can leverage existing portfolio algorithms to decide the betting, which can be re-written as

\[
\beta_t = \frac{\int_{1/(1-m)}^{1/m} \text{Wealth}_{t-1}^{\text{coin}}(\beta) \; dF(\beta)}{\int_{-1/(1-m)}^{-1/m} \text{Wealth}_{t-1}^{\text{coin}}(\beta) \; dF(\beta)},
\]

where \( \text{Wealth}_{t-1}^{\text{coin}}(\beta) \) is the wealth of a strategy with a constant betting \( \beta \) at the end of round \( t-1 \) and the prior \( F \) has been properly transformed for \( \beta \) in light of Lemma 1. Since the rest of the paper leverages the reduction above, we drop the superscript ‘portfolio’ and ‘coin’ from \( \text{Wealth}_{t}^{\text{coin}} \) hereafter.

III. A NEVER-VACUOUS TIME-UNIFORM CONCENTRATION INEQUALITY FROM A PORTFOLIO ALGORITHM

In this section, we present our new time-uniform concentration inequality and sketch the basic idea behind obtaining a concentration inequality from the regret of a betting algorithm. Our starting point is [14, Section 7.2], which showed that any regret guarantee for a valid betting algorithm can be used to construct a concentration inequality. However, they were interested in sub-exponential random variables in Banach spaces and had to rely on the expected wealth in the construction rather than the actual wealth, which does not lead to empirical variance-type concentration inequalities. For simplicity, we distill their main idea in the following theorem.

**Theorem 1.** Let \( \delta \in (0,1] \) and suppose we have an online coin-betting algorithm that in each round \( t \) produces a bet \( x_t \) on the continuous coin outcome \( Y_t \) based on the \( Y_1, \ldots, Y_{t-1} \) and guarantees the existence of \( W_1, W_2, \ldots \) such that

- \( W_0 = 1 \),
- \( W_t := W_t(Y_1, \ldots, Y_t) \geq 0, t = 1, 2, \ldots \) for any sequence \( Y_1, Y_2, \ldots \),
- \( W_{t-1} \geq E[W_t - Y_t x_t | Y_1, \ldots, Y_{t-1}], \) \( t = 1, 2, \ldots \)

Then, if

\[
\mathbb{E}[Y_t | Y_1, \ldots, Y_{t-1}] = 0, t = 1, 2, \ldots
\]

then

\[
\mathbb{P}\left\{ \max_t W_t \geq \frac{1}{\delta} \right\} \leq \delta.
\]

This theorem is immediate from i) Ville’s maximal inequality\(^2\) [20, page 84] and ii) the assumption that in each round \( t \) an online learning algorithm can only use the past to set the bet \( x_t \).

The basic idea of our construction is to design \( W_t \) to have very high values if the sequence of random variables \( Y_1, \ldots, Y_{t-1} \) does not satisfy (7). In particular, we can set \( W_t \) to be any tight non-negative lower bound of the wealth (or the exact wealth) of an online algorithm betting on the outcomes \( Y_t = X_t - \mu \), where \( X_1, \ldots, X_t \) satisfy (1). Then, Theorem 1 immediately tells us that, with high probability and uniformly over \( t \), \( W_t \) must be small. If the online betting algorithm is no-regret, from the definition of the optimal log wealth,

\(^1\)Note that Cover called this algorithm \( F \)-weighted universal portfolio.

\(^2\)\( \{ \max_t M_t \geq 1/\delta \} \leq \delta \) for any non-negative super-martingale starting from \( E[M_0] = 1 \) and any \( \delta \in (0,1] \).
Then, we have proof to Section V-B. More in details, we have a bound of 1 − δ ≤ \Pr(∀t ≥ 1, W_t ≤ 1) where W_t is a function of \mu. Since we do not know \mu, we can collect the set \mathcal{S} of all \mu’s that satisfy this inequality, which forms a confidence set. This set \mathcal{S} should contain the true \mu with probability at least 1 − δ by construction. Finally, computing \max S and \min S results in the upper and lower confidence bound for the true \mu. However, this procedure is computationally feasible only if the online betting algorithm is computationally efficient.

In online learning, it is often true that algorithms with (near) optimal performance guarantees are not computationally feasible, e.g., aggregating algorithms [21]. This means that we are in a difficult situation where one would either resort to heuristic methods without performance guarantees or methods that have weak guarantees. To get around this conundrum, we leverage a simple observation that will allow us to leverage computationally inefficient betting algorithms:

We do not need to know the plays of the betting algorithm; instead, we only need a lower bound of its wealth.

In particular, we can obtain a lower bound of the wealth of any online betting algorithm from the wealth of the best fixed strategy minus an upper bound of its regret.

Finally, while this reduction from confidence sequences to online betting would work for any betting strategy, it should be immediate to realize that an algorithm that guarantees a higher wealth will result in tighter confidence sequences. Hence, the final ingredient of our reduction is to choose an F-weighted portfolio algorithm (Eq. (4)) to be the online betting algorithm.

The reasoning above gives the following theorem, which is based on the transformation in Lemma 1 combined with the regret guarantee of an F-weighted portfolio algorithm, where F is a Dirichlet(1/2,1/2); we defer the details and precise proof to Section V-B.

**Theorem 2.** Let δ ∈ (0, 1). Assume X_1, X_2, . . . is a sequence of random variables such that for each i we have 0 ≤ X_i ≤ 1 and

\[ \mathbb{E}[X_i | X_1, \ldots, X_{i-1}] = \mu \]

almost surely. Let \[ G_t(\beta, \mu) := \sum_{i=1}^t \ln (1 + \beta(X_i - \mu)). \]

Then, we have

\[ \Pr \left\{ \max_t \max_{\beta \in [-\frac{1}{1+\delta}, \frac{1}{1+\delta}]} G_t(\beta, \mu) - R_t \geq \ln \frac{1}{\delta} \right\} \leq \delta, \]

where R_t := f(b_t, (b_t^* + 0.5), t) is a regret bound of an F-weighted portfolio algorithm with

\[ f(b, k, t) := \ln \frac{\pi \cdot b^k(1 - b)^{t-k} \Gamma(t+1)}{\Gamma(k+1/2) \Gamma(t-k+1/2)} \]

and

\[ b_t^* := \mu(1 - \mu) \left( \arg\max_{\beta \in [-\frac{1}{1+\delta}, \frac{1}{1+\delta}]} \sum_{i=1}^t \ln (1 + \beta(X_i - \mu)) \right) + \mu. \]

Moreover, R_t ≤ \ln \frac{\sqrt{\Gamma(t+1)} \Gamma(t+\frac{1}{2})}{\Gamma(\frac{1}{2})}.

One may recognize that W_t of Theorem 1 is the optimal log wealth \[ \max_{\beta \in [-\frac{1}{1+\delta}, \frac{1}{1+\delta}]} G_t(\beta, \mu) \] minus the regret R_t.

While we will provide a tight numerical computation of the confidence bound in Section III-D, let us first provide a quick-yet-loose construction of confidence bound for ease of exposition. One can simply use the loose bound of R_t ≤ \ln \frac{\sqrt{\Gamma(t+1)} \Gamma(t+\frac{1}{2})}{\Gamma(\frac{1}{2})} \leq \ln(\sqrt{\pi(t+1)}) and obtain, with probability at least 1 − δ,

\[ \forall t \geq 1, \max_{\beta \in [-\frac{1}{1+\delta}, \frac{1}{1+\delta}]} G(\beta, \mu) \leq \ln \frac{1}{\delta} + \ln(\sqrt{\pi(t+1)}). \] (9)

As explained above, one can obtain a confidence sequence for \mu by finding the max and min of the set of all \mu’s that satisfy the inequality above. Doing so numerically can be done efficiently, which we explain in Section III-D.

Next, let us focus on the maximum log wealth term in the left-hand side (9). As far as we know, this quantity is novel in the literature on concentration inequalities for bounded random variables and it is a main difference from our previous result in [14]. We show that this quantity allows recovering both KL and empirical Bernstein bounds. Later in Theorem 6, we also show that this quantity allows us to numerically calculate confidence sequences with a width strictly smaller than 1 for any δ ∈ (0, 1).

### A. Wealth Upper-Bounds the KL Divergence

Define \[ D(p, q) := p \ln \frac{p}{q} + (1 - p) \ln \frac{1-p}{1-q}. \] It may not seem obvious if the maximum log wealth is a better candidate for constructing a confidence sequence than the standard ones like Bernoulli KL-divergence based bound, see, e.g., [22, Theorem 10], which works for random variables supported in [0,1]:

\[ \Pr \left\{ \max_t t \cdot D(\hat{\mu}_t, \mu) - \ln f(t) \geq \ln \frac{1}{\delta} \right\} \leq \delta, \]

for some f(t) that grows polynomially in t or slower. In the following proposition, we show that the maximum log wealth is never worse than the KL divergence, which supports a viewpoint that the KL divergence is a special case of the maximum log wealth and that confidence bounds constructed with the maximum wealth are never worse than those with KL divergence, ignoring the minor difference in ln f(t).

**Proposition 1.** Let X_1, . . . , X_t ∈ [0,1], \[ \hat{\mu}_t = \frac{1}{t} \sum_{i=1}^t X_i, \]

and \mu ∈ [0,1]. Then,

\[ \max_{\beta \in [-\frac{1}{1+\delta}, \frac{1}{1+\delta}]} \ln \sum_{i=1}^t (1 + \beta(X_i - \mu)) \geq t \cdot D(\hat{\mu}_t, \mu), \]

where we achieve the equality if X_1, . . . , X_t ∈ {0,1} almost surely.

**Proof:** The proof is inspired by [18, Lemma 7]. Using Jensen’s inequality, we have for any X ∈ [0,1] that

\[ \ln(1 + \beta(X - \mu)) = \ln[X(1 + \beta(1 - \mu)) + (1 - X)(1 + \beta(0 - \mu))] \geq X \ln(1 + \beta(1 - \mu)) + (1 - X) \ln(1 + \beta(0 - \mu)). \]
Note that we achieve equality when $X = 1$ or $X = 0$. Then, we have
\[ \max_{\beta \in [-\frac{1}{\pi}, \frac{1}{\pi}]} \frac{1}{t} \sum_{i=1}^{t} \ln(1 + \beta(X_i - \mu)) \]
\[ \geq \max_{\beta \in [-\frac{1}{\pi}, \frac{1}{\pi}]} \sum_{i=1}^{t} X_i \ln(1 + \beta(1 - \mu)) + (1 - X_i) \ln(1 - \beta \mu) \]
\[ = \max_{\beta \in [-\frac{1}{\pi}, \frac{1}{\pi}]} t[\hat{\mu}_t \ln(1 + \beta(1 - \mu)) + (1 - \hat{\mu}_t) \ln(1 - \beta \mu)]. \]
As the RHS is concave in $\beta$, it remains to maximize the RHS over $\beta$. The solution is $\beta = \frac{\hat{\mu}_t - \mu}{\mu(1 - \mu)}$ with which the maximum becomes $t \cdot D(\hat{\mu}_t, \mu)$. \hfill $\square$

### B. Explicit Empirical Bernstein Time-Uniform Concentration

We now show that, besides implying the KL bound above, the implicit concentration in Theorem 2 also implies an empirical Bernstein time-uniform concentration. Note that a sequel paper [23] shows that the key supermartingale used for the state-of-the-art empirical Bernstein bound [24, Lemma 13] is further upper-bounded by the log wealth [23, Corollary 5], which means that the confidence sequence based on the log wealth is never worse than that of the empirical Bernstein up to logarithmic terms in $n$. This highlights the fundamental role of the betting-based martingale in establishing tight confidence sequences.

**Theorem 3.** Under the assumptions of Theorem 2, denote by $\hat{\mu}_i = \frac{1}{t} \sum_{j=1}^{i} X_j$, $V_i = \sum_{j=1}^{i} (X_j - \hat{\mu}_i)^2$, $R_i = \ln \frac{\sqrt{\hat{\mu}_i t + 1}}{\sqrt{\hat{\mu}_i t} + 1}$, and
\[ \epsilon_i = \frac{(4/3) t R_i + \sqrt{16/9} t^2 R_i^2 + 8 \sqrt{V_i R_i} (t^2 - 2 t R_i)}{2 t^2 - 4 t R_i} . \]
Then, with probability at least $1 - \delta$ uniformly for all $t$ such that $t > 2 R_t$, we have
\[ \max_{i=1, \ldots, t} \hat{\mu}_i - \epsilon_i \leq \mu \leq \min_{i=1, \ldots, t} \hat{\mu}_i + \epsilon_i . \]

We defer the proof to Section V-F. For $t$ big enough the deviation is roughly \[ \frac{(4/3) \ln \left( \sqrt{\delta / \beta} \right)}{t} + \sqrt{2 V_i R_i} \ln \left( \sqrt{\delta / \beta} \right) \], similarly to the inequalities in [9] and [25]. The second term here can be further bounded by \[ \sqrt{2 \hat{\mu}_i V_i / t} \ln \left( \sqrt{\delta / \beta} \right) + \sqrt{2 \hat{\mu}_i V_i} \ln(1 / \beta) . \]
One can easily see that, in terms of the scaling with $\ln(1 / \beta)$, the factor $\sqrt{2 \hat{\mu}_i / t}$ is the optimal one due to the central limit theorem. For the scaling with $t$, we show later in Section IV that it is possible to obtain \[ \sqrt{2 \hat{\mu}_i V_i / t} \ln(1 / \beta) + o(1) \], which matches the law of the iterated logarithm (thus asymptotically optimal) with a slightly different strategy.

### C. F-Weighted Portfolios: Optimal Wealth and Permutation-Invariance

Finally, let us justify the choice of the F-weighted portfolio algorithms in Theorem 2. First of all, F-Weighted portfolios are the only algorithm in the literature that allows us to have the maximum log wealth in Theorem 2; other non-universal betting algorithms will necessarily have a lower bound of that quantity, e.g., Online Newton Step in [26, Section 3]. However, one might wonder if there are better choices. This is readily answered: [27] showed that F-weighted portfolios achieve an anytime (i.e., time-uniform) regret upper bound that is at most $\ln \sqrt{2 \pi}$ larger than the optimal achievable regret for this problem. Also, later in Theorem 8, we show that it is possible to derive the law of the iterated logarithm using an F-weighted portfolio algorithm. Hence, F-weighted portfolios are optimal in two different ways.

Moreover, the wealth of F-weighted portfolios is invariant to the order of the market gains [19, Proposition 4], as it is clear from (5). So, the confidence interval at time $t$ obtained from the wealth of an F-weighted portfolio algorithm is independent of the order of the samples $X_1, \ldots, X_t$. We believe this is an important property to avoid a “lottery” based on the ordering of the samples, i.e., obtaining a worse or better confidence interval at a given time step just by reshuffling the samples. This is commonly considered an undesired effect, as motivated for example in [28].

More importantly, a lesser-known result in [20, pages 85–87] proves that in the case of betting on a coin the only achievable wealth processes that are invariant to the order of the outcomes are the ones from F-weighted portfolios. Hence, if one desires confidence sequences that are invariant to the order of the samples, F-weighted portfolio algorithms are the only possible choice.

In the next section, we show how to numerically invert the inequality in Theorem 2, which leads to a tighter confidence bound than (9).

### D. Tight Numerical Inversion of Theorem 2

Having explained how to quickly derive confidence sequences from the implicit concentration in Theorem 2, we now show the full power of our result showing how to numerically invert it.

For the loose version shown in (9), recall that one needs to compute max and min of the set $\mathcal{S}$ of all $\mu$’s that satisfy (9). Since $R_t := \ln(\sqrt{\pi (t + 1)})$ is independent of $\mu$, as long as $\mathcal{S}$ is an interval it suffices to perform a binary search, which takes $O(\ln(1 / \text{precision}))$ iterations for the target numerical precision. Indeed, we prove that $\mathcal{S}$ is an interval since $\max_{\beta \in \left[ -\frac{1}{\pi\sqrt{2}}, \frac{1}{\pi\sqrt{2}} \right]} G(\beta, \mu)$ is a quasi-convex function of $\mu$ as follows:

**Theorem 4.** Let $X_i \in [0, 1]$ for $i = 1, \ldots, t$. Let $\hat{\mu}_t := \frac{1}{t} \sum_{i=1}^{t} X_i$. Take $G(\beta, m)$ from Theorem 2 and define $H(m) := \max_{\beta \in \left[ -\frac{1}{\pi\sqrt{2}}, \frac{1}{\pi\sqrt{2}} \right]} G(\beta, m)$. Then, $H(m)$ is nonincreasing in $[0, \hat{\mu}_t)$ and nondecreasing in $(\hat{\mu}_t, 1]$. Hence, $H(m)$ is quasicontinuous in $[0, 1]$.

Since the maximization over $\beta$ is a one-dimensional concave problem that can be easily solved with any standard convex optimization algorithm, the numerical computation of the confidence bound can be performed efficiently.

### E. Tighter Confidence Bounds

The simple method above is based on a loose bound $R_t$ on $R_t$. Instead, we can obtain a much tighter bound on $R_t$, which

---

We were unable to find any mention of this result in prior work.
Algorithm 1 PRECISE-CO96: Portfolio REGret for Confidence SEquences Using [27]

1: **Input**: $\delta \in (0, 1)$, random variables $X_1, X_2, \ldots$ in $[0, 1]$
2: $t_0 = 0, u_0 = 1$
3: for $t = 1, 2, \ldots$ do
4: $\mu_t = \frac{1}{t+1} \sum_{i=1}^{t} X_i$
5: $b = \arg\max_{b \in [0,1]} \sum_{i=1}^{t} \ln \left( \frac{b X_m}{1-b} + \frac{1-b}{1-b} \right)$
6: $R_t = \max \{ f(b^* \cdot t - 0.5) / t, b^* \cdot t - 0.5, t \}$, where $f$ is in (8)
7: $L_t = \max \{ \ell_t \cup \{ u_{t-1} \} \}$
8: $\ell_t = \max \{ (L_t \cup \{ \ell_{t-1} \}) \}$ by binary search
9: $\mu_t = \arg\max_{\mu} \left\{ \frac{1}{t} \sum_{i=1}^{t} \ln \left( \frac{b X_m}{1-b} + \frac{1-b}{1-b} \right) \right\}$
10: $U_t = \max \{ \mu_t \}$
11: $u_t = \min \{ U_t \}$
12: end for

would lead to a tighter confidence bound. Let us define $R_t(m)$ as $R_t$ defined in Theorem 2 with $\mu$ replaced by $m$ for clarity. As explained above, to invert the concentration inequality in Theorem 2, it is enough to find the set $S$ of $m$’s that satisfies the inequality

$$\max_{b \in [0,1]} \sum_{i=1}^{t} \ln \left( 1 + \beta(X_i - m) \right) - R_t(m) \leq \frac{1}{\delta}.$$  (10)

Unlike the loose version (9), the regret bound $R_t(m)$ now depends on $m$, so the quasi-convexity of $H(m)$ is not sufficient; depending on the expression of $R_t(m)$ the set $S$ might not be an interval. Hence, we construct a tight upper bound on $R_t(m)$ that makes the left-hand side of (10) quasi-convex in $m$, which allows us to use binary search. The overall procedure is described in Algorithm 1 that we call PRECISE-CO96 (Portfolio REGret for Confidence SEquences (PRECISE) using [27]). This algorithm is guaranteed to generate a valid confidence sequence, as detailed in the following theorem.

**Theorem 5.** Let $\delta \in (0, 1)$. Assume $X_1, X_2, \ldots$ a sequence of random variables such that for each $i$ we have $0 \leq X_i \leq 1$ and

$$\mathbb{E}[X_i | X_1, \ldots, X_{i-1}] = \mu$$

almost surely. Run Algorithm 1 on $X_1, X_2, \ldots$. Then, we have $\mathbb{P}(\forall t \geq 1, \mu \in [\ell_t, u_t]) \geq 1 - \delta$.

The algorithm works by finding a lower bound $\ell_t$ and an upper bound $u_t$ for the mean $\mu$ at each time step $t$. In particular, it has to find the zero of a quasi-convex function in lines 8 and 12, which can be done by binary search. However, the function itself is defined as the maximum of a concave function on a bounded domain, which again can be solved by a binary search or any convex optimization algorithm.

F. Never-Vacuous Guarantees

We prove our claim that the confidence sequences from Algorithm 1 are never vacuous; the proof is in Section V-E.

**Theorem 6.** Under the assumptions of Theorem 5, for any $t$ we have that $u_t - \ell_t \leq u_t - \ell_t = 1 - \frac{1}{2}\delta$.

While it is common to reason about the asymptotics of confidence intervals, it is less known what the right behavior is for a small number of samples. Hence, to better understand the statement in the last theorem, it is useful to compare it with the width of the exact confidence intervals for a Binomial distribution [29]. By definition, the probability that the exact confidence interval includes $\mu$ is exactly $1 - \delta$. Of course, it would be impossible to match this bound with time-uniform results. However, we essentially match the right dependency on $\delta$, allocating $\frac{1}{2}$ one half of the probability of error on $t = 1$ and the other half on all the other $t$; see Figure 6 for comparing ours with the exact confidence bound.

G. Leveraging the Stochasticity the Data Cannot Improve the Results

As explained above, higher wealth corresponds to tighter confidence sequences. Interestingly, F-weighted portfolio algorithms that we used in the proof are designed to work well when there are no assumptions on the data-generating mechanism; i.e., they work with arbitrarily or even adversarially generated data. One might wonder if they make overly-pessimistic bets to guard against adversarial data. Could we obtain significantly better confidence bounds by leveraging the fact that the data is stochastic?

We claim that this is not likely to be true for the following reasons. Our results above imply that there is not much to improve. Specifically, Theorem 3 shows that the leading term is almost optimal — in fact, we later show a version that achieves the optimal confidence width asymptotically (Section IV). Even for the non-leading term, Theorem 6 shows that our bound is nonvacuous even with one sample, and we empirically show later in Section VI that our algorithms have superior performance in the small sample regime compared to state-of-the-art bounds. We remark that the competitors in our experiments include heuristic betting algorithms [15] that are exclusively designed to work well for stochastic data without worrying about having tight regret bounds.

Furthermore, existing results show that F-weighted portfolio algorithms almost achieve the largest possible wealth for stochastic data, not just adversarial data. Specifically, note that the wealth achieved by the algorithm is equal to the log wealth of the best constant rebalanced portfolio (CRP) minus the regret. First, we note that [30, Chapter 16] have shown that if the market gains are i.i.d., the optimal growth rate of the wealth is achieved by a CRP. Thus, as long as the regret is small compared to the log wealth of the best CRP, the algorithm must be nearly optimal in maximizing the wealth, thus leading to tight confidence bounds. Second, indeed, the regret is small. Specifically, the upper bound of the regret we use in Theorem 2 is at most logarithmic in time, while the log wealth in lines 7 and 11 of Algorithm 1 roughly grows linearly.
in time, as we can infer from Proposition 1. Hence, the use of the regret upper bound instead of the exact regret can have at best a minor effect. Moreover, for Bernoulli random variables our regret upper bound is tight. This is due to the fact that [27] have shown that the regret upper bound of the F-weighted portfolios is tight for Kelly markets, which corresponds to \( X_t \) being Bernoulli random variables in our reduction.

**H. Computationally Efficient Version**

One downside of Algorithm 1 is that its per-time-step time complexity is \( \Theta(t) \), which means that the cumulative time complexity up to time step \( t \) is \( \Theta(t^2) \). In contrast, empirical Bernstein bounds that take the form of Theorem 3 can be implemented with \( O(1) \) per-time-step complexity by making incremental updates to the first and the second moments. However, they are usually not as tight as PRECiSE-CO96 in the small sample regime. Could we obtain a confidence bound that is both tight and computationally efficient?

We answer this question affirmatively by proposing a computationally efficient version called PRECiSE-A-CO96 (PRECiSE with Approximation using [27]) described in Algorithm 2. The main idea starts from the fact that the first and second moments can be computed incrementally. Thus, if we could compute a tight lower bound on the maximum log wealth \( \max_{\mu \in [-1/(1-m), 1/m]} \sum_{i=1}^t \ln(1 + \beta(X_i - m)) \) as a function of those moments in a closed form, then we will not have to pay the linear per-time-step time complexity in \( t \). We leverage the tight inequality by [31] to obtain a closed-form lower bound on the maximum log wealth: \( G_t^\beta(\beta, m) \) (line 6 of Algorithm 2) for computing the lower confidence bound \( \ell_t \) and \( G_t^\beta(\beta, m) \) (line 10) for computing upper confidence bound \( u_t \). These lower approximations are maximized at the optimal betting of \( \beta_t(m) \) and \( \beta_t(m) \) respectively for which we have a closed form as well. This alone, however, can be loose for small \( t \). Fortunately, Proposition 1 says that the KL divergence is a valid lower bound of the maximum log wealth. We can therefore take the maximum of these two approximations as a tighter lower bound of the maximum log wealth. Note that the algorithm uses binary search, which results in \( O(\ln(1/\text{precision})) \) per-time-step time complexity if one updates \( \mu_t \) and \( V_t \) incrementally.

The following theorem shows that Algorithm 2 computes a valid confidence sequence. The proof is in Section V-G.

**Theorem 7.** Let \( \delta \in (0, 1) \). Assume \( X_1, X_2, \ldots \) a sequence of random variables such that for each \( i \) we have \( 0 \leq X_i \leq 1 \) and \( \mathbb{E}[X_i | X_1, \ldots, X_{i-1}] = \mu \) almost surely. With the notation in Algorithm 2, let \( G_t^\beta(m) = \mathbb{E} \left[ m \leq \mu_t \right] G_t^\beta(\beta_t(m), m) + \mathbb{E} \left[ m > \mu_t \right] G_t^\beta(\beta_t(m), m) \). Then, with probability at least \( 1 - \delta \), we have

\[
\mathbb{P} \left( \max_t \left( \max_m \{ G_t(\mu), D(\mu, \mu) \} - \bar{R}_t \geq \ln(1/\delta) \right) \right) \leq \delta.
\]

Furthermore, for every \( t \geq 1 \), the confidence set for time \( t \)

\[
\left\{ m \in \mathbb{R} : \max_t \{ G_t(m), D(\mu, m) \} - \bar{R}_t \geq \ln(1/\delta) \right\}
\]

is an interval.

---

**Algorithm 2 PRECiSE-A-CO96: Portfolio Regret for Confidence SEquences With Approximation Using [27]**

```plaintext
1: Input: \( \delta \in (0, 1) \), random variables \( X_1, X_2, \ldots \) in \([0, 1]\)
2: \( u_0 = 0 \), \( u_0 = 1 \)
3: for \( t = 1, 2, \ldots \) do
4: \( R_t = \ln \frac{\Theta(t+1)}{\Theta(t+2)} \)
5: \( \hat{\mu}_t = \frac{1}{t} \sum_{i=1}^t X_i \), and \( V_t = \frac{1}{t} \sum_{i=1}^t (X_i - \hat{\mu}_t)^2 \) (use online updates)
6: Let \( G_t^\beta(\beta, m) = \beta(\hat{\mu}_t - m) - (\ln(1 - \beta) - \beta m) \ln(V_t + (\hat{\mu}_t - m)^2) \)
7: Let \( \beta(m) = \frac{m(\mu_t - m) + V_t(\mu_t - m)^2}{\beta(m)} \)
8: \( m_t = \min \{ m \in [0, \mu_t] : \max \{ G_t^\beta(\beta(m), m), D(\mu_t, m) \} \leq \frac{1}{\delta} \ln(e^{\ell_t} / \delta) \} \) (use binary search)
9: \( \ell_t = \max \{ m_t, u_{t-1} \} \)
10: Let \( G_t^\beta(\beta, m) = -\beta(m - \mu_t) - (\ln(1 + (1 - m)\beta) + (1 - m)\beta \ln(V_t + (m - \mu_t)^2) \)
11: Let \( \beta(m) = \frac{1 + (m - \mu_t)^2}{\beta(m)} \)
12: \( m_t = \max \{ m \in [\mu_t, 1] : \max \{ G_t^\beta(\beta(m), m), D(\mu_t, m) \} \leq \frac{1}{\delta} \ln(e^{\ell_t} / \delta) \} \) (use binary search)
13: \( u_t = \min \{ m_t, u_{t-1} \} \)
14: Output \( \{ \ell_t, u_t \} \)
15: end for
```

We emphasize that such a trick for computational efficiency is possible thanks to our regret-based construction of the confidence sequence. Other betting-based confidence bounds, e.g., [15], cannot be turned into a computationally efficient one because they must run the actual betting algorithm, which necessarily spends \( \Omega(t) \) time complexity at time step \( t \). In contrast, we just need to evaluate the maximum wealth, and any lower approximation of it would result in a valid confidence sequence. Another interesting point is that we are taking the maximum of our lower approximation \( G_t^\beta \) with \( D(\mu_t, m) \) without using a union bound. That is, we are not using the common practice of achieving the best of two different confidence bounds by splitting the target failure probability \( \delta \) into two and instantiating the two confidence bounds with \( \delta \rightarrow \delta/2 \), which has the issue of inflating the confidence width. In fact, one can further tighten up the confidence bounds of Algorithm 2 by taking the maximum of any number of approximations of the maximum log wealth as long as they are monotonic when split at \( m = \hat{\mu}_t \).

**Remark 1.** It is possible to further tighten up the KL-divergence-based lower bound \( D(\mu_t, m) \) on the maximum log wealth as a function of \( X_{t, \min} := \min \{ X_1, \ldots, X_t \} \) and \( X_{t, \max} := \max \{ X_1, \ldots, X_t \} \). Specifically, one can prove a tighter version of Proposition 1: tighten up the first display in the proof of Proposition 1 using the fact that \( X \) lies in \([X_{t, \min}, X_{t, \max}]\) rather than \( X \in [0, 1] \).

**IV. LAW OF THE ITERATED LOGARITHM WITH PORTFOLIO ALGORITHMS**

In the previous section, we used an algorithm with an optimal regret to derive time-uniform concentration inequalities. However, does small regret imply good performance? \( F \)-weighted portfolio with \( F \) equal to the Dirichelet(1/2, 1/2) is optimal up to the additive constant \( \ln \sqrt{2\pi} \) in the log wealth.
Algorithm 3 PRECISE-R70: Portfolio REgret for Confidence Sequences Using [32]

1: **Input**: $\delta \in (0, 1)$, random variables $X_1, X_2, \ldots$ in $[0, 1]
2: \quad \ell_0 = 0, u_0 = 1
3: \quad \text{for } t = 1, 2, \ldots, \text{ do}
4: \quad \quad \mu_t = \frac{1}{\ell_t} \sum_{i=1}^{\ell_t} X_i
5: \quad \quad \text{Wealth}_t(m) = \max_{\beta \in [-1,1]} \prod_{i=1}^{\ell_t} (1 + \beta(X_i - m))
6: \quad \quad \beta_t^*(m) = \arg \max_{\beta \in [-1,1]} \prod_{i=1}^{\ell_t} (1 + \beta(X_i - m))
7: \quad \quad \eta_t = \min_{s=1, \ldots, \ell_t} (X_i - m) \text{sign}(\beta_t^*)
8: \quad \quad \tilde{V}_t(m) = \sum_{i=1}^{\ell_t} (X_i - m)^2
9: \quad \quad \Delta_t(m) := \begin{cases} \frac{1}{\ell_t} \text{min}\{\eta_t(m), \beta_t^*(m), 0\} & \frac{1}{\ell_t} \beta_t^*(m) < 1 \\ 0 & \frac{1}{\ell_t} \beta_t^*(m) \geq 1 \end{cases}
10: \quad \quad \Delta_t(m) = \Delta_t(m) - \Delta_t(m), \quad \delta_t(m) = 1
11: \quad \quad R_t(m) = \min \left\{ \frac{1}{1-\gamma} \text{Wealth}_t(m), \frac{1}{\ell_t} \beta_t^*(m), \Delta_t(m), 0 \right\}
12: \quad \quad \exp \left( -\frac{1}{2(1+\min(\eta_t(m), \beta_t^*(m)))^2} \right) \Delta_t(m) F(\beta_t^*(m))
13: \quad \text{Compute with binary search}
14: \quad \quad m_\ell = \min \{ m \in [0, \mu_t] : \text{Wealth}_t(m) \leq R_t(m) \}
15: \quad \quad m_u = \min \{ m \in [\mu_t, 1] : \text{Wealth}_t(m) \leq R_t(m) \}
16: \quad \text{Output } [\ell_t, m_u]
17: \text{end for}

So, at least if we compare ourselves with the optimal rebalanced portfolio, we are not losing much. However, a minimax regret is not a guarantee of maximum wealth because it is concerned with the worst-case scenario. In fact, there might exist a better algorithm that, for example, has a smaller regret against a smaller class of competitors that might include the optimal one for our specific problem. Indeed, this is exactly the strategy used in [14, Section 7.2] to obtain a time-uniform concentration with a $\log \log t$ dependency rather than a $\log t$. In this section, we explain how to derive a confidence bound with the law of the iterated logarithm rate using a novel portfolio algorithm.

The strategy is very simple: we just use a different mixture prior that has more mass around the optimal betting fraction, which goes to 0 as $t$ goes to infinity. Then, thanks to our novel regret bound, we can again approximate the log wealth of the algorithm with the log wealth of the constant betting fraction strategy minus the regret bound and use Theorem 1.

We will use directly the formulation in terms of betting on a continuous coin, which is

$$\text{Wealth}_t(m) = \int_{-1}^{1} \prod_{i=1}^{\ell_t} (1 + \beta(X_i - m)) dF(\beta), \quad (11)$$

where $F$ is defined similarly to the mixture prior used in [32, Example 3]:

$$F(\beta) = \frac{1}{|\beta|h(1/|\beta|)}, \quad |\beta| \leq 1.$$  

In particular, we will use $h(x) = \frac{2}{\ln e} \ln \frac{x}{\ln x} \ln^2 \frac{x}{\ln x}$ for $c = 6.6 e$. The choice of $c$ assures that $h(x)$ is decreasing in $[0, 1]$. We call the resulting algorithm as PRECISE-R70 (PRECISE using [32]), which is described in Algorithm 3.

Note that here we restrict the betting fractions to $[-1, 1]$ for simplicity. One could consider the entire interval $[-\frac{1}{1-m}, \frac{1}{m}]$ by scaling the prior to fit this interval, but it seems that this choice would result in confidence sequences that depend on $\frac{1}{t}$ and $\frac{2}{\ln t}$ that can be arbitrarily big, which may not be able to place sufficient probability mass around 0. While we show a remedy later in Remark 2, we stick to the range $[-1, 1]$ for ease of exposition.

The following theorem guarantees that Algorithm 3 has the correct coverage and also that its resulting confidence bounds match the rate of the law of the iterated logarithm, which is the asymptotically optimal rate.

**Theorem 8.** The regret of the strategy (11) is bounded by $R_t(m)$ defined in Algorithm 3. Furthermore, let $\delta \in (0, 1)$ and assume $X_1, X_2, \ldots$ be a sequence of random variables such that for each $i$ we have $0 \leq X_i \leq 1$ and $\mathbb{E}[X_i_1, X_i_2, \ldots, X_i_{i-1}] = \mu$ almost surely. Run Algorithm 3 on $X_1, X_2, \ldots$. Then, $\mathbb{P}(\forall t \geq 1, \mu \in [\ell_t, m_u]) \geq 1 - \delta$. Furthermore, we have that

$$\max \left\{ u_t - \mu_t, \mu_t - \ell_t \right\} \leq \frac{1}{t} \sqrt{2 \ln \left( \frac{3U_t}{2-t} 2V_t + \frac{4U_t}{t-2U_t} + \frac{24}{t} \ln \frac{7}{\delta} \right)},$$

where

$$V_t := \sum_{i=1}^{t} (X_i - \mu_i)^2,$$

$$U_t := -\frac{1}{2} W_{-1} \left( -\frac{20}{38} \cdot h \left( \frac{1}{2+\sqrt{V_t}/2} \right) \right) = O(\ln \frac{\ln t}{\delta}),$$

and $W_{-1}$ is the negative branch of the Lambert function. This means that $u_t - \mu_t$ and $\mu_t - \ell_t$ are both upper bounded by $\frac{1}{t} \sqrt{2 \ln \ln t}$ as $t \rightarrow \infty$.

While we aimed for a regret guarantee easy to calculate, it would be possible to have a tighter regret guarantee, obtaining numerically tighter confidence sequences. However, improvements are possible only in the small sample regime, because the confidence sequence we obtained above is already asymptotically optimal.

A similar guarantee was first obtained by [33] and a variance-oblivious one is also implied by Ville’s inequality and [34, Theorem 12]. However, we want to stress the fact that (12) is only an upper bound of the output of Algorithm 3. In practice, the confidence sequences are much tighter. Moreover, as far as we know, the regret upper bound $R_t(m)$ in Algorithm 3 is the first finite-time regret bound of an $F$-weighted portfolio algorithm with prior $F$ being the same type as [32, Example 3].

**Remark 2** (Mixing PRECISE-CO96 with PRECISE-R70). Any convex combination of betting strategies is still a betting strategy. Hence, we can use PRECISE-CO96 and PRECISE-R70 at the same time, allocating, for example, one-half of the initial wealth to each of them. The total wealth is just the sum of their wealth. In fact, this is equivalent to using
a prior that is a mixture of the one of PRECiSE-C096 and the one of PRECiSE-R70. In this way, we can guarantee both the LIL property, thanks to PRECiSE-R70, and never-vacuous confidence sequences, thanks to PRECiSE-C096. Of course, the mixing coefficient becomes a hyperparameter and it could be tuned to trade-off the performance at infinity with the one on few samples.

V. OMITTED PROOFS

A. Proof of Lemma 1

Proof: After $b$ is revealed, the wealth is updated by multiplying the previous wealth by the following:

\[ w_1 b + w_2 (1 - b) = b + b \frac{c}{m} + \left(1 - \frac{c}{1 - m}\right) (1 - b) \]
\[ = b + b \frac{c}{m} + 1 - b - \frac{c}{1 - m} + \frac{c}{1 - m} b \]
\[ = 1 + c \beta, \]

where $\beta$ is the signed betting fraction equal to $-\frac{1}{m} + \frac{1}{m} b$. Given that $b \in [0, 1]$, the range of the betting fractions is in $[-\frac{1}{m}, \frac{1}{m}]$ as we wanted. \qed

B. Proof of Theorem 2

First of all, we leverage Lemma 1 to reduce the problem of betting to a continuous coin to the one of portfolio selection with 2 stocks and solve it with a universal portfolio algorithm. Then, we prove a tighter regret for our universal portfolio algorithm.

1) From Betting on a Continuous Coin to Portfolio Selection: Define the regret on the log-wealth of an online betting algorithm receiving continuous coins $c_t \in [-m, 1 - m]$ where $m \in (0, 1)$ as

\[ \text{Regret}_t := \max_{\beta \in \left[-\frac{1}{m}, \frac{1}{m}\right]} \sum_{i=1}^{t} \ln(1 + \beta c_i) - \sum_{i=1}^{t} \ln(1 + \beta_i c_i). \]

(13)

If we know an upper bound on $\text{Regret}_t$, then it would be immediate to calculate a lower bound on the log wealth of the algorithm.

So, the only thing we need to know is the regret of the best possible betting algorithm. This can be done with Lemma 1, which reduces this problem to the portfolio selection problem with 2 stocks, where an algorithm with the minimax regret is known. We stress that we do not need to use the transformation in Lemma 1 in our algorithm. Instead, it is enough to know that such a transformation exists. This reduction allows us to use any portfolio selection algorithm to bet on an asymmetric continuous coin. Next, we show a tight empirical upper bound on the regret of Dirichlet(1/2,1/2)-weighted portfolio.

2) Data-Dependent Regret With 2 Stocks: Here, we introduce a data-dependent regret upper bound for the $F$-weighted portfolio algorithm.

Reference [27] proved that setting the mixture distribution $F$ equal to the Dirichlet(1/2,1/2) distribution gives an upper bound of the regret of

\[ \text{Regret}_t \leq \ln \frac{\sqrt{\pi^3 (t + 1)}}{\Gamma(t + \frac{3}{2})}, \]

(14)

that is optimal up to constant additive terms. As we anticipated above, this regret is tight in the case that in each round the market gains are exactly one 0 and one 1. However, for other sequences, the regret can be smaller.

So, here we derive an easy-to-calculate data-dependent regret guarantee that is tighter than (14). The following upper bound is essentially in the proofs in [27].

Theorem 9. Denote by $[b_t^*, 1 - b_t^*] = \arg\max_{b \in B} \text{Wealth}_t(b)$. Then, the regret of the Dirichlet(1/2,1/2)-weighted portfolio algorithm satisfies

\[ \text{Regret}_t \leq \max_{0 \leq k \leq t} f(b_t^*, k, t), \]

where $f(b, k, t)$ is defined in (8).

Proof: From [27, Lemma 2], in the 2 stocks case we have

\[ \max_{b \in B} \text{Wealth}_t(b) \leq \max_{0 \leq k \leq t} \frac{(b_t^*)^k (1 - b_t^*)^{t-k}}{\int_0^t b^k (1 - b)^{t-k} dF(b)}. \]

Moreover, from equation (64) in [27], we have

\[ \int_0^t b^k (1 - b)^{t-k} dF(b) = \frac{\Gamma(k + \frac{1}{2}) \Gamma(t - k + \frac{1}{2})}{\left(\Gamma(\frac{3}{2})\right)^2 \Gamma(t + 1)}. \]

Putting all together, we have the stated bound. \qed

Unfortunately, it can be verified numerically that the bound in Theorem 9 does not give rise to an interval. The reason is that $b_t^*$ depends on $m$ and the regret gets smaller when $b_t^*$ is close to 1/2, but at the same time, Wealth($b_t^*$) also decreases in this case. Hence, Wealth($b_t^*$) minus the regret upper bound is not guaranteed to be quasi-convex in $m$, and indeed we verified numerically that we can get non-quasi-convex functions.

So, we propose the following variation. For each $t$ we have to check $m$ in a range equal to the confidence set calculated at $t - 1$. Assume by induction that the confidence set with $t - 1$ samples is an interval $[\ell_{t-1}, u_{t-1}]$, we now construct a new interval. The key idea is the following one: as we vary $m$ in $[\ell_{t-1}, u_{t-1}]$, $b_t^*$ moves monotonically from a positive value to 0. We now calculate the worst value of the regret for all the $b_t^*$ in this range. This is now a value independent of $m$ and so Theorem 4 applies. The same reasoning holds for the interval $m \in (\mu_t, u_{t-1}]$.

The following Lemma characterizes the local maxima of the function appearing in the regret. To facilitate the understanding of the proof, we plot the behavior of the key function $h(b)$ defined in the following lemma for $t = 10$ in Figure 1.

Lemma 2. Let $t \geq 1$ integer and denote by

\[ f(b, k, t) = \ln \frac{b^k (1 - b)^{t-k} \Gamma(t + 1) \Gamma(1/2)^2}{\Gamma(k + 1/2) \Gamma(t - k + 1/2)}. \]
Define $h(b) = \max_{k=0, \ldots, t} f(b, k, t)$. Then, for any $0 \leq b \leq 1$, the following hold:

- $\{[tb - 0.5], [tb + 0.5]\} = \arg\max_{k=0, \ldots, t} f(b, k, t)$.
- The local maxima of the function $h(b)$ are exactly at $\{i/t\}_{i=0}^t$.
- $h(i/t)$ is nonincreasing for $i = 0, \ldots, \lceil(t+1)/2\rceil - 1$ and nondecreasing for $i = t - \lceil(t+1)/2\rceil, \ldots, t$.

**Proof:** Fix $b \in [0, 1]$ and let $k^* \in \arg\max_{x=0, \ldots, t} f(b, k, t)$. This implies that $f(b, k^* + 1, t)$ and $f(b, k^* - 1, t)$ are not larger than $f(b, k^*, t)$. For convenience, let us work with $g(b, k, t) := \exp(f(b, k, t))$. Then,

$$g(b, k^* + 1, t)/g(b, k^*, t) = \frac{b(t - k^* - 1/2)}{(k^* + 1/2)(1 - b)}.$$

This ratio is less than 1 if and only if

$$b(t - k^* - 1/2) \leq (k^* + 1/2)(1 - b),$$

that is

$$k^* \geq b(t - 1/2) - 1/2(1 - b) = bt - 1/2.$$

Moreover,

$$g(b, k^* - 1, t)/g(b, k^*, t) = \frac{(k^* - 1/2)(1 - b)}{b(t - k^* + 1/2)} \leq 1,$$

that is equivalent to

$$k^* \leq b(t + 1/2) + 1/2(1 - b) = bt + 1/2.$$

Putting the two inequalities together with the constraint that $k^*$ is an integer gives the stated expression. Moreover, the conditions above imply that the arg max becomes only one element in $\{0, \ldots, t\}$ in general, but it could be one of the two elements when $tb$ is an integer. In such an edge case, one can see that they both achieve the same objective value. This concludes the first statement of the lemma.

For the second property, for every $k \in \{0, \ldots, t\}$, let $I_k$ be the set of $b$’s for which $k \in \arg\max_{x=0, \ldots, t} f(b, k, t)$. Then, $\{I_k\}$ forms a partition of $[0, 1]$ where they share the boundary with adjacent intervals. That is, we have $\bigcup_{k=0}^t I_k = [0, 1], I_k \cap I_{k+1} = (k + 0.5)/t, \forall k \in \{0, \ldots, t\}$, and $I_k \cap I_j = \{\}$ for all $j, k \in \{0, \ldots, t\}$ with $|j - k| > 1$. On each partition $I_k$, it is easy to see that $h(b)$ is concave and its maximum is $k/t$. Hence, the local maxima of $h(b)$ are exactly at $\{k/t\}_{k=0}^t$. The third property directly follows from the proof of Lemma 4 in [27].

We can now present the upper bounds of the regret.

**Lemma 3.** For any $X_1, \ldots, X_t$ in $[0, 1]$, define $\hat{\mu}_t = \frac{1}{t} \sum_{i=1}^t X_i$, and $\phi(x) = \arg\max_{b \in [0, 1]} \sum_{i=1}^t \ln \left(b \left(1 + \frac{X_i - x}{x}\right) + (1 - b) \left(1 - \frac{X_i - x}{1 - x}\right)\right)$.

Consider running $F$-weighted portfolio with 2 stocks, Dirichlet$(1/2, 1/2)$ mixture, and the market gains $\{(1 + \frac{X_i - m}{m}, 1 - \frac{X_i - m}{1 - m})\}_{i=1}^t$. Then, for any $\ell \in [0, \hat{\mu}_t]$ and $m \in [\hat{\mu}_t, u]$, we have that its regret is upper bounded by

$$\max \left\{ \frac{f \left(\frac{\phi(\ell) \cdot t - 0.5}{t}, \frac{\phi(\ell) \cdot t - 0.5}{t}\right)}{f \left(\frac{\hat{\mu}_t \cdot t + 0.5}{t}, \frac{\hat{\mu}_t \cdot t + 0.5}{t}\right)} \right\}.$$

On the other hand, for any $u \in (\hat{\mu}_t, 1]$ and $m \in [\hat{\mu}_t, u]$, we have that the regret is upper bounded by

$$\max \left\{ \frac{f \left(\frac{\phi(u) \cdot t + 0.5}{t}, \frac{\phi(u) \cdot t + 0.5}{t}\right)}{f \left(\frac{\hat{\mu}_t \cdot t - 0.5}{t}, \frac{\hat{\mu}_t \cdot t - 0.5}{t}\right)} \right\}.$$

**Proof:** We will only prove the first upper bound because the other one is analogous.

Observe that when $m$ increases, the market gains of the first stock decrease, and the ones of the second stock increase. This means that $\phi(x)$ is nonincreasing. Hence, for $m \in [\hat{\mu}_t, u]$ we have that $\phi(m) \in [\phi(\hat{\mu}_t), \phi(\ell)]$.

Recall that $G(\beta, \hat{\mu}_t)$ is maximized at $\beta = 0$ by Theorem 4. Using our reduction in Lemma 1 and noting that $\sum_{i=1}^t \ln \left(b \left(1 + \frac{X_i - m}{m}\right) + (1 - b) \left(1 - \frac{X_i - m}{1 - m}\right)\right) = G(\beta, x)$, one can show that $\phi(\hat{\mu}_t) = \hat{\mu}_t$. Using the properties of the local maxima of the upper bound of the portfolio algorithm in Lemma 2, we have the stated bound. □

We can finally prove Theorem 2.

**Proof of Theorem 2:** It is enough to use Lemma 3 to derive a lower bound of the wealth and use Theorem 1. □

**C. Proof of Theorem 4**

**Proof:** Denote $\beta'(m) := \arg\max_{\beta \in [-\frac{1}{2}, \frac{1}{2}]} G(\beta, m)$.

The derivative of $G$ w.r.t. its first argument is

$$G'(\beta, m) = \sum_{i=1}^t \frac{X_i - m}{1 + \beta(X_i - m)}.$$

Then, we have that $G'(0, \hat{\mu}_t) = 0$. Given that $G(\beta, m)$ is concave in $\beta$, then $G(\beta, \hat{\mu}_t)$ has a maximum w.r.t. the first argument in $\beta = 0$ and the value of the function is 0.

For $m' > \hat{\mu}_t$, $G'(0, m') < 0$. Since $G(\beta, m')$ is concave in $\beta$, we have $\beta^*(m') < 0$. In the same way, for $m' < \hat{\mu}_t$ we have $\beta^*(m') > 0$. □
Let us start with \( m' > \hat{\mu}_t \) and prove that \( H \) is nondecreasing, the other side is analogous. Consider \( m_1 > m_2 > \hat{\mu}_t \). Given that \( \beta^*(m_2) < 0 \), we have
\[
H(m_2) = G(\beta^*(m_2), m_2) \leq G(\beta^*(m_2), m_1)
\leq G(\beta^*(m_1), m_1) = H(m_1),
\]
where the first inequality is due to the fact that \( G(\beta, m) \) is nondecreasing in \( m \) when \( \beta < 0 \) and the second inequality is due to the fact that the negative part of the interval \([-\frac{1}{m'^2}, \frac{1}{m^2}]\) contains the negative part of the interval \([-\frac{1}{m'^2}, \frac{1}{m^2}]\) and we know the maximum \( \beta \) is negative. \( \square \)

D. Proof of Theorem 5

Proof: The algorithm simply numerically inverts the inequality in (10), using two different upper bounds of the regret for the lower and upper values of the confidence intervals, in \( R^t_i \) and \( R^t_i \) respectively. The particular expression of the regret is given in Lemma 3. Note that lines 8 and 12 take care of the maximum over \( t \) in (10). Moreover, Theorem 4 proves that \( L_t \) and \( U_t \) are intervals, so we can use a binary search procedure. From the proof above, it should be clear that we can refine the regret upper bounds at each step of the binary search procedure. However, numerically the advantage is tiny, so for simplicity, we decided not to present this variant.

E. Proof of Theorem 6

Proof: First of all, it should be clear that the width of the confidence intervals \( \ell_t \) are always smaller than the one calculated with 1 sample, \( u_t - \ell_t \). With only one sample, the upper and lower bound have a closed form. Indeed, the argmax of (10) is achieved in \( \beta = \frac{1}{m} \) if \( X_1 - m > 0 \) and in \( \beta = -\frac{1}{m} \) for \( X_1 - m < 0 \). This implies that
\[
\ell_t = \frac{X_1}{\exp(C_1 + \ln \frac{1}{t})} \quad \text{and} \quad u_t = 1 - \frac{1 - X_1}{\exp(C_1 + \ln \frac{1}{t})}.
\]

Moreover, from a direct calculation, we also get that \( C_1 \) is equal to \( \ln 2 \). Subtracting the upper bound from the lower bound, we get the stated bound for any \( X_t \). \( \square \)

F. Proof of Theorem 3

First, we state a technical lemma.

Lemma 4. Let \( f(x) = Ax + B \ln(1 - |x|) + |x| \), where \( A \in \mathbb{R} \) and \( B \geq 0 \). Then,\( \arg \max_{x \in [-1, 1]} f(x) = \frac{A}{A+B} \) and \( \max_{x \in [-1, 1]} f(x) = B \psi(\frac{A}{B}) \geq \frac{A^2}{(4/3)|A|+2B} \), where \( \psi(x) = |x| - \ln(|x| + 1) \).

Proof: If \( B = 0 \), we have that argmax is sign\( (A) \). If \( A = 0 \), the argmax is 0. Hence, in the following, we can assume \( A \) and \( B \) to be different than 0.

We can rewrite the maximization problem as
\[
\arg \max_x f(x) = B \arg \max_x \frac{A}{B} x + \ln(1 - |x|) + |x|.
\]
From the optimality condition, we have that \( \frac{A}{B} = -\frac{\sin(x^*)}{1 - |x^*|} \) and \( \text{sign}(x^*) = 0 \), that implies \( x^* = \frac{A}{|A|+B} \). Substituting this expression in \( f \), we obtain the stated expression. The inequality is obtained by the elementary inequality \( \ln(1+x) \leq x \frac{5+x}{6+4x} \) for \( x \geq 0 \).

We can now prove Theorem 3.

Proof: For a given \( t \), set \( \epsilon_t \) equal to \( \mu - \hat{\mu}_t \), so that \( \epsilon + \mu_t \in [0,1] \). Symmetrizing [31, equation 4.12], we have for any \( |x| \leq 1 \) and \( |\beta| \leq 1 \)
\[
\ln(1 + \beta x) \geq \beta x + (\ln(1 - |\beta|) + |\beta|) x^2.
\]

Hence, for any \( \beta \in [-1, 1] \), we have
\[
\sum_{i=1}^t \ln(1 + \beta_i (X_i - \mu)) = \sum_{i=1}^t \ln(1 + \beta_i (X_i - \hat{\mu}_t - \epsilon_i))
\]
\[
\geq \beta \sum_{i=1}^t (X_i - \hat{\mu}_t - \epsilon_i)
\]
\[
+ (\ln(1 - |\beta|) + |\beta|) \left( \sum_{i=1}^t (X_i - \hat{\mu}_t)^2 + \epsilon_i^2 t - 2\epsilon_i \sum_{i=1}^t (X_i - \hat{\mu}_t) \right)
\]
\[
= -\epsilon_i \beta + (\ln(1 - |\beta|) + |\beta|) \left( \sum_{i=1}^t (X_i - \hat{\mu}_t)^2 + \epsilon_i^2 t \right).
\]

Hence, we have
\[
\max_{\beta \in [-1,1]} \sum_{i=1}^t \ln(1 + \beta(X_i - \hat{\mu}_t - \epsilon_i))
\]
\[
\geq \left( \sum_{i=1}^t (X_i - \hat{\mu}_t)^2 + \epsilon_i^2 t \right) \psi \left( \frac{|\epsilon_i| t}{\sum_{i=1}^t (X_i - \hat{\mu}_t)^2 + \epsilon_i^2 t} \right),
\]
where \( \psi(x) = |x| - \ln(|x| + 1) \) and the equality is due to Lemma 4. From the inequality in Lemma 4 we also obtain
\[
\max_{\beta \in [-1,1]} \sum_{i=1}^t \ln(1 + \beta(X_i - \hat{\mu}_t - \epsilon_i))
\]
\[
\geq \frac{\epsilon_i^2 t^2}{(4/3)|\epsilon_i| t + 2 \sum_{i=1}^t (X_i - \hat{\mu}_t)^2 + 2\epsilon_i^2 t}.
\]

Now, note that for any \( \mu \in [0,1] \) the interval \([-\frac{1}{1 + \mu}, \frac{1}{1 + \mu}] \) is contained in \([-1, 1] \). Hence, from Theorem 2, uniformly on all \( t \) with probability at least 1 - \( \delta \), we have
\[
\frac{\epsilon_i^2 t^2}{(4/3)|\epsilon_i| t + 2 \sum_{i=1}^t (X_i - \hat{\mu}_t)^2 + 2\epsilon_i^2 t} \leq \text{Regret}_t + \ln \frac{1}{\delta}.
\]
Assuming \( \epsilon_t \) positive and solving for it, we have the stated upper bound. By the symmetry of the formula, the expression for negative \( \epsilon_t \) has the opposite sign. \( \square \)

G. Proof of Theorem 7

Proof: We focus on showing that
\[
P \left( \max_t t \cdot G_t(\mu) - R_t \geq \ln(1/\delta) \right) \leq \delta
\]
and that, for every \( t \geq 1 \), the confidence set for time \( t \)
\[
\left\{ m \in \mathbb{R} : t \cdot G_t(m) - R_t \geq \ln(1/\delta) \right\}
\]
is an interval. The proof relies on the fact that \( D(\hat{\mu}_t, \mu) \) lower bounds the maximum wealth due to Proposition 1, \( D(\hat{X}_t, x) \) is piece-wise monotonic when split at \( x = \hat{\mu}_t \), and the
maximum of two nonincreasing (nondecreasing) function is
nonincreasing (nondecreasing) respectively.

First, we show a tight lower bound on the log weight that
depends on the sign of $\beta$. Define $c_i := X_i - m$. Throughout the
proof, we drop the subscript $t$ from $V_t, G^t_1, G^t_2$, etc. to reduce clutter.

If $\beta \in [0, 1/m]$, using [31, Eq. 4.11] with $\lambda = \beta m \in [0, 1]$ and
$\xi = c_i/m \geq -1$, we have

$$\ln(1 + \beta c_i) \geq \beta c_i - \beta^2 c_i^2 = \frac{\beta^2 c_i - \beta(1 - \beta m) - \beta m}{\beta^2 m^2},$$

where we define $\ln(0)$ as $-\infty$. Recall that $\bar{V} = \frac{1}{t} \sum_{i=1}^{t} (X_i -
\bar{\mu}_t)^2$. Then, with some algebra, one can show that

$$\max_{\beta \in [0,1/m]} \frac{1}{t} \sum_{i=1}^{t} \ln(1 + \beta c_i) \geq \max_{\beta \in [0,1/m]} G^t_1(\beta, m).$$

A simple calculation tells us that the maximum is achieved at
$\bar{\beta}_t(m)$ (defined in Algorithm 2).

If $\beta \in [-1/(1 - m), 0]$, we can employ a similar argument to derive

$$\max_{\beta \in [-1/(1 - m), 0]} \frac{1}{t} \sum_{i=1}^{t} \ln(1 + \beta c_i) \geq \max_{\beta \in [-1/(1 - m), 0]} G^u(\beta, m).$$

Once again, it is easy to show that the maximum is achieved at
$\bar{\beta}_u(m)$ (defined in Algorithm 2).

Let $\beta^*(m) = \max_{\beta \in [-1/(1 - m), 1/m]} \sum_{i=1}^{t} \ln(1 + \beta(X_i - m))$. Examining the gradient of the objective function at $\beta = 0$, one can see that $\bar{\mu}_t \geq \bar{\mu}$ implies that $\beta^*(\mu) \geq 0$ and $\bar{\mu}_t \leq \bar{\mu}$ implies that $\beta^*(\mu) \leq 0$.

If $\bar{\mu}_t \geq \bar{\mu}$, then using $\beta^*(\mu) \geq 0$ we have

$$\sum_{i=1}^{t} \ln(1 + \beta^*(\mu) \cdot (X_i - \mu)) = \max_{\beta \in [0,1/\mu]} \sum_{i=1}^{t} \ln(1 + \beta c_i) \geq \max_{\beta \in [0,1/\mu]} G^t_1(\beta, \mu).$$

Otherwise, we have

$$\sum_{i=1}^{t} \ln(1 + \beta^*(\mu) \cdot (X_i - \mu)) \geq \max_{\beta \in [-1/(1 - \mu), 0]} G^u(\beta, \mu).$$

Hence, from Theorem 2, with probability at least $1 - \delta$, we have for all $t$

$$G_t(\mu) \leq \frac{1}{t} \ln(e^{\text{Regret}_t} / \delta) \leq \frac{1}{t} \ln(e^{\bar{R}_t} / \delta).$$

It remains to show that the confidence set stated in the theorem forms an interval. In this proof, we focus on showing the
lower side only (i.e., the confidence set intersecting with $[0, \bar{\mu}_t]$), as the proof for the upper side is symmetric.

Let $\beta(m) = \arg \max_{\beta \in [0,1/m]} G^t_1(\beta, m)$ and $F(m) = \max_{\beta \in [0,1/m]} G^t_1(\beta, m)$. It suffices to show that $F$ is non-increasing in $[0, \bar{\mu}_t]$. That is, we claim that if $m_1 < m_2$ then

$$F(m_1) = G^t_1(\beta(m_1), m_1) \geq G^t_1(\beta(m_2), m_2) = F(m_2).$$

To prove the claim, we will show below that $G^t_1(\beta(m_2), m_1) \geq G^t_1(\beta(m_2), m_2)$. If this is true, then we have

$$F(m_1) = G^t_1(\beta(m_1), m_1) \geq G^t_1(\beta(m_2), m_1) \geq G^t_1(\beta(m_2), m_2) = F(m_2),$$

where the first inequality is due to the definition of $\beta(m_1)$ and the fact that $[0,1/m] \supseteq [0,1/m_2]$. Now, let us show that $G^t_1(\beta(m_2), m_1) \geq G^t_1(\beta(m_2), m_2)$.

To see this, it suffices to show that $G^t(\beta, m)$ is nonincreasing in $m \in [0, m_2]$ where $\beta \in [0, 1/m_2]$. Therefore,

$$\frac{\partial}{\partial m} G^t(\beta, m) = \frac{1}{m^3} m^2 \left(1 - \beta m \right) \left(1 + 2 \ln(1 - \beta m) / \beta m^2 \right) \left(\bar{V} + (\bar{\mu}_t - m)^2 \right) =: A$$

It suffices to show that $A \geq 0$. Using the fact that $q := \beta m \in [0, 1]$, we have

$$A = \frac{1}{1 - q} + 1 + \frac{2 \ln(1 - q)}{q}.$$

Let $y = 1 - q$. Using the elementary inequality $\ln(1 - q) = \ln(y) \geq -\frac{1-y}{y}$, $\forall y \in [0, 1]$, we have that $\ln(y) \geq -\frac{1-y}{y} \geq -\frac{1}{y}$, $\frac{1+y}{2} = -\frac{1}{2(1-q)}$. With this bound, we have

$$A \geq \frac{1}{1 - q} + 1 + \frac{2 - q}{1 - q} = 0.$$

\hfill \square

H. Proof of Theorem 8

First, we state two technical lemmas.

Lemma 5. Assume $\beta \in [0, 1]$, and $\beta \leq \frac{1}{4}(1 + \beta^2)$. If

$A \geq 4$, then $\beta \leq \frac{2}{A - 2 + \sqrt{(A - 2)^2 - 4}}$. Furthermore, if $A \geq 5$, $\beta \leq \frac{2}{A - 2 + \sqrt{(A - 2)^2 - 4}}$,

$\frac{A^2 + 2A - \sqrt{A^2(A - 4)}}{A - 2 + \sqrt{(A - 2)^2 - 4}}$. Note that $\frac{A^2 + 2A - \sqrt{A^2(A - 4)}}{A - 2 + \sqrt{(A - 2)^2 - 4}}$ is nonincreasing in $A$. Using the bound $A \geq 5$ concludes the proof.

Lemma 6. Let $A > 0$, $\ln(A) + \frac{2}{A} \geq 1$, and $x \leq A \ln(x) + B$.

Then,

$$x \leq -A \cdot W_{-1} \left(1 - A \exp \left(\frac{B}{A}\right)\right) \leq A \ln A + B + A \sqrt{2 \left(\ln A + \frac{B}{A} - 1\right)},$$

where $W_{-1}$ is the negative branch of the Lambert function and the first inequality is tight.

Proof: To obtain the first inequality we observe that

$$x \leq A \ln x + B \iff \frac{x}{A} \leq \ln \left(\frac{x}{A}\right) + \ln A + \frac{B}{A}.$$
Hence, we can solve the associated equality directly obtaining that the biggest solution is given by \( \frac{x}{\lambda} = W_{-1} \left( -\frac{1}{\lambda} \exp \left( -\frac{1}{\lambda} \right) \right) \), which also gives the solution of the inequality. The second one is from the lower bound on the Lambert function in [35]. □

Next, we show that the confidence sequences derived by the wealth of any F-weighted portfolio algorithm in (5) are always intervals.

**Theorem 10.** Let \( c_i \in [0,1] \). Then, the function \( F(m) = \int_1^t \prod_{i=1}^t (1 + \beta(c_i - m)) \) is convex in \( m \in [0,1] \).

**Proof:** For a fixed \( \beta \in [-1, 1] \), define \( G(m) = \prod_{i=1}^t (1 + \beta(c_i - m)) = \exp(\sum_{i=1}^t \ln(1 + \beta(c_i - m))) \). We have that

\[
G''(m) = \exp \left( \sum_{i=1}^t \ln(1 + \beta(c_i - m)) \right) 
\times \left[ \left( \sum_{i=1}^t \frac{-\beta}{1 + \beta(c_i - m)} \right)^2 - \sum_{i=1}^t \frac{\beta^2}{(1 + \beta(c_i - m))^2} \right] \geq 0.
\]

Hence, \( F(m) \) is convex. □

We can now prove the theorem.

**Proof of Theorem 8:** For brevity of notation, we drop the dependency on \( m \) on the quantities defined in the algorithm, i.e., \( V_1 := V_1(m) \). Also, for a fixed \( m \), define \( \theta_i := \sum_{i=1}^t (X_i - m) \) and \( \text{Wealth}_i(\beta) := \prod_{i=1}^t (1 + \beta(c_i - m)) \). From the above, it is easy to verify that \( V_t = V_t(\beta^*_t) + \frac{1}{2} \theta_t^2 \).

Note that from the log-concavity of the wealth function, we have for any \( \beta_2 > \beta_1 \geq 0 \) that

\[
\int_{\beta_1}^{\beta_2} \text{Wealth}(\beta) F(\beta) \, d\beta \geq \text{Wealth}(\beta_2) \int_{\beta_1}^{\beta_2} \text{Wealth}(\beta) \, d\beta \\
= (\beta_2 - \beta_1) F(\beta_2) \int_{0}^{1} \text{Wealth}(\beta_1 (1 - a) + a \beta_2) \, d\beta \\
\geq (\beta_2 - \beta_1) F(\beta_2) \int_{0}^{1} \text{Wealth}^{1-a}(\beta_1) \text{Wealth}^a(\beta_2) \, d\beta \\
= F(\beta_2)(\beta_2 - \beta_1) \frac{\text{Wealth}(\beta_2) - \text{Wealth}(\beta_1)}{\ln \left( \frac{\text{Wealth}(\beta_2)}{\text{Wealth}(\beta_1)} \right)}.
\]

Our first term in the regret is obtained using the fact that \( \frac{x-1}{\ln x} \geq \sqrt{x} \), obtaining

\[
\text{Wealth}_i \geq F(\beta^*_t) \frac{\text{Wealth}(\beta^*_t) - \text{Wealth}(0)}{\ln \left( \frac{\text{Wealth}(\beta^*_t)}{\text{Wealth}(0)} \right)} \\
= \frac{F(\beta^*_t) \text{Wealth}(\beta^*_t) - 1}{\ln \text{Wealth}(\beta^*_t)} \\
\geq \sqrt{\text{Wealth}_i(\beta^*_t) |\beta^*_t| F(\beta^*_t)}.
\]  

(17)

For the second term in the regret, we use a Taylor expansion of the log wealth. Denote by \( f(\beta) = \ln \text{Wealth}(\beta) \). Hence, when \( |\beta^*_t| < 1 \), for some \( \beta \) between \( \beta^*_t \) and \( \beta^*_t - \Delta_t \text{sign}(\beta^*_t) \), we have

\[
f(\beta^*_t - \Delta_t \text{sign}(\beta^*_t)) = f(\beta^*_t) - \Delta_t \text{sign}(\beta^*_t) f'(\beta^*_t) + \frac{\Delta_t^2}{2} f''(\beta^*_t).
\]

From the definition of \( q_t \) and \( \Delta_t \), in the algorithm, for any \( \beta \) between \( \beta^*_t \) and \( \beta^*_t - \Delta_t \text{sign}(\beta^*_t) \), we have

\[
f''(\beta) = -\sum_{i=1}^t \frac{(X_i - m)^2}{(1 + \beta(X_i - m))^2} \geq -\sum_{i=1}^t \frac{(X_i - m)^2}{(1 + \beta q_t)^2}.
\]

Hence, for \( |\beta_t^*| < 1 \) we have

\[
f(\beta^*_t - \Delta_t \text{sign}(\beta^*_t)) \geq f(\beta^*_t) + \frac{\Delta_t^2}{2} f''(\beta) \\
\geq f(\beta^*_t) - \Delta_t^2 V_t \\
= \frac{2}{V_t (1 + \min(\beta_t^* q_t, 0))^2} \\
\]

where the presence of the min is necessary to use \( \beta_t^* \) in the denominator. Using this expression in the integral of the wealth, we have

\[
\int_{-\infty}^1 \text{Wealth}_i(\beta) F(\beta) \, d\beta \\
\geq \text{Wealth}_i(\beta^*_t) \Delta_t \exp \left( -\frac{\Delta_t^2 V_t}{2(1 + \min(\beta_t^* q_t, 0))^2} \right) F(\beta^*_t).
\]

(18)

Putting together (17) and (18), we have the expression of the regret of the portfolio algorithm (i.e., \( R_t(m) \) in the algorithm).

Now, we turn our attention to the expression of the confidence sequences. In the following, we safely assume that \( \theta_t > \sqrt{2V_t} \) since otherwise we obtain the desired bound.

First, we need to study \( \beta^*_t \). If \( |\beta^*_t| < 1 \), then using the Taylor’s remainder theorem, we have \( f'(0) + \beta_t^* f''(\beta) = f(\beta^*_t) \), where \( \beta \) is between 0 and \( \beta_t^* \). This implies that

\[
\sum_{i=1}^t (X_i - m) - \beta_t^* \sum_{i=1}^t \frac{(X_i - m)^2}{(1 + \beta(X_i - m))^2} = 0
\]

for some \( \beta \) between 0 and \( \beta_t^* \). This implies that \( |\beta_t^*| \leq |\theta_t| + 5 \left| \frac{\theta_t}{V_t} \right|^2 \). Using Lemma 5 to solve the second inequality, if \( V_t / \theta_t > 5 \) then we obtain \( |\beta_t^*| \leq |\theta_t| + 5 \left( \frac{\theta_t}{V_t} \right)^2 \).

Also, solving the first inequality, we get

\[
|\beta_t^*| \geq \frac{1}{2} \left( \frac{V_t}{\theta_t} + 2 - \sqrt{\frac{V_t}{\theta_t} + \frac{V_t}{\theta_t} + 4} \right) \\
= \frac{2}{V_t/\theta_t + 2 - \sqrt{\frac{V_t}{\theta_t} + 2}} \geq \frac{|\theta_t|}{2|\theta_t| + V_t}.
\]

(19)

If \( |\beta^*_t| = 1 \), then there are two cases: either the absolute value of the unconstrained maximizer is 1 or it is bigger than 1. In the first case, \( f'(\beta^*_t) > 0 \) if \( \beta_t^* = 1 \) and \( f'(\beta^*_t) < 0 \) if \( \beta_t^* = -1 \). Using the fact that \( \beta_t^* \) and \( \theta_t \) have the same sign, reasoning as above in both cases we have

\[
\sum_{i=1}^t (X_i - m) - \sum_{i=1}^t \frac{(X_i - m)^2}{(1 + \beta(X_i - m))^2} \geq 0
\]

for some \( |\beta| \leq 1 \), which implies \( V_t / |\theta_t| \leq 4 \).

We now do a case analysis.

**Case 1.** \( V_t \geq 5|\theta_t| \)
Our analysis above for the case of $|\beta^*_t| = 1$ implies that $V_t \leq 4|\theta_t|$, which contradicts the condition of Case 1. Thus, $|\beta^*_t| < 1$ which from (19) implies that $|\beta^*_t| \geq \frac{\theta_t}{\sqrt{\theta_t^2 + 2V_t}}$.

Given that $\Delta_t \leq \frac{1}{\min\{\theta_t^2, \beta_t^2\}}$, from (18) we obtain

$$\ln \text{Wealth}_t(\beta^*_t) - \ln \text{Wealth}_t \leq \ln \left( \frac{\sqrt{\frac{\theta_t}{\Delta_t}}}{\sqrt{F(\beta^*_t)}} \right).$$

So, from Ville’s inequality, with probability at least $1 - \delta$, we have that, $\forall t \geq 1,$

$$\max_{\beta \in [-1, 1]} \ln \text{Wealth}_t(\beta) \leq \ln \left( \frac{\sqrt{\frac{\theta_t}{\Delta_t}}}{\sqrt{F(\beta^*_t)}} \right) + \ln \frac{1}{\delta}.$$ \hspace{1cm} (20)

Now, it remains to figure out the range of $m$ that satisfies the inequality above.

Define $\beta^*_t := \frac{\theta_t}{(3/4)|\theta_t| + 2V_t}$. Using Lemma 4 and (15), we have

$$\ln \text{Wealth}_t(\beta^*_t) \geq \ln \text{Wealth}_t(\beta^*_t) \geq \frac{\theta_t^2}{(4/3)|\theta_t| + 2V_t}.$$ \hspace{1cm} (21)

We now need a lower bound for $\Delta_t$. If $\Delta_t = \Delta_t$, then $\Delta_t \geq \frac{1}{\sqrt{V_t}}$. Indeed, if $\Delta_t \neq \Delta_t$, then by its definition we have $\Delta_t \geq \frac{|\beta^*_t|}{1 - |\beta^*_t|}$.

$$\frac{|\beta^*_t|}{1 - |\beta^*_t|} \geq \frac{|\theta_t|}{|\beta_t| + V_t} \geq \frac{\sqrt{2V_t}}{\frac{|\theta_t|}{2V_t}} \geq \frac{1}{\sqrt{V_t}} \Rightarrow 1 - |\beta_t^*| < |\beta_t^*|.$$ \hspace{1cm} (22)

Hence, in both cases we have $\Delta_t \geq \frac{1}{\sqrt{V_t}}$. Thus, using (20) and the lower bound of $\Delta_t$, we obtain

$$\frac{\theta_t^2}{\frac{1}{3}|\theta_t| + 2V_t} \leq \ln \left( \frac{\sqrt{\frac{\theta_t}{\Delta_t}}}{1 - |\beta^*_t|} \right).$$

Since $|\beta^*_t| \leq \frac{|\theta_t|}{\sqrt{V_t}} + 5 \left( \frac{\theta_t}{V_t} \right)^2$, we have $|\beta^*_t| \leq \frac{|\theta_t|}{\sqrt{V_t}} + 5 \left( \frac{\theta_t}{V_t} \right)^2$. So, using $1/(1 - |\beta_t^*|) \leq \frac{3}{5}$, $|\beta_t^*| \leq \frac{5 |\theta_t|}{V_t}$, and $h(\beta^*_t) \leq h(\frac{\theta_t}{5|\theta_t| + 2V_t})$, we have

$$\frac{\theta_t^2}{\frac{1}{3}|\theta_t| + 2V_t} \leq \ln \left( \frac{10\sqrt{\frac{2|\beta_t^*|}{\theta_t} + \sqrt{\frac{\theta_t}{2V_t}}}}{3\delta} \right).$$

$$\leq \ln \left( \frac{10\sqrt{\frac{2|\beta_t^*|}{\theta_t} + \sqrt{\frac{\theta_t}{2V_t}}}}{3\delta} \right).$$

From the definition of $h(x)$, we have that the minimum of $h$ in $[\sqrt{2}, 1]$ is greater than 6. So, using Lemma 6 with $x = \frac{\theta_t^2}{\frac{1}{3}|\theta_t| + 2V_t}$, $A = \frac{1}{2}$, and $B = \ln \left( \frac{20}{33} h \left( \frac{1}{2 + \sqrt{\theta_t^2/2}} \right) \right)$, we have $\frac{\theta_t^2}{\frac{1}{3}|\theta_t| + 2V_t} \leq U_t$, where

$$U_t := \frac{1}{2} W_{-1} \left( \frac{2}{\frac{2}{3 \delta} - h \left( \frac{2}{3 \delta} \right) \left( \frac{e^{-2/(3 \delta)} - 1}{2} \right)^{3/2}} \right).$$

Finally, use the fact that $V_t = V_t(\hat{\mu}_t) + \theta_t^2/t$ in $\theta_t^2 \leq \frac{(4/3)|\theta_t|}{t} + 2V_t U_t$ to obtain $|\theta_t| \leq \frac{4\delta U_t}{2t \sqrt{\theta_t^2} + \sqrt{2V_t \theta_t^2}}.$

Case 2: $V_t/|\theta_t| < 5$

In this case, either $|\beta_t^*| = 1$ or $|\beta_t^*| < 1$ and from (19) we have that $|\beta_t^*| \geq \frac{1}{2 + |\theta_t|} \geq \frac{1}{2}$. Hence, from (17) and Ville’s inequality we have

$$\ln(\text{Wealth}_t(\beta_t^*)) \leq \frac{1}{2} \ln(\text{Wealth}_t(\beta_t^*)) + \ln \frac{7}{F(1/\delta)} \Rightarrow \frac{1}{2} \ln(\text{Wealth}_t(\beta_t^*)) \leq \ln \frac{7}{F(1/\delta)}.$$ (Lemma 4 and (15))

$$\Rightarrow |\theta_t| < 24 \ln \frac{7}{F(1/\delta)}. \hspace{1cm} (V_t < 5|\theta_t|)$$

In this case, we do not have a guarantee that the set of $m$ that satisfies the inequalities above is an interval. However, from Theorem 10, we know that if you could invert exactly the wealth inequality we would obtain intervals. Hence, given that we have derived an upper bound on the regret, any interval $[\ell_t, u_t]$ found by the algorithm is always a valid one.

Finally, let us now focus on the asymptotic behavior of the term $U_t$ that appears in Case 1. Denote by $C_t = \ln \left( \frac{20}{33} \cdot h \left( \frac{1}{2 + \sqrt{\theta_t^2/2}} \right) \right)$ and by $C_t' = \ln \left( \frac{20}{33} \cdot h \left( \frac{1}{2 + \sqrt{\theta_t^2/2}} \right) \right)$. Using the second inequality of Lemma 6 and $V_t \leq t$, we have $U_t \leq C_t + \sqrt{C_t - \frac{1}{2}} \leq 2C_t'$, hence, given that and $h(x) \leq K_1 \ln^2 (1/x)$ for a suitable constant $K_1$, we have $U_t = O(\ln \ln t)$ as $t \to \infty$. Now, we bound $V_t$ using the fact that $\theta_t^2 \leq \frac{\theta_t}{3|\theta_t|} U_t + 2V_t U_t \leq 3V_t U_t$ which implies for $t$ large enough and for a suitable universal constant $K_2$ that $V_t \leq \frac{1}{3V_t U_t} \leq \frac{1}{K_2 \ln \frac{\theta_t}{\sqrt{V_t}}}$. Denoting by

$$C_t'' = \ln \left( \frac{20}{33} \cdot h \left( \frac{1}{2 + \sqrt{\theta_t^2/2}} \right) \right),$$

we have $U_t \leq C_t'' + \sqrt{C_t''}$ and $\lim_{t \to \infty} \frac{1}{\ln \left( \frac{1}{2 + \sqrt{\theta_t^2/2}} \right)} \leq \frac{1}{4C_t''} = 1.$

VI. NUMERICAL EVALUATION

To test the empirical performance of our proposed procedures, we consider the same set of distributions in [15]. They consider four different sequences of i.i.d. random variables: Bernoulli(0.1) (Figure 2), Bernoulli(0.5) (Figure 3), Beta(1,1) (Figure 4), and Beta(10,30) (Figure 5). We repeat each experiment 10 times and use the first 5 repetitions in which all algorithms result in a valid confidence bound at all time steps. This is because all the algorithms/inequalities do fail
with probability at most $\delta$, so reporting an average over all the runs is misleading.

As baseline methods, we consider the numerical inversion of the betting algorithms ConBO+LBOW, Hedged, and dKelly in [15], and the empirical Bernstein LIL concentration in [2, Equation 4.2] for sub-Bernoulli random variables. We implemented the PRECiSE algorithms in Matlab, our code is available at https://github.com/bremen79/precise. We used a modified Newton update with projections for lines 5 and 9 of PRECiSE-CO96 and PRECiSE-R70. For binary search, we stop when the interval width is less than $10^{-4}$. For the algorithms in [15], we used their Python code.\textsuperscript{4} For dKelly we used their default setting of 10 different bets, while for ConBo and Hedged we used a discretization of 1000 points. For all methods, we set $\delta = 0.05$ for the target failure probability.

From the empirical results shown in Figure 2-5, we see that PRECiSE-CO96 dominates all the other methods in the small sample regime. It is particularly striking that it already has a nonvacuous confidence bound (i.e., the width being strictly smaller than 1) with only one sample for any $\delta \in (0, 1)$.

\textsuperscript{4}https://github.com/WannabeSmith/confseq
as proved in Theorem 6. Such a behavior would be impossible with the empirical Bernstein-type results like [9] and [36] taking the style of Theorem 3 since the lower order term $1/t \ln(1/\delta)$ becomes greater than 1 for a small enough $\delta$. Specifically, the never-vacuous property is possible because we do not restrict the range of the betting fraction as in previous work but use the full range $[-1/m, 1/m]$, allowing a fast increase of the wealth even with a few samples. This full range can be handled thanks to the $F$-weighted portfolio algorithm that works even with unbounded loss functions.

When the number of samples increases, the methods in [15] have only marginal gains, which seem to disappear in the large sample regime. This is probably due to the fixed discretization used in all of their methods, which does not allow the correct behavior at infinity, especially in the dKelly method. Moreover, when the number of bets increases to infinity in dKelly, we recover the $F$-weighted portfolio algorithm with the uniform mixture, which has a worse regret than that of the Dirichlet(1/2, 1/2) mixture [27].

Finally, with a large enough number of random variables the LIL confidence sequences produced by PRECISE-R70 dominate all the other methods, even the one in [2], at least with $10^5$ observations.

One common concern on time-uniform confidence intervals is that they can be too conservative in practice. To study this issue, we compared the behavior of PRECISE-CO96 with a small number of samples with the so-called “exact” confidence intervals in [37], see Figure 6. It is known that the confidence intervals calculated with the method in [37] are not the tightest ones, see, e.g., the discussion in [29], yet they are the easiest ones to compute for Bernoulli random variables. On the other hand, the confidence intervals in [37] are not uniform over time, so we are providing an unfair advantage to [37]. Surprisingly, PRECISE-CO96 closely tracks the exact confidence interval from the very first sample without knowledge of the distribution at hand. Hence, a practitioner might be reassured by the fact that our bounds do not lose too much compared to the approach in [37] yet enjoys a time-uniform guarantee, providing the benefit of being able to decide the number of observations (i.e., determine when to stop collecting observations in sequential experiments) in a data-dependent way.

VII. HISTORY AND RELATED WORK

As noted in [3], the interest on the connection between betting and probability exploded in the last couple of years. However, recent studies may have missed the fact that most of these ideas seem to have been discovered in the information theory literature. Hence, we find it important to discuss here not only related work but also a detailed history of these ideas.

A. Betting and Portfolio Selection Algorithms

The distributional approach to betting and gambling was pioneered by [38]. This approach assumes that the market gains are i.i.d. from a (known) distribution. [38] also noted the equivalence between gambling and prediction with log loss. Moreover, given the well-known connection between universal compression, prediction with log loss, and gambling [27], the literature on compression and prediction with log loss, e.g., [21], [39], [40], [41], [42], and [43], effectively deals with gambling too. In particular, the best achievable regret for a betting game on the outcomes of a coin with a known number of rounds was reported by [39]. Asymptotic optimality in the stochastic case was obtained by a mixture of fixed betting strategies based on the Jeffreys prior of the Bernoulli distribution by [44] and analyzed in finite-time in [45]. References [46], [47], and [48] removed the distributional assumptions and focused on the problem of betting on a bounded outcome (even in multiple dimensions). Later, the online setting, i.e., without distributional assumptions, became standard in the universal compression literature.

Reference [34] showed that it is possible to design an online betting algorithm with a $\ln T$ regret against a weaker competitor. Reference [18] introduced a generic potential framework to design and analyze online algorithms for the same problem. Their work started a line of research on coin-betting algorithms with efficient updates, e.g., [26] and [49], but an outsider to this subfield might have missed the fact that all of them are known to be suboptimal. Indeed, none of them guarantee any regret for betting on a continuous coin $g_t \in [-1, 1]$ against the best constant betting fraction. For example, the known regret bound (e.g., [18, Section 4]) of the KT estimator is relative to the betting fraction $\frac{1}{T} \sum_{t=1}^T g_t$, which is the best constant betting fraction for a discrete coin $g_t \in \{-1, 1\}$ rather
than that of the continuous coin. In fact, a minimax optimal algorithm for continuous coins is already known. To see this, one can use the folklore knowledge that symmetric betting on a bounded symmetric outcome can be easily reduced to portfolio selection with two stocks. In turn, [27] improved upon [19] and proposed a universal portfolio algorithm that, for the first time, achieves the minimax regret without assumptions over the market gains. On a related note, betting on an asymmetric bounded outcome can be also reduced to portfolio selection with 2 stocks as we show in Lemma 1. Lastly, note that the family of online betting algorithms is larger than one might think. Indeed, [17, Theorem 5.16] proved that any online linear optimization algorithm is equivalent to a betting algorithm with initial money equal to its regret against the null competitor.

B. Testing and Betting Algorithms

The connection between betting and testing is very old. The first work on the connection between betting strategies and probability is in the PhD thesis of [20]. He proved that an event depending on an infinite sequence of outcomes has a zero probability iff it is possible to achieve infinite wealth by betting on the outcomes of the sequence. In particular, he defined a martingale as the wealth of a betting strategy on a fair coin. [20, page 34] also has the very first appearance of the wealth process of a betting strategy based on the Laplace method of counts, which is more than 30 years before [50] uses it in coding. However, as [51, page 431] points out, “Ville […] did not consider statistical testing, and his work has had little or no influence on mathematical statistics.” The ideas of Ville were later used to design ideal tests for randomness of infinite sequences [52], [53], [54] where “ideal” means that none of these tests are computable.

We found two independent papers that explicitly link hypothesis testing on a finite sequence of outcomes to betting. One is [55, Example 3], where an optimal betting strategy is defined in terms of the null hypothesis. The other one is [56] that constructs confidence sequences from novel betting schemes and explicitly recognize the connection to the sequential probability ratio test [57]. However, while in the information theory literature these ideas flourished and gave birth to results on coding, compression, minimum description length, and gambling, they seem to have disappeared from the statistics community for 30 years.6 One might argue that any result on martingales is still related to gambling, but, even assuming that this view is correct, it ignores the fact that the gambling view stresses various aspects that are missing from the martingale literature including the computability of the strategies, their optimality, the adversarial nature of the outcomes, and the gambler’s wealth as evidence.

In fact, in the statistics literature gambling strategies reappeared only in the ’90-’00s thanks to the book and papers by Shafer and Vovk. In particular, [12], [58] aimed to found the notion of probability based on a game-theoretic ground through betting schemes. They also proposed the idea of using non-negative martingales as test martingales [51], [59]. However, in the foundational approach in [12] all the betting strategies do not have closed-form expressions and cannot be easily implemented. Moreover, [12] does not contain any explicit concentration, while the first concentration for game-theoretic probability derived by a betting scheme is in [60], which derives a game-theoretic Hoeffding’s inequality.

On the other hand, in the information theory literature, the use of compression schemes to test the randomness of a finite string of symbols became a standard strategy [61], [62], [63], [64]; there is even a software package made available by the American National Institute of Standards and Technology [65]. Given the connection between compression and gambling [27], these approaches can also be considered as tests based on betting.

Even if the idea of numerically deriving valid confidence sequences from any betting algorithm does not appear in print in the work of Shafer and Vovk, it is an “obvious” corollary of their results.7 That said, the first paper to consider an implementable strategy for testing through betting is by [13]. Directly building on [12] and [59], Reference [13] proposes to construct testing martingales and confidence sequences for

6It is remarkable that the work of [55] was submitted to Annals of Statistics and probably rejected, as it can be inferred from the footnote on its first page.

7G. Shafer, 2022, personal communication.
bounded random variables as uniform mixtures of constant betting strategies. Reference [13] also showed empirically the good performance of the proposed approach in a simple statistical test. Reference [14] show how to easily derive a Law of the Iterated Logarithm (LIL) for sub-Gaussian random vectors in Banach spaces from the regret of a one-dimensional betting algorithm. However, given that they were aiming for confidence sequences in closed form, their bounds do not depend on the maximum log wealth and so they cannot obtain the non-vacuous bounds we obtain in this work. In turn, the work of [14] was based on the seminal work in [66] that showed an equivalence between the regret guarantee of online learning algorithms with linear losses and concentration inequalities. However, the proof technique in [14] is different from the one in [66] and it is specific to online algorithms that guarantee a non-negative exponential wealth for biased inputs. In particular, it allows us to derive time-uniform concentrations including those with the rate of the law of the iterated logarithm, which is not possible with the method in [66]. Reference [15] seem to follow the same approach as [13] but propose a number of heuristic betting algorithms to maximize the wealth as well as a discrete version of the uniform mixture that appeared in [13]. They show very good empirical results and asymptotic rates of the obtained confidence sequences. As in previous work, e.g., [15] and [26] guarantee non-negative wealth for their betting heuristics by restricting the range of the allowed betting fractions to, e.g., 1/2 or 2/3. This restriction is exactly the reason why their confidence sequences are vacuous in the small sample regime, as shown in the experiments in Section VI. Moreover, they also need heuristic betting schemes to satisfy strict conditions in order to guarantee that the confidence sets are intervals. On the computational complexity side, their algorithms require running a number of betting schemes in parallel, which results in running the betting algorithm $O(\frac{1}{\text{precision}})$ times where each run of the betting algorithm takes $\Theta(t)$ complexity. It is worth noting that the idea of designing heuristics gambling schemes goes back at least to [67] who used neural networks, and it is present in the online learning literature too, see [68] for a recent review.

Notably, [15] explicitly motivate the heuristic betting with the idea that regret guarantees “are loose in the constants and not competitive in practice compared to our direct approach” and “the resulting concentration inequalities [from regret-based approaches] are not tight in practice” [15, Appendix D]. Here, we argue the opposite: The use of a non-regret-based betting algorithm is a step back. Indeed, there is no need to use heuristic schemes in the hope of having a low computational complexity because it is straightforward to construct ideal betting algorithms based on portfolio strategies and evaluate tight lower bounds of their wealth using a regret analysis. Moreover, the existence of a regret guarantee allows us to construct confidence sequences with $O(1)$ complexity per time step.

C. Principles for Constructing Betting Algorithms

With the notable exception of [14], none of the described works above propose a unifying general way to construct betting strategies. In fact, even recently [51, page 424] writes “How should the statistician choose the strategy for Skeptic? An obvious goal is to obtain small warranty sets. But a strategy that produces the smallest warranty set for one $N$ and one warranty level $1/\alpha$ will not generally do so for other values of these parameters. So any choice will be a balancing act. How to perform this balancing act is an important topic for further research.”

Actually, in 2019 two papers independently proposed a similar idea as a guiding principle for betting strategies for testing. Reference [14] proposed to use betting algorithms designed to minimize regret with respect to the best constant betting strategy in hindsight for the specific application of confidence sequences. Reference [3] instead proposed to construct betting strategies for testing composite, but parametric, null hypotheses by maximizing the expected log wealth. The connection between the two approaches is rooted in the fact that the expected log wealth is maximized by a constant betting strategy [69]. However, the approach in [3] do not seem to be immediately extendable to the non-parametric setting. Moreover, our work explicitly addresses the “balancing act” mentioned by [51], obtaining tight confidence sequences with one sample and at infinity.

D. Method of Mixtures in Concentration and Betting Proofs

A powerful approach to prove time-uniform concentration inequality is through the use of the so-called method of mixtures or pseudo-maximization [32], [70]. In the simplest case, the key idea is to use a sequence of zero-mean random variables $Y_i \in \mathbb{R}$ with appropriate conditions to construct a supermartingale as

$$\int \exp \left[ \frac{t}{\lambda} \left( \lambda Y_i - \frac{\lambda^2 Y_i^2}{2} \right) \right] dF(\lambda)$$

for a mixture distribution $F$. Then, use Ville’s inequality to infer that $\sum_{t=1}^{\infty} Y_i$ is controlled by high probability. However, the very same idea is also the core one for universal compression [39] and betting algorithms [19], [55], where the betting algorithm is constructed as a mixture of betting schemes that bet a fixed fraction of the current wealth. The connection is far from being accidental: [12], [71] describe how the concept of martingale and the wealth of a bettor can be completely unified. More in detail, for bounded random variables the expression (21) is nothing else than a second-order Taylor approximation of the growth rate of the wealth for a $F$-weighted portfolio algorithm. However, we also introduce for the first time an integration interval that depends on $m$, thanks to Lemma 1.

Since the preliminary version of our paper has been available, there have been a few papers that use the same portfolio/betting based on the $F$-weighted mixture we have proposed in (4) and (6) to derive confidence sequences. Reference [23] leverage the same $F$-weighted mixture and develop PAC-Bayes style concentration inequalities. Reference [72] consider the same mixture but evaluate the log wealth

---

8The arXiv version of this paper appeared online on 27th October 2021.
achieved by a universal portfolio algorithm directly with dynamic programming and obtain a constant per-time-step time complexity. Reference [73] study the near-optimality of the betting-based confidence sequences and focus on a special case of our F-weighted mixture equipped with the uniform prior (Definition 2.3 therein). Note that the uniform prior results in a regret bound with the factor of 1 for the $\ln(n)$ term rather than $\frac{1}{2}$ of ours that uses the Dirichlet($1/2, 1/2$) prior, leading to a looser confidence sequence.

VIII. DISCUSSION

We have presented our new time-uniform concentrations and confidence sequences through a straightforward reduction from betting on a continuous coin to portfolio selection with two stocks. While the portfolio algorithm plays a key role in the construction, one does not need to run a portfolio algorithm explicitly — showing the existence of an online algorithm with a regret bound is enough. Numerically inverting our concentration of measure results, gives non-vacuous confidence sequences that are tight with one sample and at infinity. We conclude our paper by discussing future work.

A. Yet Another Mixture

We have considered the Dirichlet($1/2, 1/2$) and the mixture in [32]. However, other choices are still possible! In fact, we could choose the mixture as a function of $\mu$. This might seem to be an odd choice but it is perfectly legal. Indeed, Theorem 2 uses the knowledge of $\mu$ in deciding the interval of the betting fraction.

B. Running the F-Weighted Universal Portfolio

One might wonder why we need to lower bound the wealth of a universal portfolio algorithm, through an upper bound of its regret, instead of just running it. Indeed, the F-weighted universal portfolio with two stocks with Dirichlet($1/2, 1/2$) and uniform mixture can be implemented with an algorithm whose complexity per time step is linear in the number of the received samples (i.e., cumulative time complexity up to time step $t$ is $O(t^2)$) [27]. However, the same approach cannot be applied to the mixture one in Theorem 8 since the prior does not allow an efficient evaluation of the integral.²

ACKNOWLEDGMENT

The authors thank Wouter Koolen for his blog post on how to prove the LIL bound that inspired this line of work in 2017; Peter Grünwald and Glenn Shafer for suggestions and discussion on prior work; and Blair Bilodeau, Erik Learned-Miller, Erik Ordentlich, Dávid Pál, Aaditya Ramdas, Vladimir Vovk, and Ian Waudby-Smith for comments and feedback on a previous version of this manuscript.

²Since the preliminary version of our paper has been available, [72] have proposed a dynamic programming approach to directly evaluate the wealth achieved by the F-weighted universal portfolio.

REFERENCES

[1] D. A. Darling and H. Robbins, “Iterated logarithm inequalities,” Proc. Nat. Acad. Sci. USA, vol. 57, no. 5, pp. 1188–1192, May 1967.
[2] S. R. Howard, A. Ramdas, J. McAuliffe, and J. Sekhon, “Time-uniform, nonparametric, nonasymptotic confidence sequences,” Ann. Statist., vol. 49, no. 2, pp. 1055–1080, Apr. 2021.
[3] P. Grünwald, R. de Heide, and W. Koolen, “Safe testing,” 2019, arXiv:1906.07801.
[4] R. Kohavi, D. Tang, and Y. Xu, Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing. Cambridge, U.K.: Cambridge Univ. Press, 2020.
[5] D. Precup, “Eligibility traces for off-policy policy evaluation,” in Proc. Int. Conf. Mach. Learn. (ICML), 2000, pp. 759–766.
[6] L. Li, W. Chu, J. Langford, and X. Wang, “Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms,” in Proc. 4th ACM Int. Conf. Web Search Data Mining, Feb. 2011, pp. 297–306.
[7] N. Karampatziakis, P. Mineiro, and A. Ramdas, “Off-policy confidence sequences,” in Proc. Int. Conf. Mach. Learn., 2021, pp. 5301–5310.
[8] J.-Y. Audibert, R. Munos, and C. Szepesvári, “Tuning bandit algorithms in stochastic environments,” in Proc. Int. Conf. Algorithmic Learn. Theory (ALT), 2007, pp. 150–165.
[9] A. Maurer and M. Pontil, “Empirical Bernstein bounds and sample variance penalization,” in Proc. Conf. Learn. Theory (COLT), 2009.
[10] W. Hoeffding, “Probability inequalities for sums of bounded random variables,” J. Amer. Stat. Assoc., vol. 58, no. 301, pp. 13–30, Mar. 1963.
[11] R. Nowak and E. Tánczos, “Tighter confidence intervals for rating systems,” 2019, arXiv:1912.03528.
[12] G. Shafer and V. Vovk, Probability and Finance: It’s Only a Game! Hoboken, NJ, USA: Wiley, 2001.
[13] H. Hendriks, “Test Mattingales for bounded random variables,” 2021, arXiv:2109.08923.
[14] K.-S. Jun and F. Orabona, “Parameter-free online convex optimization with sub-exponential noise,” in Proc. Conf. Learn. Theory, 2019, pp. 1802–1823.
[15] I. Waudby-Smith and A. Ramdas, “Estimating means of bounded random variables by betting,” 2020, arXiv:2010.09686.
[16] N. Cesa-Bianchi and G. Lugosi, Prediction, Learning, and Games. Cambridge, U.K.: Cambridge Univ. Press, 2006.
[17] F. Orabona, “A modern introduction to online learning,” 2019, arXiv:1912.13213.
[18] F. Orabona and D. Pál, “Coin betting and parameter-free online learning,” in Advances in Neural Information Processing Systems, vol. 29, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2016, pp. 577–585.
[19] T. M. Cover, “Universal portfolios,” Math. Finance, vol. 1, no. 1, pp. 1–29, Jan. 1991.
[20] J. Ville, Étude Critique de la Notion de Collectif. Paris, France: Gauthier-Villars, 1939.
[21] V. G. Vovk, “Aggregating strategies,” in Proc. Conf. Comput. Learn. Theory, 1990, pp. 371–386.
[22] A. Garivier and O. Cappe, “The KL-UCB algorithm for bounded stochastic bandits and beyond,” in Proc. Conf. Learn. Theory (COLT), 2011, pp. 359–376.
[23] K. Jang, K. Jun, I. Kuzborskij, and F. Orabona, “Tighter PAC-Bayes bounds through coin-betting,” in Proc. 36th Annua. Conf. Learn. Theory (COLT), vol. 195, Bengaluru, India, Jul. 2023, pp. 2240–2264.
[24] Z. Mhammedi, P. Grünwald, and B. Guedj, “PAC-Bayes un-expected Bernstein inequality,” in Proc. Adv. Neural Inf. Process. Syst., vol. 32, 2019, pp. 1–12.
[25] J.-Y. Audibert, R. Munos, and C. Szepesvári, “Exploration–exploitation tradeoff using variance estimates in multi-armed bandits,” Theor. Comput. Sci., vol. 410, no. 19, pp. 1876–1902, Apr. 2009.
[26] A. Cutkosky and F. Orabona, “Black-box reductions for parameter-free online learning in Banach spaces,” in Proc. Conf. Learn. Theory, 2018, pp. 1493–1529.
[27] T. M. Cover and E. Ordentlich, “Universal portfolios with side information,” IEEE Trans. Inf. Theory, vol. 42, no. 2, pp. 348–363, Mar. 1996.
[28] N. Meinshausen, L. Meier, and P. Bühlmann, “P-values for high-dimensional regression,” J. Amer. Stat. Assoc., vol. 104, no. 488, pp. 1671–1681, Dec. 2009.
[29] C. R. Blyth and H. A. Still, “Binomial confidence intervals,” J. Amer. Statist. Assoc., vol. 78, no. 381, pp. 108–116, 1983.
[30] T. M. Cover and J. A. Thomas, Elements of Information Theory. Hoboken, NJ, USA: Wiley, 2006.
