SAOFTRL: A Novel Adaptive Algorithmic Framework for Enhancing Online Portfolio Selection

Runhao Shi and Daniel P. Palomar, *Fellow, IEEE*

***Abstract*—Strongly Adaptive meta-algorithms (SA-meta) are popular in online portfolio selection due to their resilience in adversarial environments and adaptability to market changes. However, their application is often limited by high variance in errors, stemming from calculations over small intervals with limited observations. To address this limitation, we introduce the Strongly Adaptive Optimistic Follow-the-Regularized-Leader (SAOFTRL), an advanced framework that integrates the Optimistic Follow-the-Regularized-Leader (OFTRL) strategy into SA-meta algorithms to stabilize performance. SAOFTRL is distinguished by its novel regret bound, which provides a theoretical guarantee of worst-case performance in challenging scenarios. Additionally, we reimagine SAOFTRL within a mean-variance portfolio (MVP) framework, enhanced with shrinkage estimators and adaptive rolling windows, thereby ensuring reliable average-case performance. For practical deployment, we present an efficient SAOFTRL implementation utilizing the Successive Convex Approximation (SCA) method. Empirical evaluations demonstrate SAOFTRL's superior performance and expedited convergence when compared to existing benchmarks, confirming its effectiveness and efficiency in dynamic market conditions.**

***Index Terms*—Optimistic follow-the-regularized-leader, strongly adaptive meta-algorithm, mean-variance portfolio, universal portfolio.**

I. Introduction

SINCE Kelly's investment theory's inception [1], online portfolio selection has emerged as a pivotal area in financial engineering. Advances in online learning techniques have enhanced strategies to rival the best constant rebalanced portfolio (BCRP) in hindsight under an adversarial environment[1], without relying on assumptions about data distribution [2], [3].

The key performance metric in online portfolio selection is the regret, which quantifies the cost incurred by an online strategy in determining the optimal portfolio in real time. It reflects the worst-case performance of the online portfolio strategy under an adversarial environment. This concept was first benchmarked by Cover's universal portfolio, achieving a regret of $O(\log T)$ over T trading periods [2], [4], [5]. However, this model faces scalability challenges with large high-dimensional datasets [6], [7], [8]. While subsequent strategies have been more efficient, they often compromise on regret bounds [9]. Addressing this, Agarwal et al. [10] combined Cover's principles with online convex programming [11], leading to significant advancements in optimal and efficient portfolio construction. Hazan et al. [12], [13] then introduced the Online Newton Step (ONS) employing the property of exp-concave functions to achieve $O(N \log T)$ regret on N assets with manageable complexity. Building on this, they further employed the Follow-the-Regularized-Leader (FTRL) algorithm, achieving logarithmic regret in terms of the variance of gradients that surpass ONS in regret bounds [14]. To leverage the advantages of correctly predicting future signals, an advanced variant of FTRL called Optimistic FTRL (OFTRL) has emerged [15], [16]. OFTRL incorporates prediction signals into the decision-making process, offering enhanced performance with accurate predictions and maintaining robustness against incorrect ones.

Given the nonstationarity of the stock market, Hazan et al. proposed an adaptive meta-algorithm with the concept of adaptive regret [17]. This metric better reflects the market's dynamic nature compared to static regret [18], [19], [20], [21], [22]. Daniely et al. [23] generalized the concept of adaptive regret and proposed the Strongly Adaptive meta-algorithm (SA-meta), enabling the transformation of low static regret algorithms into strongly adaptive ones. This method involves calculating multiple portfolio instances over selected intervals and aggregating them based on past performance. Jun et al. [24] further improved the SA-meta's regret bounds using coin betting techniques, achieving optimal theoretical performance.

In the dynamic landscape of online portfolio selection, SA-meta faces challenges stemming from the significant variance in short-term calculations. This variance often leads to inconsistent performance in real-world markets. We observe that OFTRL effectively mitigates variance error resulting from limited observations by introducing a deliberate bias.

Received 22 March 2024; revised 4 July 2024 and 1 October 2024; accepted 1 November 2024. Date of publication 11 November 2024; date of current version 25 November 2024. This work was supported by Hong Kong GRF 16206123 Research Grant. The associate editor coordinating the review of this article and approving it for publication was Dr. Mingyi Hong. *(Corresponding author: Runhao Shi.)*

The authors are with Hong Kong University of Science and Technology (HKUST), Clear Water Bay, Kowloon, Hong Kong (e-mail: rshiaf@connect.ust.hk; palomar@ust.hk).

Digital Object Identifier 10.1109/TSP.2024.3495696

[1]An adversarial environment is designed or assumed to be as unfavorable as possible for the algorithm.

Expanding on this notion, we have integrated OFTRL into SA-meta, giving rise to the Strongly Adaptive Optimistic Follow-the-Regularized-Leader (SAOFTRL) approach. SAOFTRL is equipped with a unified regret bound closely related to the regret of OFTRL, benefiting from improved performance for accurate estimations. This integration serves to stabilize performance during short-term periods by striking a balance between variance and bias.

Recent advances in regret analysis for OFTRL focus on the case where lag-1 signals are used for prediction [25], [26]. However, the inherent volatility of the stock market often makes such predictions suboptimal. Therefore, we propose a robust variant of OFTRL that uses the average of historical signals for prediction. We further derive a novel regret bound for OFTRL. Our analysis reveals that accurate predictions substantially enhance performance, while incorrect predictions have a negligible effect on the algorithm's worst-case performance guarantee. Moreover, we improve the computational efficiency of OFTRL by utilizing the Successive Convex Approximation (SCA) method [27], without compromising regret bound. Specifically, we have optimized the time complexity from a previous standard of $O(NT + \log(1/\epsilon)N^{3.5})$ to a more efficient $O\left(\tau_{\text{iter}}\left(NT + N\log N\right)\right)$, where ϵ and τ_{iter} are parameters linked to the convergence gap of the algorithm.

While the regret bound for SAOFTRL ensures robustness in adversarial environments, it may not fully encompass the multifaceted complexities of the market, which extend beyond adversarial conditions. To address this limitation, we establish a connection between SAOFTRL and the mean-variance portfolio (MVP) model [28], known for its focus on average-case market scenarios. Regret in SAOFTRL measures worst-case performance, which may be overly conservative for real markets, whereas MVP estimates expected profit based on commonly employed model assumptions. This provides a more precise representation of average performance expectations. Our research shows that SAOFTRL can be equivalently viewed as an advanced version of the MVP model, incorporating shrinkage estimators and adaptive lookback windows, as illustrated in Fig. 1. This dual analysis enables SAOFTRL to align with MVP's average-case expectations while preserving the worst-case performance guarantees of online learning techniques.

To validate our methods, we conduct extensive simulations using real market datasets. The results show significant improvements in computational efficiency and effectiveness. By comparing the Sharpe ratio and maximum drawdown, we demonstrate the outstanding average-case performance of our methods. Additionally, comparing the regret of SAOFTRL with benchmarks highlighted its robustness in maintaining worst-case performance guarantees.

In summary, the main contributions of this paper are:

- We develop a versatile adaptive SAOFTRL framework that integrates the benefits of SA-meta and OFTRL for portfolio selection in dynamic environments.
- We introduce a novel regret bound for OFTRL utilizing average historical signals as the prediction, providing a theoretical worst-case performance guarantee under adversarial conditions. Our analysis shows that this regret can be decreased for accurate predictions while maintaining robustness for incorrect estimations.
- We propose an efficient algorithm for SAOFTRL using SCA, which significantly reduces time complexity while maintaining a consistent regret bound.
- We establish the connection between SAOFTRL and MVP, which enables a more precise representation of average performance expectations in real market scenarios.

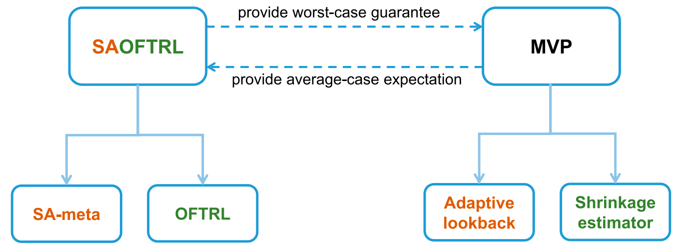


Fig. 1. Dual analysis of SAOFTRL and MVP: In SAOFTRL, SA-meta functions as an adaptive lookback mechanism in MVP, while OFTRL operates as MVP with a shrinkage estimator.

The structure of this paper is organized as follows. In Section II, we introduce the fundamental concepts that form the basis of this study. Section III presents the SAOFTRL framework along with a novel regret bound for the underlying OFTRL strategy. In Section IV, we present an efficient algorithm for OFTRL without compromising regret bound. Section V provides a new analytical perspective for SAOFTRL by bridging the connection to MVP. Extensive experiments to showcase the performance and efficiency of our approach are provided in Section VI. Finally, we conclude this paper and explore potential future directions in Section VII.

II. Preliminaries

In this section, we introduce the fundamental concepts underpinning our research, laying the groundwork for subsequent discussions.

A. Follow-the-Regularized-Leader

Online portfolio selection involves investors making sequential investment decisions to approximate the performance of an ideal strategy, known as the Best Constant Rebalanced Portfolio (BCRP), determined in hindsight. BCRP is a static portfolio determined by maximizing the log cumulative return using future prices:

$$\mathbf{b}^\star = \arg\min_{\mathbf{b}\in\mathcal{W}} \sum_{t=1}^{T} f_t(\mathbf{b}).$$

Here, $\mathbf{b}$ is the portfolio vector, $\mathcal{W}$ is the feasible set, and $f_t(\mathbf{b}) = -\log(1 + \mathbf{r}_t^\top \mathbf{b})$ represents the negative log return, where $\mathbf{r}_t$[2] is the asset return vector at time t. Online portfolio selection strategies sequentially select the portfolio $\mathbf{b}_{t+1}$ based on historical information $\{f_1, \ldots, f_t\}$ to match the cumulative

[2]The return $r_{t,i}$ of an asset i at time t is defined as $r_{t,i} = \frac{p_{t,i} - p_{t-1,i}}{p_{t-1,i}}$, where $p_{t,i}$ denotes the price of asset i at time t.

return of the BCRP. The effectiveness of such a strategy is quantified using the concept of regret, defined as:

$$R_{[T]} = \sum_{t=1}^{T} f_t(\mathbf{b}_t) - \sum_{t=1}^{T} f_t(\mathbf{b}^\star).$$

The regret $R_{[T]}$ quantifies the discrepancy between the log cumulative return of the selected portfolio and the BCRP. A strategy attains universality if $R_{[T]} = o(T)$, suggesting that the average performance gap reduces to zero as T approaches infinity.

The Follow-the-Leader (FTL) strategy, which sets $\mathbf{b}_t$ as the action minimizing past cumulative losses (the Leader), often fails to achieve optimal regret due to instability [29]. This leads to the adoption of the Follow-the-Regularized-Leader (FTRL) strategy. FTRL involves minimizing the cumulative past loss plus a regularization term (the Regularized Leader) to enhance stability, as shown below:

$$\mathbf{b}_{t+1} = \arg\min_{\mathbf{b}\in\mathcal{W}} \sum_{\tau=1}^{t} f_\tau(\mathbf{b}) + \frac{\beta}{2}\left\|\mathbf{b}\right\|_2^2, \tag{1}$$

where $\beta > 1$ is the weight of the regularization term. The FTRL strategy achieves one of the tightest known regret bounds in the realm of online portfolio selection, which is $O(N \log Q_{[T]})$. Here, $Q_{[T]}$ represents the variance of the gradient series $\{\nabla f_t(\mathbf{b}_t)\}$, a critical measure in this context [14]. Compared to the well-known Online Newton Step (ONS) that attains regret of $O(N \log T)$ [13], FTRL has priority in retaining the worst-case robustness since $Q_{[T]} \leq T$ after appropriate normalization.

B. Optimistic Follow-the-Regularized-Leader

The Optimistic Follow-the-Regularized-Leader (OFTRL) strategy is an advanced variant of the FTRL strategy [15], [16]. OFTRL distinguishes itself by integrating predictions of future losses into the decision-making process. The portfolio for the subsequent period is determined as follows:

$$\mathbf{b}_{t+1} = \arg\min_{\mathbf{b}\in\mathcal{W}} \sum_{\tau=1}^{t} f_\tau(\mathbf{b}) + \tilde{f}_{t+1}(\mathbf{b}) + \frac{\beta}{2}\left\|\mathbf{b}\right\|_2^2, \tag{2}$$

where $\tilde{f}_{t+1}(\mathbf{b})$ denotes the predicted loss function. Since the regret bound for OFTRL is directly linked to the gradient of $\tilde{f}_{t+1}(\mathbf{b}_t)$ [30, Theorem 7.35], we only need to consider the first-order information about the predicted loss function. We can represent $\tilde{f}_{t+1}(\mathbf{b})$ as:

$$\tilde{f}_{t+1}(\mathbf{b}) = \langle \mathbf{m}_{t+1}, \mathbf{b}\rangle, \tag{3}$$

where $\mathbf{m}_{t+1} := \nabla \tilde{f}_{t+1}(\mathbf{b})$ represents the gradient for the predicted loss function.

C. Strongly Adaptive Algorithm

The dynamic nature of trading environments often results in varied optimal strategies over time. As such, standard regret measures, which compare performance with the best static action, are less effective in identifying a superior portfolio. Addressing this, [17] introduced adaptive regret, later expanded into strongly adaptive regret by [23], to evaluate the tracking accuracy to the optimal portfolio over any contiguous time interval I[3]. Strongly adaptive regret is defined as:

$$\text{SA-R}_T^{(\tau)} = \max_{I\subseteq[T]:|I|=\tau} R_I, \tag{4}$$

where R_I is the static regret over interval $I = [q, s]$:

$$R_I = \sum_{t\in I} f_t(\mathbf{b}_t) - \min_{\mathbf{b}\in\mathcal{W}} \sum_{t\in I} f_t(\mathbf{b}).$$

$\text{SA-R}_T^{(\tau)}$ measures tracking accuracy against the BCRP for any interval of length τ, which encompasses the static regret when $\tau = T$. A Strongly Adaptive meta-algorithm (SA-meta) can transform a low-regret algorithm into a strongly adaptive one. Specifically, if an algorithm has a static regret of $R_I = o(\tau)$ over an interval I of length τ, the meta-algorithm is considered strongly adaptive if $\text{SA-R}_T^{(\tau)} = O(\text{poly}(\log T)\cdot R_I)$. Practically, for a market transitioning from a bear to a bull phase, strongly adaptive algorithms would adeptly align with the optimal portfolio in each phase.

Various SA-meta, as demonstrated in [17], [23], [24], [31], [32], can transform standard low-regret algorithms into strongly adaptive ones. In this paper, we adopt the Coin Betting for Changing Environment (CBCE) method from [24], known for its excellent strongly adaptive regret performance. For a given low regret online learning algorithm $\mathcal{A}$ with regret $R_I^{\mathcal{A}}$, CBCE [24, Lemma 2] can achieve the following strongly adaptive regret:

$$\text{SA-R}_T^{(\tau)} = O\left(\sqrt{\tau \log T} + \log\tau \cdot R_I^{\mathcal{A}}\right), \tag{5}$$

where $I = [q, s] \subseteq [T]$ denotes the interval with length $|I| = \tau$.

III. Proposed Method

In this section, we present the Strongly Adaptive Optimistic Follow-the-Regularized-Leader (SAOFTRL) framework. Additionally, we propose a novel regret bound specifically designed for the underlying OFTRL strategy in the context of online portfolio selection, where the average of past signals is used for prediction. This bound is intrinsically connected to the regret bound for FTRL, demonstrating superior performance for accurate predictions and negligible cost for inaccurate predictions in OFTRL.

A. Strongly Adaptive Optimistic Follow-the-Regularized-Leader

To address the dynamic and complex nature of financial markets, we introduce an advanced integration of SA-meta and OFTRL into the domain of portfolio selection. Our approach, named as the Strongly Adaptive Optimistic Follow-the-Regularized-Leader (SAOFTRL) algorithm, is outlined in Algorithm 1.

[3]Compared to adaptive regret, strongly adaptive regret places greater emphasis on the dependence on the interval length $|I|$.

Algorithm 1 Strongly Adaptive Optimistic Follow-the-Regularized-Leader (SAOFTRL) for Portfolio Selection

1: **Parameter:** $\beta > 1$
2: Initialize $\mathbf{b}_1 = \frac{1}{N}\mathbf{1} \in \mathbb{R}^N$, $p_1^{([1,1])} = 1$
3: **for** $t = 1, \ldots, T$ **do**
4: For each $I \in \text{Active}(t)$, compute instantaneous regret

$$\ell_t^{(I)} = f_t(\mathbf{b}_t) - f_t\left(\mathbf{b}_t^{(I)}\right) \tag{6}$$

5: Update weight $p_{t+1}^{(I)}$ based on $\ell_t^{(I)}$ according to (9)-(11)
6: For each $I = [q, s] \in \text{Active}(t+1)$, compute

$$\mathbf{b}_{t+1}^{(I)} = \arg\min_{\mathbf{b}\in\mathcal{W}} \sum_{\tau=q}^{t} f_\tau(\mathbf{b}) + \langle \mathbf{m}_{t+1}^{(I)}, \mathbf{b}\rangle + \frac{\beta}{2}\|\mathbf{b}\|_2^2 \tag{7}$$

7: Compute the combined portfolio

$$\mathbf{b}_{t+1} = \sum_{I\in\text{Active}(t+1)} p_{t+1}^{(I)} \cdot \mathbf{b}_{t+1}^{(I)} \tag{8}$$

8: **end for**
9: **Output:** $\{\mathbf{b}_t\}$

In essence, SAOFTRL computes the portfolio instances $\mathbf{b}_{t+1}^{(I)}$ for each interval I within a set of structured intervals, termed $\text{Active}(t+1)$, at any given time $t+1$. The weight assignment for these portfolios is determined based on their historical performance, favoring portfolios with better performance. The portfolio is finally aggregated by combining the portfolio instances according to their weights.

a) Active Intervals: The active intervals $\text{Active}(t)$ at any time t are derived from a collection of geometric covering (GC) intervals $\mathcal{I}$, as outlined in [23]. Specifically, GC intervals are defined as:

$$\mathcal{I} := \bigcup_{k\in 1,2,\ldots} \mathcal{I}_k,$$

with $\mathcal{I}_k$ consisting of intervals $[i\cdot 2^k, (i+1)\cdot 2^k - 1]$ for $i \in \mathbb{N}$. This configuration of $\mathcal{I}_k$ segments $\mathbb{N}\backslash\{1, \ldots, 2^k\}$ into sequential intervals of equal length 2^k, as detailed in Table I. The set $\text{Active}(t)$ comprises GC intervals that include t:

$$\text{Active}(t) := \{I \in \mathcal{I} : t \in I\}.$$

A notable feature of this method is its computational efficiency. At any time t, the number of active intervals is $\lfloor \log(t) \rfloor + 1$, indicating only $O(\log(t))$ calculations of OFTRL per round. This logarithmic scalability is crucial for the efficiency of the algorithm in rapidly changing stock markets.

b) Weighting Method: The crucial difference among these SA-meta algorithms is their interval weight update mechanisms, as specified in Line 5 of Algorithm 1. In this paper, we adopt the CBCE method. CBCE involves scaling the instantaneous regret $\ell_t^{(I)}$ to the $[-1, 1]$ range and updating weights accordingly:

$$\tilde{\ell}_t^{(I)} = \begin{cases} \ell_t^{(I)}/\tilde{\ell}_{\max} & \hat{p}_t^{(I)} > 0 \\ \left[\ell_t^{(I)}/\tilde{\ell}_{\max}\right]_+ & \text{otherwise,} \end{cases} \tag{9}$$

TABLE I
GEOMETRIC COVERING (GC) INTERVALS $\mathcal{I}$ [23], WHERE EACH INTERVAL IS DENOTED BY [ ]

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$\mathcal{I}_0$	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	...
$\mathcal{I}_1$		[	]	[	]	[	]	[	]	[	]	[	]	[	]	[	]	[	...
$\mathcal{I}_2$				[			]	[			]	[			]	[			...
$\mathcal{I}_3$								[							]	[			...
$\mathcal{I}_4$																[			...

where $[\cdot]_+$ denotes $\max\{\cdot, 0\}$, $\tilde{\ell}_{\max}$ is a scaling number, and $\hat{p}_t^{(I)}$ is the adjusted weight determined by the previous instantaneous regret. The adjusted weight for the next round $\hat{p}_{t+1}^{(I)}$ for each $I = [q, s] \in \text{Active}(t+1)$ is then calculated by

$$\hat{p}_{t+1}^{(I)} = \frac{\sum_{\tau=q}^{t} \tilde{\ell}_\tau^{(I)}}{(t-q+2)}\left(1 + \sum_{\tau=q}^{t} \hat{p}_\tau^{(I)} \tilde{\ell}_\tau^{(I)}\right). \tag{10}$$

The interval weight for each $I = [q, s] \in \text{Active}(t+1)$ is then calculated using a normalization process with a predefined scaling factor $\pi^{(I)} = \frac{1}{q^2\lfloor 1+\log_2 q\rfloor}$:

$$p_{t+1}^{(I)} = \begin{cases} C_1^{-1}\pi^{(I)}\left[\hat{p}_{t+1}^{(I)}\right]_+ & \text{if } C_1 > 0 \\ C_2^{-1}\pi^{(I)} & \text{otherwise,} \end{cases} \tag{11}$$

where C_1 and C_2 are normalization factors. This method ensures the theoretical performance for SAOFTRL by reflecting the interval performance and maintaining a balanced distribution among all active intervals.

c) Regret Analysis: The worst-case performance analysis for SAOFTRL leverages Equation (5) with the underlying low regret algorithm $\mathcal{A}$ as OFTRL. Consequently, the strongly adaptive regret for SAOFTRL is primarily determined by the regret bound of the underlying OFTRL strategy. In the following section, we will present an in-depth analysis of the regret bound for OFTRL within the context of online portfolio selection. Furthermore, we will introduce a novel regret bound for OFTRL that depends on the variance of the gradients. This bound underscores the robust performance of OFTRL, even when the prediction of the next gradient is inaccurate.

B. OFTRL With Logarithmic Variance Bound

To further analyze the regret bound of OFTRL, we introduce the following assumptions:

Assumption 1: The gradients of any loss function f_t are bounded, with $\sup_{\mathbf{b}\in\mathcal{W}} \|\nabla f_t(\mathbf{b})\|_2 \le G$ for all $t \in [T]$.

Assumption 2: The domain $\mathcal{W}$ has a bounded diameter, meaning $\|\mathbf{b} - \mathbf{u}\|_2 \le B$ for all $\mathbf{b}, \mathbf{u} \in \mathcal{W}$.

Suppose the feasible set $\mathcal{W}$ is a simplex, the diameter can be conveniently set as $B = \sqrt{2}$ in Assumption 2. Furthermore, Assumption 1 can be easily satisfied by ensuring that the stock price does not fall to half its original value. We denote $\mathbf{g}_t := \nabla f_t(\mathbf{b}_t)$ as the gradient of the loss function at time t. For convenience, instead of considering the regret on interval I, we calculate the regret on $[T]$. Under Assumptions 1 and 2, we establish the following theorem, which provides the regret bound for OFTRL in online portfolio selection for any prediction term $\mathbf{m}_{t+1}$:

Theorem 1: Under Assumption 1 and 2, the regret of OFTRL for online portfolio selection is bounded by

$$R_{[T]}^{\mathrm{OFTRL}} \leq \frac{1}{2}\sum_{t=1}^{T}\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2 + \frac{\beta}{2}B^2, \tag{12}$$

where $\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2 = (\mathbf{g}_t - \mathbf{m}_t)^\top \mathbf{A}_t^{-1}(\mathbf{g}_t - \mathbf{m}_t)$, $\mathbf{A}_t = \beta\mathbf{I} + \alpha\sum_{\tau=1}^{t}\mathbf{g}_\tau\mathbf{g}_\tau^\top$, and $\alpha = \frac{1}{8GB}$.

The proof, based on [30, Lemma 7.1], can be found in Appendix A. Intriguingly, the FTRL method for portfolio selection [14], which becomes identical to OFTRL when $\tilde{f}_{t+1}(\mathbf{b}) = 0$, exhibits a similar regret bound:

$$R_{[T]}^{\mathrm{FTRL}} \leq \frac{1}{2}\sum_{t=1}^{T}\|\mathbf{g}_t\|_{\mathbf{A}_t^{-1}}^2 + \frac{\beta}{2}B^2 = O\left(N\log Q_{[T]}\right). \tag{13}$$

Here $Q_{[T]} = \sum_{t=1}^{T}\|\mathbf{g}_t - \bar{\mathbf{g}}\|_2^2$ represents the variance of the gradients, where $\bar{\mathbf{g}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_t$. By comparing the regret bounds of OFTRL and FTRL, we observe that OFTRL, with an accurate prediction of the gradient, can significantly minimize the term $\sum_{t=1}^{T}\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2$, thus achieving better performance compared to FTRL. Consequently, when the predictions are perfect, the regret is a constant value (no regret).

Recent advancements in OFTRL have primarily relied on a lag-1 gradient as a predictive signal in subsequent iterations with a corresponding regret bound to guarantee the worst-case performance theoretically [25], [26]. However, the inherent volatility of the stock market often makes relying solely on the most recent signal for gradient prediction suboptimal in general. Therefore, we would like to improve the prediction in OFTRL for the stock market and construct a regret bound that reflects the worst-case performance of utilizing such a prediction.

A robust alternative for the prediction signal, favored by investors, is to use the historical average of data points:

$$\mathbf{m}_{t+1} = \frac{1}{t}\sum_{s=1}^{t}\mathbf{g}_s.$$

We further establish a regret bound for OFTRL with $\mathbf{m}_{t+1}$ as the historical average signal, which is related to the variance of gradients. The details of this regret bound are provided in Theorem 2.

Theorem 2: Utilizing $\mathbf{m}_{t+1} = \frac{1}{t}\sum_{s=1}^{t}\mathbf{g}_s$ in OFTRL results in a regret bound over $[T]$:

$$R_{[T]}^{\mathrm{OFTRL}} = O\left(N\log\left(Q_{[T]} + \log T\right)\right), \tag{14}$$

where N is the number of assets, and $Q_{[T]}$ is the variance of gradients defined as $\sum_{t=1}^{T}\|\mathbf{g}_t - \bar{\mathbf{g}}\|_2^2$ with $\bar{\mathbf{g}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_t$.

This theorem facilitates a direct calculation of the regret bound for OFTRL over any specific interval $I = [q, s]$ with $\mathbf{m}_{t+1}^{(I)} = \frac{1}{t-q+1}\sum_{\tau=q}^{t}\mathbf{g}_\tau^{(I)}$:

$$R_I^{\mathrm{OFTRL}} = O\left(N\log\left(Q_I + \log|I|\right)\right), \tag{15}$$

where Q_I represents the variance of the gradients over interval I. The foundational proof of Theorem 2 builds on the regret bound established in Theorem 1. To effectively bound $\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2$, insights from Lemma 1 is leveraged, with detailed proofs provided in Appendices B.

Lemma 1: Given $\mathbf{m}_s = \frac{1}{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_\tau$ with $\mathbf{m}_1 = \mathbf{0} \in \mathbb{R}^N$, it follows that $\sum_{s=1}^{t}(\mathbf{g}_s - \mathbf{m}_s)(\mathbf{g}_s - \mathbf{m}_s)^\top \leq 8\sum_{s=1}^{t}\mathbf{g}_s\mathbf{g}_s^\top$.

Let $\tilde{\mathbf{g}}_s = \mathbf{g}_s - \mathbf{m}_s$ and define $\tilde{\mathbf{A}}_t = \beta\mathbf{I} + \alpha\sum_{s=1}^{t}\tilde{\mathbf{g}}_s\tilde{\mathbf{g}}_s^\top$. In light of Lemma 1, for $\mathbf{m}_s = \frac{1}{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_\tau$, we directly establish the following inequality:

$$\sum_{s=1}^{t}\tilde{\mathbf{g}}_s\tilde{\mathbf{g}}_s^\top \leq 8\sum_{s=1}^{t}\mathbf{g}_s\mathbf{g}_s^\top,$$

implying that $\tilde{\mathbf{A}}_t \preceq 8\mathbf{A}_t$. Setting $\tilde{\mathbf{A}}_0 = \beta\mathbf{I}$, the inequality is reformulated as:

$$\begin{aligned}
\sum_{t=1}^{T}\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2 &= \sum_{t=1}^{T}\tilde{\mathbf{g}}_t^\top\mathbf{A}_t^{-1}\tilde{\mathbf{g}}_t \\
&\leq \frac{8}{\alpha}\sum_{t=1}^{T}\left(\sqrt{\alpha}\tilde{\mathbf{g}}_t\right)^\top\tilde{\mathbf{A}}_t^{-1}\left(\sqrt{\alpha}\tilde{\mathbf{g}}_t\right) \\
&\leq \frac{8}{\alpha}\log\left[\frac{\det\left(\tilde{\mathbf{A}}_T\right)}{\det\left(\tilde{\mathbf{A}}_0\right)}\right] \\
&\leq \frac{8N}{\alpha}\log\left(\beta + \alpha\sum_{t=1}^{T}\|\tilde{\mathbf{g}}_t\|_2^2\right),
\end{aligned}$$

where the third inequality utilizes [33, Lemma 11.11]. Defining $\tilde{Q}_{[T]} = \sum_{t=1}^{T}\|\tilde{\mathbf{g}}_t\|_2^2$, and basedv on Lemma 3 in Appendix C, we deduce:

$$\sum_{t=1}^{T}\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2 = O\left(N\log\left(Q_{[T]} + \log T\right)\right).$$

When comparing this regret bound with that of the FTRL algorithm in Equation (13), the additional overhead incurred due to incorrect predictions is negligible, amounting to $o(\log\log T)$. As a result, OFTRL with a historical average for gradient prediction proves to be a highly adaptable and resilient strategy, especially in the dynamic and ever-changing landscape of financial markets.

Our proposed method not only achieves a problem-dependent regret guarantee, which improves performance when the data is stable and exhibits low variance, but also ensures that the minimax regret rates are safeguarded in worst-case scenarios. For exp-concave functions in online portfolio selection tasks, the optimal minimax regret bound is well-established as $O(N\log T)$, representing the best achievable performance for any algorithm in general cases [13], [34], [35]. Under Assumption 1, the variance term $Q_{[T]}$ can be further bounded as follows:

$$Q_{[T]} = \sum_{t=1}^{T}\|\mathbf{g}_t\|_2^2 + T\|\bar{\mathbf{g}}\|_2^2 - 2\sum_{t=1}^{T}\langle\mathbf{g}_t, \bar{\mathbf{g}}\rangle \leq 4TG^2.$$

Consequently, this results in the worst-case regret bound:

$$R_{[T]}^{\mathrm{OFTRL}} = O(N\log(T + \log T)) = O(N\log T),$$

which aligns with the established minimax regret lower bound. Thus, our OFTRL algorithm meets the theoretical minimax bound in the worst case, while allowing for improved performance in more favorable conditions.

IV. Efficient Algorithm

The SAOFTRL algorithm, as outlined in Algorithm 1, involves solving the optimization problem of OFTRL multiple times at each iteration (see Line 6). Assuming the feasible set is a simplex, the problem can be formulated as follows:

$$\begin{aligned} \underset{\mathbf{b}\in\mathbb{R}^N}{\text{minimize}} \quad & \sum_{\tau=1}^{t} f_\tau(\mathbf{b}) + \langle \mathbf{m}_{t+1}, \mathbf{b}\rangle + \frac{\beta}{2}\|\mathbf{b}\|_2^2 \\ \text{subject to} \quad & \mathbf{b} \geq 0, \quad \mathbf{1}^\mathsf{T}\mathbf{b} = 1. \end{aligned} \tag{16}$$

Although this problem could be solved directly by using the Interior Point Method (IPM) with an off-the-shelf solver [36], it generally incurs a high-order time complexity as $O(NT + \log(1/\epsilon)\, N^{3.5})$, with ϵ as the convergence tolerance.

To enhance efficiency, we introduce an efficient implementation algorithm that incorporates the Successive Convex Approximation (SCA) method to accelerate computation [27], [37], [38]. The SCA method is a strategy for reducing time complexity by iteratively optimizing a more manageable surrogate function for the objective function until convergence. We denote the objective function of problem (16) as $h_t(\mathbf{b})$. To apply the SCA method, the surrogate function denoted as $\tilde{h}_t(\mathbf{b}\,|\,\mathbf{b}^k)$ should be strongly convex and satisfies the property that $\nabla\tilde{h}_t(\mathbf{b}^k\,|\,\mathbf{b}^k) = \nabla h_t(\mathbf{b}^k)$.

Let's start by approximating the logarithmic part of the objective function with a quadratic function. We define the sum of loss functions as $F_t(\mathbf{b}) = \sum_{\tau=1}^{t} f_\tau(\mathbf{b})$. To approximate $F_t(\mathbf{b})$ around $\mathbf{b}^k$, we utilize the following quadratic function:

$$\begin{aligned} F_t(\mathbf{b}) \approx\ & F_t\left(\mathbf{b}^k\right) + \left(\mathbf{b}-\mathbf{b}^k\right)^\mathsf{T} \nabla F_t\left(\mathbf{b}^k\right) \\ & + \frac{1}{2}\left(\mathbf{b}-\mathbf{b}^k\right)^\mathsf{T} \mathbf{L}_1 \left(\mathbf{b}-\mathbf{b}^k\right). \end{aligned}$$

Here, we set $\mathbf{L}_1 = \boldsymbol{\mu}\boldsymbol{\mu}^\mathsf{T}$ with $\boldsymbol{\mu} = \frac{1}{t}\sum_{\tau=1}^{t}\mathbf{r}_\tau$. By incorporating the quadratic approximation of the logarithmic terms into the objective function, we obtain the surrogate function $\tilde{h}_t(\mathbf{b}\,|\,\mathbf{b}^k)$:

$$\tilde{h}_t(\mathbf{b}\,|\,\mathbf{b}^k) = \frac{1}{2}\mathbf{b}^\mathsf{T}(\mathbf{L}_1 + \beta\mathbf{I})\,\mathbf{b} - \mathbf{v}_t^{k,\mathsf{T}}\mathbf{b} + \text{const},$$

where $\mathbf{v}_t^k = \mathbf{L}_1\mathbf{b}^k - \mathbf{m}_{t+1} + \sum_{\tau=1}^{t}\frac{\mathbf{r}_\tau}{1+\mathbf{r}_\tau^\mathsf{T}\mathbf{b}^k}$. The surrogate function $\tilde{h}_t(\mathbf{b}\,|\,\mathbf{b}^k)$ constructed above satisfies the criteria for the surrogate function of SCA. Specifically, it's easy to verify that $\tilde{h}_t(\mathbf{b}\,|\,\mathbf{b}^k)$ is strongly convex and $\nabla\tilde{h}_t(\mathbf{b}^k\,|\,\mathbf{b}^k) = \nabla h_t(\mathbf{b}^k)$. To further simplify the calculation, we can majorize the quadratic term of $\tilde{h}_t(\mathbf{b}\,|\,\mathbf{b}^k)$ using the following lemma.

Lemma 2 [39]: Let $\mathbf{L}$ be a real symmetric matrix and $\mathbf{M}$ another real symmetric matrix such that $\mathbf{M} \succeq \mathbf{L}$. Then for any point $\mathbf{b}^k \in \mathbb{R}^N$ the quadratic function $\mathbf{b}^\mathsf{T}\mathbf{L}\mathbf{b}$ is majorized at $\mathbf{b}^k$ by $\mathbf{b}^\mathsf{T}\mathbf{M}\mathbf{b} + 2\mathbf{b}^\mathsf{T}(\mathbf{L}-\mathbf{M})\mathbf{b}^k + \mathbf{b}^{k,\mathsf{T}}(\mathbf{M}-\mathbf{L})\mathbf{b}^k$.

Based on Lemma 2, if we set $\mathbf{M}_1 = \lambda_{\max}^{(\mathbf{L}_1)}\mathbf{I} = \boldsymbol{\mu}^\mathsf{T}\boldsymbol{\mu}\mathbf{I}$, then $\mathbf{M}_1 \succeq \mathbf{L}_1$ holds. Thus we can and further majorize $\tilde{h}_t\left(\mathbf{b}\,|\,\mathbf{b}^k\right)$ at $\mathbf{b}^k$ by $\hat{h}(\mathbf{b}|\mathbf{b}^k)$, which is given by:

$$\begin{aligned} \hat{h}(\mathbf{b}|\mathbf{b}^k) =\ & -\mathbf{b}^\mathsf{T}\left(\lambda_{\max}^{(\mathbf{L}_1)}\mathbf{b}^k - \mathbf{m}_{t+1} + \sum_{\tau=1}^{t}\frac{\mathbf{r}_\tau}{1+\mathbf{r}_\tau^\mathsf{T}\mathbf{b}^k}\right) \\ & + \frac{\beta + \lambda_{\max}^{(\mathbf{L}_1)}}{2}\mathbf{b}^\mathsf{T}\mathbf{b} + \text{const}. \end{aligned} \tag{17}$$

It's easy to verify that $\hat{h}(\mathbf{b}^k|\mathbf{b}^k) = \nabla h_t(\mathbf{b}^k)$. By simplifying $\hat{h}(\mathbf{b}|\mathbf{b}^k)$, we obtain an equivalent, simplified surrogate problem at the $(k+1)$-th iteration:

$$\begin{aligned} \underset{\mathbf{b}\in\mathbb{R}^N}{\text{minimize}} \quad & \mathbf{b}^\mathsf{T}\mathbf{b} - 2\mathbf{d}_t^{k,\mathsf{T}}\mathbf{b} \\ \text{subject to} \quad & \mathbf{b} \geq 0, \quad \mathbf{1}^\mathsf{T}\mathbf{b} = 1, \end{aligned} \tag{18}$$

where

$$\mathbf{d}_t^k = \frac{1}{\beta + \lambda_{\max}^{(\mathbf{L}_1)}}\left(\lambda_{\max}^{(\mathbf{L}_1)}\mathbf{b}^k - \mathbf{m}_{t+1} + \sum_{\tau=1}^{t}\frac{\mathbf{r}_\tau}{1+\mathbf{r}_\tau^\mathsf{T}\mathbf{b}^k}\right). \tag{19}$$

The optimization problem (18) possesses a closed-form solution, which can be directly computed using a water-filling algorithm [40]. This solution is articulated in the following proposition:

Proposition 3: Consider the optimization problem:

$$\begin{aligned} \underset{\mathbf{b}\in\mathbb{R}^N}{\text{minimize}} \quad & \mathbf{b}^\mathsf{T}\mathbf{b} - 2\mathbf{q}^\mathsf{T}\mathbf{b} \\ \text{subject to} \quad & \mathbf{b} \geq 0, \quad \mathbf{1}^\mathsf{T}\mathbf{b} = 1. \end{aligned} \tag{20}$$

This problem has a closed-form solution given by:

$$b_i^\star = [q_i + \kappa]_+ \quad i = 1, \ldots, N, \tag{21}$$

where $\kappa = \frac{1}{\rho}\left(1 - \sum_{i=1}^{\rho} q_{[i]}\right)$ with:

$$\rho = \max\left\{1 \leq j \leq N : q_{[j]} + \frac{1}{j}\left(1 - \sum_{i=1}^{j} q_{[i]}\right) > 0\right\},$$

and $q_{[i]}$ are the sorted elements of $\mathbf{q}$, arranged such that $q_{[1]} \geq q_{[2]} \geq \cdots \geq q_{[N]}$.

With this closed-form solution, the overall procedure for efficiently optimizing Problem (16) can be effectively encapsulated in Algorithm 2. Unlike previous work that reduces computational complexity at the expense of deteriorating the regret bound [14], our efficient implementation enhances computational efficiency without compromising the regret bound. By setting $\gamma^{k+1} = \gamma^k(1 - \epsilon\gamma^k)$ with $\epsilon \in (0, 1)$ and $\gamma^0 < 1/\epsilon$, Algorithm 2 converges to the optimal point of Problem (16). This convergence is analyzed in Proposition 4. The computational efficiency of Algorithm 2 is notable. For each iteration, two key steps are performed: computing $\mathbf{d}_t^k$ and solving Problem (18). The former has a computational cost of $O(tN)$, while the latter requires $O(N\log N)$. Overall, the algorithm's time complexity is $O(\tau_{\text{iter}}(NT + N\log N))$, where τ_{iter} is the number of iterations to convergence.

Proposition 4: With $\gamma^k \in (0, 1]$, $\gamma^k \to 0$ and $\sum_k \gamma^k = +\infty$, Algorithm 2 converges in a finite number of iterations to an

Algorithm 2 Efficient OFTRL with SCA (OFTRL-SCA)

Initialize $k = 1$, $\mathbf{b}_{t+1}^1 = \mathbf{b}_t$ and $\{\gamma^k\}$
repeat:
  Compute $\mathbf{d}_t^k$ according to (19)
  Solve (18) by Proposition 3 and obtain $\widehat{\mathbf{b}}_{t+1}^{k+1}$
  Compute $\mathbf{b}_{t+1}^{k+1} = \mathbf{b}_{t+1}^k + \gamma^k \left(\widehat{\mathbf{b}}_{t+1}^{k+1} - \mathbf{b}_{t+1}^k\right)$
  $k \leftarrow k+1$
until convergence
Output $\mathbf{b}_{t+1} = \mathbf{b}_{t+1}^k$

optimal solution of (16) or every limit point of the sequence $\{\mathbf{b}_{t+1}^k\}_{k=1}^{\infty}$ (at least one such point exists) is an optimal solution of (16).

Proof: Since [27, Assumptions A1-A4] hold and (16) is a convex problem, the proof for Proposition (4) follows directly from [27, Theorem 3]. □

In practice, truncation errors can arise due to the need for computational efficiency for large-scale problems. Despite this, we demonstrate that under specific termination conditions for Algorithm 2, the same order of regret as the exact solution can be maintained.

Let $\tilde{\mathbf{b}}_{t+1} = \mathbf{b}_{t+1}^k$ be the point generated by Algorithm 2 based on a predefined stopping criterion. The regret with inexact solutions is given by:

$$\tilde{R}_{[T]}^{\text{OFTRL}} = \sum_{t=1}^{T} f_t(\tilde{\mathbf{b}}_t) - \min_{\mathbf{b}\in\mathcal{W}} \sum_{t=1}^{T} f_t(\mathbf{b}). \tag{22}$$

Since f_t is convex, we can bound the difference between inexact point $\tilde{\mathbf{b}}_t$ and exact point $\mathbf{b}_t$ under Assumption 1 as:

$$f_t(\tilde{\mathbf{b}}_t) - f_t(\mathbf{b}_t) \le \langle \nabla f_t(\tilde{\mathbf{b}}_t), \tilde{\mathbf{b}}_t - \mathbf{b}_t\rangle \le G \left\|\tilde{\mathbf{b}}_t - \mathbf{b}_t\right\|_2.$$

Thus, the regret with inexact solutions can be bounded by:

$$\tilde{R}_{[T]}^{\text{OFTRL}} \le R_{[T]}^{\text{OFTRL}} + G \sum_{t=1}^{T} \left\|\tilde{\mathbf{b}}_t - \mathbf{b}_t\right\|_2, \tag{23}$$

where $R_{[T]}^{\text{OFTRL}}$ is the regret for exact solutions. To bound the gap between exact and inexact points, we introduce the following stopping criterion for Algorithm 2:

$$\left\|\widehat{\mathbf{b}}_{t+1}^k - \mathbf{b}_{t+1}^{k-1}\right\|_2 \le C_{\text{th}}/T^2, \tag{24}$$

where C_{th} is a predefined threshold constant for Algorithm 2. By enforcing this stopping criterion, we ensure that the regret with inexact solutions remains within the same order as the regret with exact solutions, as stated in Proposition 5.

Proposition 5: By setting the stopping criterion as $\left\|\widehat{\mathbf{b}}_{t+1}^k - \mathbf{b}_{t+1}^{k-1}\right\|_2 \le C_{\text{th}}/T^2$, with C_{th} being a predefined threshold constant for Algorithm 2, we have

$$\sum_{t=1}^{T} \left\|\tilde{\mathbf{b}}_t - \mathbf{b}_t\right\|_2 \le C_{\text{e}}, \tag{25}$$

where $\tilde{\mathbf{b}}_t = \mathbf{b}_t^k$ and C_{e} is a constant.

The proof for Proposition 5 is provided in Appendix D. Therefore, we demonstrate that even with inexact solutions produced by Algorithm 2, the same order of regret can still be ensured by appropriately choosing the stopping criterion.

V. Connection With Mean-Variance Portfolio

In this section, we establish the connection between the SAOFTRL framework and the mean-variance portfolio. Specifically, in Section V-A, we demonstrate that the underlying OFTRL strategy in the SAOFTRL framework can be viewed as the popular mean-variance portfolio with shrinkage estimators. Furthermore, in Section V-B, we illustrate how the use of active intervals in SAOFTRL offers an adaptive approach to selecting the lookback window.

A. Connection With Shrinkage Estimators

In a mean-variance portfolio, the objective is to maximize the expected portfolio return while maintaining a specified level of risk. This involves assessing the risks and expected returns of assets using the sample covariance matrix $\boldsymbol{\Sigma}$ and sample mean $\boldsymbol{\mu}$ for the asset returns and minimizing the following optimization problem:

$$\begin{aligned} \underset{\mathbf{b}\in\mathbb{R}^N}{\text{minimize}} \quad & \frac{1}{2}\mathbf{b}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{b} - \mathbf{b}^\mathsf{T}\boldsymbol{\mu} \\ \text{subject to} \quad & \mathbf{b} \in \mathcal{W}. \end{aligned} \tag{26}$$

However, the effectiveness of the resulting portfolios is significantly influenced by the accuracy of these estimators. Notably, in cases of small sample sizes, the variance of the estimator becomes a dominant source of error [41], [42], [43].

It's well known that such variance can be mitigated by introducing a deliberate bias, thereby achieving a trade-off for a reduced error [44]. This strategy involves shrinking the estimator towards specific target values as follows:

$$\hat{\theta}^{\text{sh}} = (1-\rho)\hat{\theta} + \rho\theta^{\text{tar}}. \tag{27}$$

In this equation, $\hat{\theta}$ is the empirical estimate (usually sample mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$), and θ^{tar} denotes the predefined target values based on prior information or assumptions about the parameter. The shrinkage process outcome, $\hat{\theta}^{\text{sh}}$, is a combination of the empirical estimate and the target value, moderated by the shrinkage trade-off parameter ρ. The critical challenge is in selecting an optimal shrinkage trade-off parameter ρ and target θ^{tar} to minimize the mean squared error (MSE) or other criteria [45].

Our study reveals an intriguing parallel between the predictive component in the OFTRL strategy and the target value of the shrinkage estimator in the mean-variance portfolio. To demonstrate this connection, we approximate the loss function of OFTRL using the sample mean and covariance matrix. Considering the feasible set defined as a simplex $\mathcal{W} = \{\mathbf{b} \in \mathbb{R}^N \mid \mathbf{1}^\mathsf{T}\mathbf{b} = 1, \mathbf{b} > 0\}$, we employ a second-order Taylor expansion at $\mathbf{b}^k$. Since the absolute value of each element in the stock return vector $\mathbf{r}_t$ is small, with $|r_{t,i}| \ll 1$, we have

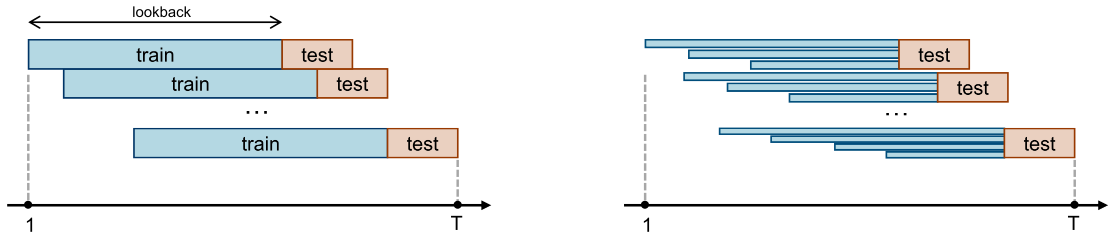


Fig. 2. Comparison of different learning schemes. **Left**: the learning scheme of portfolio selection with a fixed lookback sliding window; **Right**: the learning scheme of portfolio selection with multiple lookback sliding windows.

$|\mathbf{r}_t^\top \mathbf{b}| \ll 1$. Therefore, by carefully selecting $\mathbf{b}^k$, we can approximate $1 + \mathbf{r}_t^\top \mathbf{b}^k \approx 1$, and the loss function can be simplified as:

$$\begin{aligned} f_t(\mathbf{b}) &\approx -\log\left(1+\mathbf{r}_t^\top \mathbf{b}^k\right) - (\mathbf{b}-\mathbf{b}^k)^\top \left(\frac{\mathbf{r}_t}{1+\mathbf{r}_t^\top \mathbf{b}^k}\right) \\ &\quad + \frac{1}{2}\left(\mathbf{b}-\mathbf{b}^k\right)^\top \left(\frac{\mathbf{r}_t\mathbf{r}_t^\top}{\left(1+\mathbf{r}_t^\top \mathbf{b}^k\right)^2}\right)\left(\mathbf{b}-\mathbf{b}^k\right) \\ &\approx -\mathbf{b}^\top \mathbf{r}_t + \frac{1}{2}\mathbf{b}^\top \mathbf{r}_t \mathbf{r}_t^\top \mathbf{b} + \text{const.} \end{aligned}$$

Applying this to the OFTRL objective function yields:

$$\begin{aligned} &\sum_{\tau=1}^{t} f_\tau(\mathbf{b}) + \tilde{f}_{t+1}(\mathbf{b}) + \frac{\beta}{2}\|\mathbf{b}\|_2^2 \\ &\approx \frac{1}{2}\mathbf{b}^\top\left(\sum_{\tau=1}^{t}\mathbf{r}_\tau\mathbf{r}_\tau^\top + \beta\mathbf{I}\right)\mathbf{b} - \mathbf{b}^\top\left(\sum_{\tau=1}^{t}\mathbf{r}_\tau - \mathbf{m}_{t+1}\right) \\ &\approx \frac{1}{2}\mathbf{b}^\top\left(t\boldsymbol{\Sigma} + \beta\mathbf{I}\right)\mathbf{b} - \mathbf{b}^\top\left(t\boldsymbol{\mu} - \mathbf{m}_{t+1}\right). \end{aligned} \tag{28}$$

Here, the sample covariance matrix $\boldsymbol{\Sigma}$ is calculated by assuming $\mathbb{E}[\mathbf{r}] = 0$. As a result, OFTRL emerges as a specialized MVP that incorporates shrinkage estimators. In this setup, the target value of the mean and covariance matrix are represented by $-\mathbf{m}_{t+1}$ and $\beta\mathbf{I}$, respectively, while the shrinkage trade-off factor is $\rho = \frac{1}{t+1}$. To further substantiate our findings, we conduct experiments in Section VI-B to validate the equivalence numerically.

Recalling that SAOFTRL calculates portfolio instances over active intervals, which are derived from the GC intervals listed in Table I. It is worth noting that these GC intervals often consist of numerous small intervals, which can contribute to high variance in the resulting portfolio. However, the utilization of the OFTRL effectively mitigates this issue by incorporating predictive elements and regularization terms. This leads to a reduction in variance and ultimately stabilizes the performance of the resulting portfolio.

Moreover, OFTRL inherently provides a theoretically justified choice of the shrinkage factor $\rho = \frac{1}{t+1}$ for t data samples. This selection ensures a worst-case performance guarantee, as stated in Theorem 1, even when dealing with uninformative target values. This simplicity in determining the shrinkage factor adds to the practical appeal of OFTRL.

B. Connection With Adaptive Lookback Windows

In managing mean-variance portfolios with streaming data, a prevalent approach involves employing a rolling window with a fixed lookback period [46], as illustrated in the left-hand side of Fig. 2. The choice of an optimal lookback size is pivotal. Larger lookback periods are beneficial as they yield more reliable estimates for the expected return $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. However, excessively large lookback periods can lead to issues due to the nonstationary nature of financial data. Older data may no longer accurately reflect current return and variance trends, which can adversely impact portfolio performance. Therefore, finding the right balance in the lookback size is a critical yet complex task in the management of mean-variance portfolios.

Our proposed SAOFTRL framework provides an innovative solution to address this challenge. Specifically, SAOFTRL allows for dynamic adjustment of the lookback period for each round, as shown on the right-hand side of Fig. 2. In each round, SAOFTRL runs multiple portfolio instances with different lookback sizes and aggregates them based on past performance. The length of the lookback size utilized for each round is strategically distributed on a logarithmic scale to ensure optimal computational efficiency. Importantly, as stated in Equation (5), the design of the set of lookback sizes enables rapid convergence to the BCRP over any arbitrary intervals $I \subseteq [T]$. This guarantees portfolio performance and provides a significant advancement in adaptive portfolio management.

This strong correspondence between SAOFTRL and MVP offers dual advantages. Firstly, it provides a worst-case performance guarantee under the SAOFTRL strategy, ensuring robust performance even in challenging scenarios. Secondly, it offers an average-case performance guarantee under the mean-variance framework, providing confidence for the portfolio selection in the general stock market cases.

VI. Numeric Experiments

This section presents extensive simulations to demonstrate the effectiveness and efficiency of our proposed methods.

A. Performance Comparison With Benchmarks

To validate the effectiveness of our proposed methods, we conducted simulations using the stock lists of the S&P 500 and NASDAQ 100 indices. For each stock list, we sampled 80 datasets from January 1st, 2010 to January 1st, 2023. Each dataset consisted of 40 assets observed over a 1-year period,

TABLE II
HYPERPARAMETERS OF THE DATASETS AND BACKTESTING

	Hyper-Parameters	Value
dataset	# datasets	80
	# assets	40
	# samples per dataset	252
	data period	2010-01-01 $\sim$ 2023-01-01
backtesting	training history	189
	optimize every	1
	rebalance every	1
	price name	adjusted
	# bars per year	252
	cost	buy = 0, sell = 0

amounting to 252 observations. The training history for each dataset spanned 9 months, equivalent to 189 observations. The backtesting process was facilitated by utilizing the 'portfolioBacktest' R package [47]. Further details and specific parameters can be found in Table II.

We utilize the annualized Sharpe ratio and max drawdown to illustrate the average-case performance in the stock market. Specifically, the annualized Sharpe ratio [48] is our primary metric for performance assessment:

$$\text{Sharpe ratio} = \sqrt{252} \cdot \frac{\mathbb{E}[\mathbf{b}^{\mathsf{T}}\mathbf{r}]}{\sqrt{\text{Var}\left[\mathbf{b}^{\mathsf{T}}\mathbf{r}\right]}}. \tag{29}$$

This ratio measures the annualized return per unit of annualized risk, and we complement it with the max drawdown (maximum loss from a peak to a trough) for a comprehensive risk assessment.

To further illustrate the worst-case performance of different methods, we also present the comparison of average strongly adaptive (SA) regret, with the value of SA-$\mathrm{R}_T^{(\tau)}/\tau$, versus different interval length τ. This metric indicates the largest average gap compared to the oracle during all consecutive trading periods with the same length τ, serving as a worst-case performance indicator in changing environments.

Our benchmarks include the Equal Weights (EW) portfolio, FTL [13], Online Newton Step (ONS) [12], FTRL [14], and various OFTRL strategies with different $\mathbf{m}_{t+1}$ predictions: last ($\mathbf{m}_{t+1} = \mathbf{g}_t$) [25] and average ($\mathbf{m}_{t+1} = \frac{1}{t}\sum_{\tau=1}^{t} \mathbf{g}_\tau$). We also include Optimistic Follow-the-Leader (OFTL) for comparison, which has the same formulation as OFTRL but without regularization terms [49]. The $\mathbf{m}_{t+1}$ settings in OFTL in OFTL are the same as in OFTRL for simulations. We compare the above benchmarks with the proposed SAOFTRL framework under different $\mathbf{m}_{t+1}$ settings: zero, last, and average.

In our simulations, we systematically vary the regularization weight β for the FTRL, OFTRL, and SAOFTRL methods to assess the Sharpe ratio and max-drawdown. This is done to evaluate their performance and examine the sensitivity of the parameters. We compare their median Sharpe ratios and maximum drawdowns across 80 datasets, as depicted in Fig. 3. Interestingly, OFTRL with average historical gradients (OFTRL-avg) demonstrates a performance nearly identical to that of FTRL. This similarity can be attributed to the fact that in OFTRL-avg, the average historical gradients tend to converge to a stationary point over long-term data periods, diminishing the impact of the prediction term and aligning its performance with FTRL.

Furthermore, we observe that the Sharpe ratios of OFTRL with different predictors are worse than the EW strategy. This can be attributed to the relatively small expected returns μ in the real stock market. The small expected returns make the portfolio updating process more conservative, resulting in a slower convergence speed to the optimal portfolio compared to a more aggressive strategy.

Our proposed SAOFTRL-avg method outperforms all others, including SAOFTRL-zero. As discussed in Section V-A, this superior performance is largely due to the utilization of historical averages, which effectively reduces variance in errors for small intervals, thereby enhancing overall performance. Although the naive EW portfolio is generally difficult to outperform, as highlighted in the literature [50], our experiments show that, in most cases, our proposed methods achieve a higher Sharpe ratio compared to EW. Overall, SAOFTRL strategies exhibit a stable equilibrium between risk and return compared to other methods. Notably, SAOFTRL-avg consistently excels in the Sharpe ratio across various β settings and maintains the lowest maximum drawdown in most scenarios, underscoring its robustness.

We then fix the regularization weight $\beta = 6$ for the FTRL, OFTRL, and SAOFTRL methods to compare the average SA regret across different interval lengths τ. We analyze the median values across 80 datasets, as shown in Fig. 3. The SA-regret curves of different methods in the NASDAQ 100 dataset are nearly identical. This behavior is due to the specific characteristics of the NASDAQ data, which appears to have a stable market trend during the evaluation period. In such scenarios, the additional benefits of the regularization term in OFTRL and SAOFTRL, which primarily aids in more volatile or adversarial environments, do not significantly affect the outcome. Consequently, this results in similar behavior under such stable conditions. In datasets with higher volatility or adversarial conditions, as shown in our results for the S&P 500 dataset, the advantages of OFTRL and SAOFTRL become more apparent, as the regularization term helps manage the increased variance in the data. Thus, while the regret performance of different methods may be similar in stable markets like NASDAQ during this period, methods with regularization terms remain more robust across a variety of challenging market conditions, which is consistent with our theoretical analysis. The exceptional performance of SAOFTRL across different datasets underscores its superb worst-case performance in changing environments.

B. Validation of Equivalence Between OFTRL and MVP With Shrinkage Estimators

To verify our finding in Section V-A, we conduct a comprehensive comparison between the optimal points obtained through OFTRL and those derived from the MVP with shrinkage estimators. Our analysis focuses on evaluating the differences between these optimal points based on the ℓ_1 norm, considering various data dimensions from $N = 100$ to

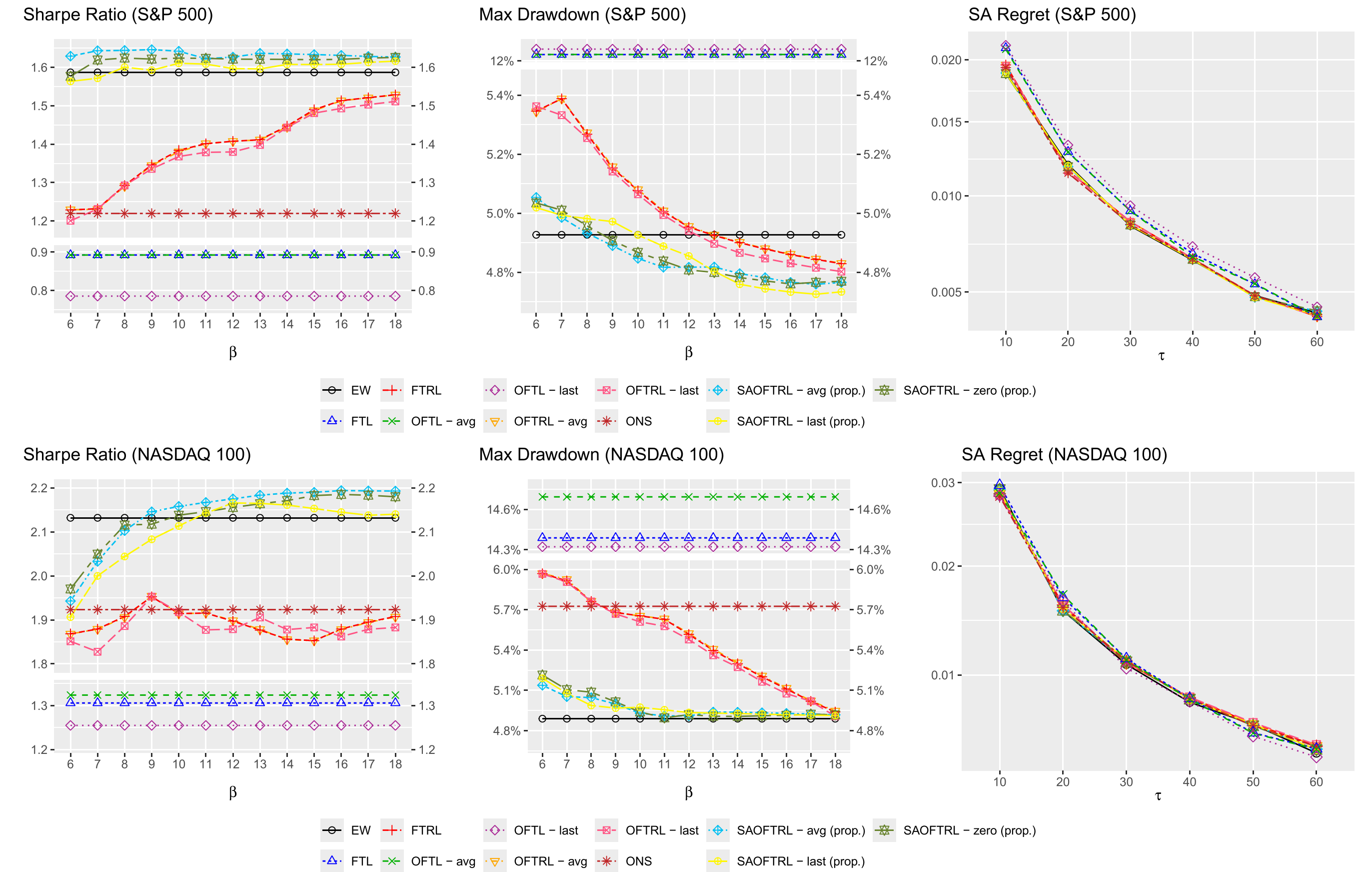


Fig. 3. Performance comparison on S &P 500 (**Top**) and NASDAQ 100 (**Bottom**) datasets. **Left** and **Middle**: Sharpe ratio and max drawdown across varying regularization weights β; **Right**: SA regret over different interval lengths τ with fixed $\beta = 6$.

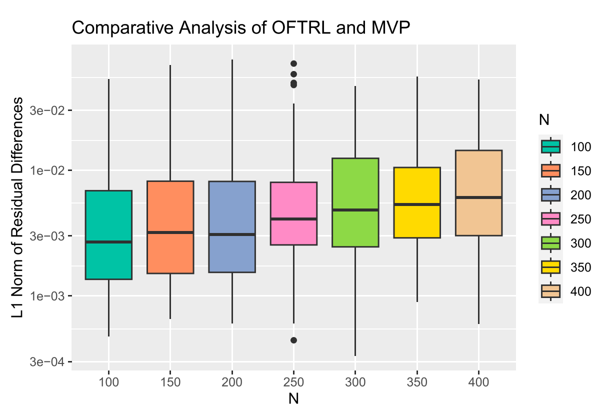


Fig. 4. Comparison of ℓ_1 norm residuals between optimal points from OFTRL and MVP with shrinkage estimators across different data dimensions N.

$N = 400$. For each dimension, we generate 100 datasets, each containing $T = 400$ observations. The datasets are generated by sampling $\mathbf{r}_t \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} \in \mathbb{R}^N$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{N\times N}$ are the sample mean and sample covariance matrix calculated using real market data from the S&P 500.

The resulting ℓ_1 norm residuals are then visualized using a boxplot, as shown in Fig. 4. The observed distribution patterns demonstrate that the distance between the optimal points generated by OFTRL and MVP consistently remains small, even in high-dimensional data scenarios. This consistent proximity across different dimensions robustly demonstrates the strong correspondence between OFTRL and MVP.

C. Speed Comparison of Different Acceleration Schemes

This section evaluates the computational efficiency of our proposed efficient method in Section IV against benchmarks across varying problem dimensions N. We specifically compare the CPU time of our efficient implementation with the widely-used convex problem solver 'CVXR' [51] and Projected Gradient Descent (PGD). Given that OFTRL can be approximated as a mean-variance portfolio (MVP) with specific shrinkage estimators, as discussed in Section V-A, we include solving MVP with an efficient QP solver 'quadprog' [52] in our efficiency comparison. The experiments are carried out on a PC equipped with a 13th Gen Intel(R) Core(TM) i7-13700 CPU and 16GB of RAM, running R 4.1.3.

In 100 randomized trials across datasets with varying data dimensions N, we observe CPU time for each method as depicted in Fig. 5. The results show that OFTRL-SCA consumes the least CPU time compared to other methods. OFTRL-SCA demonstrates a significant speed advantage, being almost 100

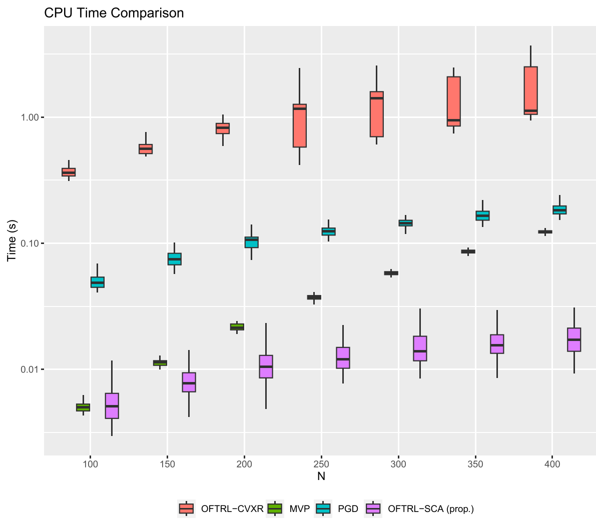


Fig. 5. CPU time comparison of 100 randomized trials on datasets with N assets.

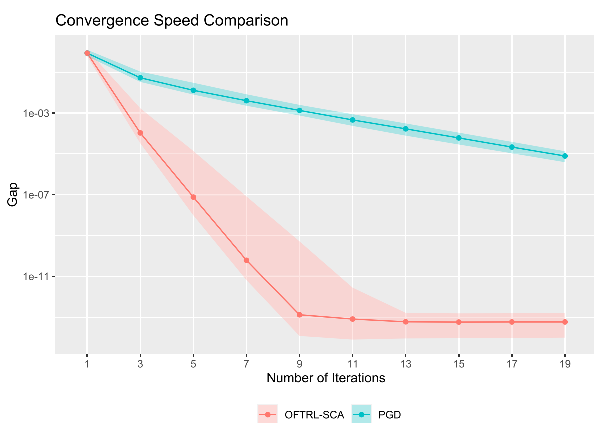


Fig. 6. Convergence speed comparison over 100 randomized trials on datasets with 400 assets, with the gap indicating the distance from the optimal point.

times faster than CVXR and 10 times faster than MVP and PGD for high-dimensional data.

We further compare the convergence speed between OFTRL-SCA and PGD over 100 randomized trials, as illustrated in Fig. 6. The gap indicates the distance from the optimal point. We observe that OFTRL-SCA consistently converges to the optimal point after a few iterations, while PGD maintains a slower convergence speed across all trials. This clearly demonstrates that OFTRL-SCA accelerates the optimization process, leading to quicker convergence to the optimal point.

These results highlight the computational efficiency of our proposed efficient methods, particularly OFTRL-SCA, compared to existing benchmarks. The faster convergence and reduced computation time make our methods highly practical and applicable in real-world scenarios with large-scale portfolio optimization problems.

VII. Conclusion and Future Directions

In this paper, we introduce a novel framework called SAOFTRL with efficient implementation for online portfolio selection in volatile stock markets. Our framework offers a worst-case performance guarantee, providing a novel regret. Moreover, we establish an intriguing equivalence between SAOFTRL and MVP, which instills confidence in its average performance within the MVP context. Through extensive numerical experiments, we demonstrate the efficiency and efficacy of our proposed methods.

Furthermore, our findings highlight the potential of making more informed predictions in portfolio management while still maintaining a safeguard for worst-case performance, as stated in Theorem 1. This opens up avenues for further research in improving time series predictions using methods such as ARMA, ARIMA, Kalman filter, and other online time series techniques [53], [54], [55]. Integrating these time series prediction methods into SAOFTRL has the potential to enhance the regret guarantee and overall effectiveness of the algorithm. This integrative approach represents a fruitful direction for future research, aiming to combine robust theoretical foundations with practical, data-driven prediction techniques.

Appendix

A. Proof of Theorem 1

We can interpret the OFTRL as FTRL with a regularization term $\tilde{\psi}_t(\mathbf{b}) = \psi_t(\mathbf{b}) + \tilde{f}_t(\mathbf{b})$, where $\psi_t(\mathbf{b}) = \frac{\beta}{2}\|\mathbf{b}\|_2^2$ and $\tilde{f}_t(\mathbf{b}) = \langle \mathbf{m}_t, \mathbf{b}\rangle$. Let $\tilde{F}_t(\mathbf{b}) = \sum_{s=1}^{t-1} f_s(\mathbf{b}) + \tilde{\psi}_t(\mathbf{b})$ and $\bar{F}_t(\mathbf{b}) = \sum_{s=1}^{t-1} f_s(\mathbf{b}) + \psi_t(\mathbf{b})$. According to the update rule for OFTRL, we have $\mathbf{b}_t = \arg\min_{\mathbf{b}\in\mathcal{W}} \tilde{F}_t(\mathbf{b})$. Since $\tilde{f}_{T+1}(\mathbf{b})$ has no influence on the algorithm, we can set it to the null function. According to [30, Lemma 7.1], we have

$$
\begin{aligned}
R(\mathbf{u}) &= \sum_{t=1}^{T} f_t(\mathbf{b}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}) \\
&= \tilde{\psi}_{T+1}(\mathbf{u}) - \min_{\mathbf{b}\in\mathcal{W}} \tilde{\psi}_1(\mathbf{b}) + \underbrace{\tilde{F}_{T+1}(\mathbf{b}_{T+1}) - \tilde{F}_{T+1}(\mathbf{u})}_{\leq 0} \\
&\quad + \sum_{t=1}^{T} \left[\tilde{F}_t(\mathbf{b}_t) + f_t(\mathbf{b}_t) - \tilde{F}_{t+1}(\mathbf{b}_{t+1})\right] \\
&\leq \tilde{\psi}_{T+1}(\mathbf{u}) - \min_{\mathbf{b}\in\mathcal{W}} \tilde{\psi}_1(\mathbf{b}) \\
&\quad + \sum_{t=1}^{T} \left[\tilde{F}_t(\mathbf{b}_t) + f_t(\mathbf{b}_t) - \tilde{F}_{t+1}(\mathbf{b}_{t+1})\right] \\
&= \psi_{T+1}(\mathbf{u}) + \tilde{f}_{T+1}(\mathbf{u}) - \min_{\mathbf{b}\in\mathcal{W}} \left(\psi_1(\mathbf{b}) + \tilde{f}_1(\mathbf{b})\right) \\
&\quad + \sum_{t=1}^{T} \left[\bar{F}_t(\mathbf{b}_t) + \tilde{f}_t(\mathbf{b}_t) + f_t(\mathbf{b}_t)\right] \\
&\quad - \sum_{t=1}^{T} \left[\bar{F}_{t+1}(\mathbf{b}_{t+1}) + \tilde{f}_{t+1}(\mathbf{b}_{t+1})\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \psi_{T+1}(\mathbf{u}) - \psi_1(\mathbf{b}_1) \\
&\quad + \sum_{t=1}^{T} \left[\bar{F}_t(\mathbf{b}_t) + f_t(\mathbf{b}_t) - \bar{F}_{t+1}(\mathbf{b}_{t+1})\right] \\
&= \psi_{T+1}(\mathbf{u}) - \psi_1(\mathbf{b}_1) + \sum_{t=1}^{T} \left[\bar{F}_{t+1}(\mathbf{b}_t) - \bar{F}_{t+1}(\mathbf{b}_{t+1})\right] \\
&\leq \sum_{t=1}^{T} \left[\bar{F}_{t+1}(\mathbf{b}_t) - \bar{F}_{t+1}(\mathbf{b}_{t+1})\right] + \frac{\beta}{2}B^2.
\end{aligned}
$$

Now we focus on bounding the term $\bar{F}_{t+1}(\mathbf{b}_t) - \bar{F}_{t+1}(\mathbf{b}_{t+1})$. Observe that $\bar{F}_{t+1}(\mathbf{b})$ is an exp-concave function, hence we have:

$$
\begin{aligned}
\bar{F}_{t+1}(\mathbf{b}_t) - \bar{F}_{t+1}(\mathbf{b}_{t+1}) &\leq \langle \mathbf{g}_t', \mathbf{b}_t - \mathbf{b}_{t+1}\rangle \\
&\quad - \frac{1}{2}(\mathbf{b}_t - \mathbf{b}_{t+1})^\mathsf{T}\mathbf{A}_t(\mathbf{b}_t - \mathbf{b}_{t+1}),
\end{aligned}
$$

where $\mathbf{g}_t' = \nabla \bar{F}_{t+1}(\mathbf{b}_t)$ and $\mathbf{A}_t = \beta\mathbf{I} + \sum_{\tau=1}^{t} \alpha \mathbf{g}_\tau \mathbf{g}_\tau^\mathsf{T}$ with $\mathbf{g}_t = \nabla f_t(\mathbf{b}_t)$ and $\alpha = \frac{1}{8GB}$. As a shorthand, we denote $\|\mathbf{b}_t - \mathbf{b}_{t+1}\|_{\mathbf{A}_t} = (\mathbf{b}_t - \mathbf{b}_{t+1})^\mathsf{T}\mathbf{A}_t(\mathbf{b}_t - \mathbf{b}_{t+1})$. Note that $\psi_{t+1}(\mathbf{b}) = \psi_t(\mathbf{b})$, we have $\bar{F}_{t+1}(\mathbf{b}_t) = \bar{F}_t(\mathbf{b}_t) + f_t(\mathbf{b}_t)$, thus $\mathbf{g}_t' = \nabla \bar{F}_t(\mathbf{b}_t) + \nabla f_t(\mathbf{b}_t)$. According to the updating rule of OFTRL, we have $\mathbf{b}_t = \arg\min_{\mathbf{b}\in\mathcal{W}} \bar{F}_t(\mathbf{b}) + \tilde{f}_t(\mathbf{b})$, thus $\mathbf{0} = \nabla \bar{F}_t(\mathbf{b}_t) + \nabla \tilde{f}_t(\mathbf{b}_t)$. Therefore, using [30, Theorem 2.16], we have $\mathbf{g}_t' = \nabla f_t(\mathbf{b}_t) - \nabla \tilde{f}_t(\mathbf{b}_t) = \mathbf{g}_t - \mathbf{m}_t$. Therefore, we have

$$
\begin{aligned}
\bar{F}_{t+1}(\mathbf{b}_t) - \bar{F}_{t+1}(\mathbf{b}_{t+1}) &\leq \langle \mathbf{g}_t - \mathbf{m}_t, \mathbf{b}_t - \mathbf{b}_{t+1}\rangle \\
&\quad - \frac{1}{2}\|\mathbf{b}_t - \mathbf{b}_{t+1}\|_{\mathbf{A}_t}.
\end{aligned}
$$

With the property of dual norm such that $\langle a, b\rangle \leq \|a\|_* \|b\| \leq \frac{1}{2c}\|a\|_*^2 + \frac{c}{2}\|b\|^2$, we have:

$$
\langle \mathbf{g}_t - \mathbf{m}_t, \mathbf{b}_t - \mathbf{b}_{t+1}\rangle \leq \frac{1}{2}\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2 + \frac{1}{2}\|\mathbf{b}_t - \mathbf{b}_{t+1}\|_{\mathbf{A}_t}^2.
$$

Summing over 1 to T, we have

$$
\begin{aligned}
&\sum_{t=1}^{T}\left[\bar{F}_{t+1}(\mathbf{b}_t) - \bar{F}_{t+1}(\mathbf{b}_{t+1})\right] \\
&\leq \sum_{t=1}^{T}\left(\langle \mathbf{g}_t - \mathbf{m}_t, \mathbf{b}_t - \mathbf{b}_{t+1}\rangle - \frac{1}{2}\|\mathbf{b}_t - \mathbf{b}_{t+1}\|_{\mathbf{A}_t}\right) \\
&\leq \sum_{t=1}^{T}\left(\frac{1}{2}\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2\right).
\end{aligned}
$$

Consequently, the regret for OFTRL over the interval $[T]$ is given by:

$$
R_{[T]}^{\mathrm{OFTRL}} \leq \frac{1}{2}\sum_{t=1}^{T}\|\mathbf{g}_t - \mathbf{m}_t\|_{\mathbf{A}_t^{-1}}^2 + \frac{\beta}{2}B^2. \tag{30}
$$

B. Proof of Lemma 1

First of all, we have

$$
\begin{aligned}
&\sum_{s=1}^{t}(\mathbf{g}_s - \mathbf{m}_s)(\mathbf{g}_s - \mathbf{m}_s)^\mathsf{T} \\
&= \mathbf{g}_1\mathbf{g}_1^\mathsf{T} + \sum_{s=2}^{t}\left(\mathbf{g}_s - \frac{1}{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_\tau\right)\left(\mathbf{g}_s - \frac{1}{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_\tau\right)^\mathsf{T} \\
&= \mathbf{g}_1\mathbf{g}_1^\mathsf{T} + \sum_{s=2}^{t}\left(\mathbf{g}_s\mathbf{g}_s^\mathsf{T} - \frac{1}{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_s\mathbf{g}_\tau^\mathsf{T} - \frac{1}{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_\tau\mathbf{g}_s^\mathsf{T}\right) \\
&\quad + \sum_{s=2}^{t}\left(\frac{1}{(s-1)^2}\sum_{r=1}^{s-1}\mathbf{g}_r\sum_{\tau=1}^{s-1}\mathbf{g}_\tau^\mathsf{T}\right) \\
&= \sum_{s=1}^{t}\mathbf{g}_s\mathbf{g}_s^\mathsf{T} + \sum_{s=2}^{t}\sum_{r=1}^{s-1}\left(-\frac{1}{s-1}\left[\mathbf{g}_s\mathbf{g}_r^\mathsf{T} + \mathbf{g}_r\mathbf{g}_s^\mathsf{T}\right]\right) \\
&\quad + \sum_{s=2}^{t}\frac{1}{(s-1)^2}\sum_{r=1}^{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_r\mathbf{g}_\tau^\mathsf{T}.
\end{aligned} \tag{31}
$$

The last term of Equation (31) could be written as:

$$
\begin{aligned}
&\sum_{s=2}^{t}\frac{1}{(s-1)^2}\sum_{r=1}^{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_r\mathbf{g}_\tau^\mathsf{T} \\
&= \sum_{s=2}^{t}\frac{1}{(s-1)^2}\sum_{\tau=1}^{s-1}\mathbf{g}_\tau\mathbf{g}_\tau^\mathsf{T} \\
&\quad + \sum_{s=2}^{t}\frac{1}{(s-1)^2}\sum_{r=1}^{s-1}\left(\sum_{\tau=1}^{r-1}\mathbf{g}_r\mathbf{g}_\tau^\mathsf{T} + \sum_{\tau=r+1}^{s-1}\mathbf{g}_r\mathbf{g}_\tau^\mathsf{T}\right) \\
&= \sum_{s=1}^{t-1}\left(\sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\mathbf{g}_s\mathbf{g}_s^\mathsf{T} \\
&\quad + \sum_{s=2}^{t-1}\sum_{r=1}^{s-1}\left(\sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\left[\mathbf{g}_r\mathbf{g}_s^\mathsf{T} + \mathbf{g}_s\mathbf{g}_r^\mathsf{T}\right].
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&\sum_{s=1}^{t}(\mathbf{g}_s - \mathbf{m}_s)(\mathbf{g}_s - \mathbf{m}_s)^\mathsf{T} \\
&= \sum_{s=1}^{t}\left(1 + \sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\mathbf{g}_s\mathbf{g}_s^\mathsf{T} \\
&\quad + \sum_{s=2}^{t}\sum_{r=1}^{s-1}\left(-\frac{1}{s-1} + \sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\left[\mathbf{g}_r\mathbf{g}_s^\mathsf{T} + \mathbf{g}_s\mathbf{g}_r^\mathsf{T}\right].
\end{aligned}
$$

For $s > 2$, we have

$$
\frac{1}{s} - \frac{1}{t} = \int_s^t \frac{1}{x^2}dx \leq \sum_{\tau=s}^{t-1}\frac{1}{\tau^2} \leq \int_{s-1}^{t-1}\frac{1}{x^2}dx = \frac{1}{s-1} - \frac{1}{t-1},
$$

thus

$$
\left|-\frac{1}{s-1} + \sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right| \leq \frac{1}{(s-1)^2} + \frac{1}{t}.
$$

Since $(\mathbf{g}_r + \mathbf{g}_s)(\mathbf{g}_r + \mathbf{g}_s)^\mathsf{T} \succeq 0$ and $(\mathbf{g}_r - \mathbf{g}_s)(\mathbf{g}_r - \mathbf{g}_s)^\mathsf{T} \succeq 0$, thus we have $\mathbf{g}_r\mathbf{g}_r^\mathsf{T} + \mathbf{g}_s\mathbf{g}_s^\mathsf{T} \succeq \pm\left(\mathbf{g}_r\mathbf{g}_s^\mathsf{T} + \mathbf{g}_s\mathbf{g}_r^\mathsf{T}\right)$. Therefore, we have

$$
\begin{aligned}
&\sum_{s=1}^{t}(\mathbf{g}_s - \mathbf{m}_s)(\mathbf{g}_s - \mathbf{m}_s)^\mathsf{T} \\
&= \sum_{s=1}^{t}\left(1 + \sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\mathbf{g}_s\mathbf{g}_s^\mathsf{T} \\
&\quad + \sum_{s=2}^{t}\sum_{r=1}^{s-1}\left(-\frac{1}{s-1} + \sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\left[\mathbf{g}_r\mathbf{g}_s^\mathsf{T} + \mathbf{g}_s\mathbf{g}_r^\mathsf{T}\right]
\end{aligned}
$$

$$
\begin{aligned}
&\le \sum_{s=1}^{t}\left(1+\sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} \\
&\quad + \sum_{s=2}^{t}\sum_{r=1}^{s-1}\left|-\frac{1}{s-1}+\sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right|\left[\mathbf{g}_r\mathbf{g}_r^{\mathsf{T}}+\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}}\right] \\
&\le \underbrace{\sum_{s=1}^{t}\left(1+\sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}}}_{①} \\
&\quad + \underbrace{\sum_{s=2}^{t}\sum_{r=1}^{s-1}\left(\frac{1}{(s-1)^2}+\frac{1}{t}\right)\left[\mathbf{g}_r\mathbf{g}_r^{\mathsf{T}}+\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}}\right]}_{②}.
\end{aligned}
$$

Since

$$
1+\sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\le 2+\int_{s}^{t-1}\frac{1}{x^2}dx = 2+\frac{1}{s}-\frac{1}{t-1}\le 3,
$$

we have

$$
① = \sum_{s=1}^{t}\left(1+\sum_{\tau=s}^{t-1}\frac{1}{\tau^2}\right)\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} \le 3\sum_{s=1}^{t}\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}}.
$$

Besides, we have

$$
\begin{aligned}
② &= \sum_{s=2}^{t}\sum_{r=1}^{s-1}\left(\frac{1}{(s-1)^2}+\frac{1}{t}\right)\left[\mathbf{g}_r\mathbf{g}_r^{\mathsf{T}}+\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}}\right] \\
&= \sum_{s=2}^{t}\left(\frac{1}{s-1}+\frac{s-1}{t}\right)\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} \\
&\quad + \sum_{s=2}^{t}\sum_{r=1}^{s-1}\left(\frac{1}{(s-1)^2}+\frac{1}{t}\right)\mathbf{g}_r\mathbf{g}_r^{\mathsf{T}} \\
&\le 2\sum_{s=2}^{t}\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} + \sum_{s=2}^{t}\left(\frac{1}{(s-1)^2}+\frac{1}{t}\right)\sum_{r=1}^{s-1}\mathbf{g}_r\mathbf{g}_r^{\mathsf{T}} \\
&= 2\sum_{s=2}^{t}\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} + \sum_{s=1}^{t}\left(\sum_{r=s}^{t-1}\left(\frac{1}{r^2}+\frac{1}{t}\right)\right)\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} \\
&\le 2\sum_{s=2}^{t}\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} + 3\sum_{s=1}^{t}\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}} \\
&\le 5\sum_{s=1}^{t}\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}}.
\end{aligned}
$$

To summarize, for $\mathbf{m}_s = \frac{1}{s-1}\sum_{\tau=1}^{s-1}\mathbf{g}_\tau$, we have

$$
\sum_{s=1}^{t}\left(\mathbf{g}_s-\mathbf{m}_s\right)\left(\mathbf{g}_s-\mathbf{m}_s\right)^{\mathsf{T}} = ① + ② \preceq 8\sum_{s=1}^{t}\mathbf{g}_s\mathbf{g}_s^{\mathsf{T}}.
$$

C. Supplementary Lemmas and Proofs

Lemma 3: Let $\bar{\mathbf{g}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_t$, $Q_{[T]} = \sum_{t=1}^{T}\|\mathbf{g}_t-\bar{\mathbf{g}}\|_2^2$, $\tilde{Q}_{[T]} = \sum_{t=1}^{T}\|\mathbf{g}_t-\mathbf{m}_t\|_2^2$ with $\mathbf{m}_t = \frac{1}{t-1}\sum_{\tau=1}^{t-1}\mathbf{g}_\tau$, we have $\tilde{Q}_{[T]} \le Q_{[T]} + 4G^2(\log T+1)$.

Proof: Consider the following loss function for an online learning algorithm:

$$
c_t(\mathbf{x}) = \|\mathbf{g}_t-\mathbf{x}\|_2^2. \tag{32}
$$

Our goal is to minimize the cumulative loss $\sum_{t=1}^{T}c_t(\mathbf{x})$. It's easy to see that the best-fixed point in hindsight is $\mathbf{x}^\star = \frac{1}{T}\sum_{t=1}^{\mathsf{T}}\mathbf{g}_t = \bar{\mathbf{g}}$. Consider the Follow-the-Leader strategy [29], where we select the next point that minimizes the loss over the past rounds $\sum_{\tau=1}^{t}c_\tau(\mathbf{x})$, which is exactly $\mathbf{x}_{t+1} = \frac{1}{t}\sum_{\tau=1}^{t}\mathbf{g}_t = \mathbf{m}_{t+1}$. Then we have

$$
\begin{aligned}
R_{[T]} &= \sum_{t=1}^{T}c_t(\mathbf{m}_t)-c_t(\bar{\mathbf{g}}) \\
&= \underbrace{\sum_{t=1}^{T}\|\mathbf{g}_t-\mathbf{m}_t\|_2^2}_{\tilde{Q}_{[T]}} - \underbrace{\sum_{t=1}^{T}\|\mathbf{g}_t-\bar{\mathbf{g}}\|_2^2}_{Q_{[T]}} \\
&\le \sum_{t=1}^{T}c_t(\mathbf{m}_t)-c_t(\mathbf{m}_{t+1}) \\
&= \sum_{t=1}^{T}\left(1-\left(1-\frac{1}{t}\right)^2\right)\|\mathbf{g}_t-\mathbf{m}_t\|_2^2 \\
&\le \sum_{t=1}^{T}\frac{1}{t}\|\mathbf{g}_t-\mathbf{m}_t\|_2^2 \\
&\le 4G^2\left(\log T+1\right),
\end{aligned}
$$

where the third inequality is from [29, Lemma 2.1] and the fourth equality follows from $\mathbf{m}_{t+1} = \frac{t-1}{t}\mathbf{m}_t + \frac{1}{t}\mathbf{g}_t$. □

D. Proof of Proposition 5

The objective function in Equation (16), denoted as $h_t(\mathbf{b})$, is β-strongly convex and smooth. For simplicity, we drop subscripts of $h_t(\mathbf{b})$ and $\mathbf{b}_{t+1}^{k+1}$ and denote as $h(\mathbf{b})$ and $\mathbf{b}^{k+1}$, respectively. Let $\mathbf{b}^\star$ represent the optimal point that minimizes $h(\mathbf{b})$. Given that $h(\mathbf{b})$ is convex and L-smooth, and $\mathbf{b}^{k+1} = \mathbf{b}^k + \gamma^k(\widehat{\mathbf{b}}^{k+1}-\mathbf{b}^k)$, we have the following:

$$
\begin{aligned}
h\left(\mathbf{b}^{k+1}\right) &\le h\left(\mathbf{b}^k\right) + \gamma^k\nabla h\left(\mathbf{b}^k\right)^{\mathsf{T}}\left(\widehat{\mathbf{b}}^{k+1}-\mathbf{b}^k\right) \\
&\quad + \frac{L}{2}\left(\gamma^k\right)^2\left\|\widehat{\mathbf{b}}^{k+1}-\mathbf{b}^k\right\|_2^2 \\
&\le h(\mathbf{b}^\star) - \nabla h(\mathbf{b}^k)^{\mathsf{T}}\left(\mathbf{b}^\star-\mathbf{b}^k\right) \\
&\quad + \gamma^k\nabla h\left(\mathbf{b}^k\right)^{\mathsf{T}}\left(\widehat{\mathbf{b}}^{k+1}-\mathbf{b}^k\right) \\
&\quad + \frac{L}{2}\left(\gamma^k\right)^2\left\|\widehat{\mathbf{b}}^{k+1}-\mathbf{b}^k\right\|_2^2,
\end{aligned}
$$

where the last inequality follows by the convexity of $h(\mathbf{b})$. Given that the feasible set is bounded, we assume $\left\|\nabla\hat{h}(\widehat{\mathbf{b}}^{k+1};\mathbf{b}^k)\right\|_2 \le Q$, where $\hat{h}(\widehat{\mathbf{b}}^{k+1};\mathbf{b}^k)$ is defined in Equation (17). Since $\nabla\hat{h}(\mathbf{b}^k;\mathbf{b}^k) = \nabla h(\mathbf{b}^k)$ and $\widehat{\mathbf{b}}^{k+1} = \arg\min_{\mathbf{b}\in\mathcal{W}}\hat{h}(\mathbf{b};\mathbf{b}^k)$, we obtain

$$
\begin{aligned}
&-\nabla h(\mathbf{b}^k)^{\mathsf{T}}\left(\mathbf{b}^\star-\mathbf{b}^k\right) \\
&\quad = \nabla\hat{h}(\mathbf{b}^k;\mathbf{b}^k)^{\mathsf{T}}\left(\mathbf{b}^k-\mathbf{b}^\star\right) \\
&\quad = \nabla\hat{h}(\widehat{\mathbf{b}}^{k+1};\mathbf{b}^k)^{\mathsf{T}}\left(\mathbf{b}^k-\widehat{\mathbf{b}}^{k+1}+\widehat{\mathbf{b}}^{k+1}-\mathbf{b}^\star\right) \\
&\qquad + \left\langle\nabla\hat{h}(\mathbf{b}^k;\mathbf{b}^k)-\nabla\hat{h}(\widehat{\mathbf{b}}^{k+1};\mathbf{b}^k),\mathbf{b}^k-\mathbf{b}^\star\right\rangle
\end{aligned}
$$

$$
\begin{aligned}
&\leq \nabla \hat{h}(\widehat{\mathbf{b}}^{k+1};\mathbf{b}^k)^{\mathsf{T}} \left(\mathbf{b}^k - \widehat{\mathbf{b}}^{k+1}\right) \\
&\quad + \left\langle \nabla \hat{h}(\mathbf{b}^k;\mathbf{b}^k) - \nabla \hat{h}(\widehat{\mathbf{b}}^{k+1};\mathbf{b}^k), \mathbf{b}^k - \mathbf{b}^\star \right\rangle \\
&\leq Q \left\| \widehat{\mathbf{b}}^{k+1} - \mathbf{b}^k \right\|_2 + \left(\beta + \lambda_{\max}^{(\mathbf{L}_1)}\right) \\
&\quad \times \left\| \widehat{\mathbf{b}}^{k+1} - \mathbf{b}^k \right\|_2 \left\| \mathbf{b}^k - \mathbf{b}^\star \right\|_2 \\
&\leq \left(Q + B\left(\beta + \lambda_{\max}^{(\mathbf{L}_1)}\right)\right) \left\| \widehat{\mathbf{b}}^{k+1} - \mathbf{b}^k \right\|_2,
\end{aligned}
$$

where the second inequality is followed by Cauchy–Schwarz inequality and $\hat{h}(\mathbf{b};\mathbf{b}^k)$ is $\left(\beta + \lambda_{\max}^{(\mathbf{L}_1)}\right)$-smooth. Similarly, we have

$$
\nabla h\left(\mathbf{b}^k\right)^{\mathsf{T}} \left(\widehat{\mathbf{b}}^{k+1} - \mathbf{b}^k\right) \leq \left(\beta + \lambda_{\max}^{(\mathbf{L}_1)}\right) \left\| \widehat{\mathbf{b}}^{k+1} - \mathbf{b}^k \right\|_2^2.
$$

Summarizing these results and considering $\gamma^k < 1$, we conclude that

$$
h(\mathbf{b}^{k+1}) - h(\mathbf{b}^\star) \leq Z_{\mathrm{c}} \left\| \widehat{\mathbf{b}}^{k+1} - \mathbf{b}^k \right\|_2, \tag{33}
$$

where $Z_{\mathrm{c}} = Q + 2B\left(\beta + \lambda_{\max}^{(\mathbf{L}_1)} + L/4\right)$. Let $\mathbf{b}_{t+1}$ denote the optimal point for $h_t(\mathbf{b})$ at t-th iteration. We have:

$$
\begin{aligned}
\left\| \mathbf{b}_{t+1}^{k+1} - \mathbf{b}_{t+1} \right\|_2^2 &\leq \frac{2}{\beta}\left(h_t(\mathbf{b}_{t+1}^{k+1}) - h_t(\mathbf{b}_{t+1})\right) \\
&\leq \frac{2}{\beta} Z_{\mathrm{c}} \left\| \widehat{\mathbf{b}}_{t+1}^{k+1} - \mathbf{b}_{t+1}^{k} \right\|_2.
\end{aligned}
$$

Therefore, by setting the stopping criterion as $\left\| \widehat{\mathbf{b}}_{t+1}^{k} - \mathbf{b}_{t+1}^{k-1} \right\|_2 \leq C_{\mathrm{th}}/T^2$ with C_{th} being a predefined constant, we have:

$$
\sum_{t=1}^{T} \left\| \mathbf{b}_t^k - \mathbf{b}_t \right\|_2 \leq C_{\mathrm{e}},
$$

where $C_{\mathrm{e}} = \sqrt{\frac{2Z_{\mathrm{c}}C_{\mathrm{th}}}{\beta}}$.

REFERENCES

[1] J. Kelly, “A new interpretation of information rate,” *Bell Syst. Tech. J.*, vol. 35, no. 4, pp. 917–926, 1956.

[2] T. M. Cover, “Universal portfolios,” *Math. Finance*, vol. 1, no. 1, pp. 1–29, Jan. 1991.

[3] T. Orton, “An introduction to regret minimization in algorithmic trading: A survey of universal portfolio techniques,” 2021, *arXiv:2105.13126*.

[4] T. M. Cover and E. Ordentlich, “Universal portfolios with side information,” *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 348–363, Mar. 1996.

[5] A. Blum and A. Kalai, “Universal portfolios with and without transaction costs,” *Mach. Learn.*, vol. 35, no. 3, pp. 193–205, 1999.

[6] A. T. Kalai and S. Vempala, “Efficient algorithms for universal portfolios,” *J. Mach. Learn. Res.*, vol. 3, pp. 423–440, Mar. 2003.

[7] D. Applegate and R. Kannan, “Sampling and integration of near log-concave functions,” in *Proc. 23rd Annu. ACM Symp. Theory Comput.*, 1991, pp. 156–163.

[8] A. Frieze and R. Kannan, “Log-Sobolev inequalities and sampling from log-concave distributions,” *Ann. Appl. Probability*, vol. 9, no. 1, pp. 14–26, 1999.

[9] D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth, “On-line portfolio selection using multiplicative updates,” *Math. Finance*, vol. 8, no. 4, pp. 325–347, 1998.

[10] A. Agarwal and E. Hazan, “New algorithms for repeated play and universal portfolio management,” Princeton University, Technical Report TR-740-05, Tech. Rep., 2005.

[11] M. Zinkevich, “Online convex programming and generalized infinitesimal gradient ascent,” in *Proc. 20th Int. Conf. Mach. Learn.*, 2003, pp. 928–936.

[12] A. Agarwal, E. Hazan, S. Kale, and R. E. Schapire, “Algorithms for portfolio management based on the Newton method,” in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 9–16.

[13] E. Hazan, A. Agarwal, and S. Kale, “Logarithmic regret algorithms for online convex optimization,” *Mach. Learn.*, vol. 69, no. 2, pp. 169–192, 2007.

[14] E. Hazan and S. Kale, “An online portfolio selection algorithm with regret logarithmic in price variation,” *Math. Finance*, vol. 25, no. 2, pp. 288–310, 2012.

[15] A. Rakhlin and K. Sridharan, “Online learning with predictable sequences,” in *Proc. Conf. Learn. Theory*, PMLR, 2013, pp. 993–1019.

[16] A. Rakhlin and K. Sridharan, “Optimization, learning, and games with predictable sequences,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26, 2013, pp. 3066–3074.

[17] E. Hazan and C. Seshadhri, “Adaptive algorithms for online decision problems,” ECCC Tech. Rep. TR07-088, Sep. 2007. [Online]. Available: https://eccc.weizmann.ac.il/eccc-reports/2007/TR07-088/index.html

[18] Y. Freund, R. E. Schapire, Y. Singer, and M. K. Warmuth, “Using and combining predictors that specialize,” in *Proc. 29th Annu. ACM Symp. Theory Comput.*, 1997, pp. 334–343.

[19] M. Herbster and M. K. Warmuth, “Tracking the best expert,” *Mach. Learn.*, vol. 32, no. 2, pp. 151–178, 1998.

[20] E. Lehrer, “A wide range no-regret theorem,” *Games Econ. Behav.*, vol. 42, no. 1, pp. 101–115, 2003.

[21] A. Blum and Y. Mansour, “From external to internal regret.” *J. Mach. Learn. Res.*, vol. 8, no. 6, pp. 1307–1324, 2007.

[22] E. Hazan and C. Seshadhri, “Efficient learning algorithms for changing environments,” in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 393–400.

[23] A. Daniely, A. Gonen, and S. Shalev-Shwartz, “Strongly adaptive online learning,” in *Proc. Int. Conf. Mach. Learn.,* PMLR, 2015, pp. 1405–1411.

[24] K.-S. Jun, F. Orabona, S. Wright, and R. Willett, “Improved strongly adaptive online learning using coin betting,” in *Proc. Artif. Intell. Statist.,* PMLR, 2017, pp. 943–951.

[25] S. Sachs, H. Hadiji, T. van Erven, and C. Guzmán, “Between stochastic and adversarial online convex optimization: Improved regret bounds via smoothness,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 691–702.

[26] S. Chen, W.-W. Tu, P. Zhao, and L. Zhang, “Optimistic online mirror descent for bridging stochastic and adversarial online convex optimization,” in *Proc. Int. Conf. Mach. Learn.*, 2023, pp. 691–702.

[27] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, “Decomposition by partial linearization: Parallel optimization of multi-agent systems,” *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[28] H. Markowitz, “Portfolio selection,” *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[29] S. Shalev-Shwartz et al., “Online learning and online convex optimization,” *Found. Trends® Mach. Learn.*, vol. 4, no. 2, pp. 107–194, 2012.

[30] F. Orabona, “A modern introduction to online learning,” 2021, *arXiv:1912.13213*.

[31] K.-S. Jun, F. Orabona, S. Wright, and R. Willett, “Online learning for changing environments using coin betting,” *Electron. J. Statist.*, vol. 11, no. 2, pp. 5282–5310, 2017.

[32] L. Zhang, T.-Y. Liu, and Z.-H. Zhou, “Adaptive regret of convex and smooth functions,” in *Proc. Int. Conf. Mach. Learn.,* PMLR, 2019, pp. 7414–7423.

[33] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. Cambridge, U.K.: Cambridge University Press, 2006.

[34] E. Ordentlich and T. M. Cover, “The cost of achieving the best portfolio in hindsight,” *Math. Oper. Res.*, vol. 23, no. 4, pp. 960–982, 1998.

[35] J. Abernethy, P. L. Bartlett, A. Rakhlin, and A. Tewari, “Optimal strategies and minimax lower bounds for online convex games,” in *Proc. 21st Annu. Conf. Learn. Theory*, 2008, pp. 414–424.

[36] A. Nemirovski, “Interior point polynomial time methods in convex programming,” *Lecture Notes*, vol. 42, no. 16, pp. 3215–3224, 2004.

[37] Y. Sun, P. Babu, and D. P. Palomar, “Majorization-minimization algorithms in signal processing, communications, machine learning,” *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[38] G. Scutari and Y. Sun, "Parallel and distributed successive convex approximation methods for big-data optimization," *Lecture Notes in Mathematics,* C.I.M.E, Springer Verlag series, 2018.

[39] J. Song, P. Babu, and D. P. Palomar, "Optimization methods for designing sequences with low autocorrelation sidelobes," *IEEE Trans. Signal Process.*, vol. 63, no. 15, pp. 3998–4009, Aug. 2015.

[40] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 686–695, Feb. 2005.

[41] A. Meucci, *Risk and Asset Allocation,* vol. 1. Amsterdam, Netherlands: Elsevier, 2005.

[42] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *J. Multivariate Anal.*, vol. 88, no. 2, pp. 365–411, 2004.

[43] O. Ledoit and M. Wolf, "Analytical nonlinear shrinkage of large-dimensional covariance matrices," *Ann. Statist.*, vol. 48, no. 5, pp. 3043–3065, 2020.

[44] Y. C. Eldar et al., "Rethinking biased estimation: Improving maximum likelihood and the Cramér–Rao bound," *Found. Trends® Signal Process.*, vol. 1, no. 4, pp. 305–449, 2008.

[45] Y. Feng et al., "A signal processing perspective on financial engineering," *Found. Trends® Signal Process.*, vol. 9, nos. 1–2, pp. 1–231, 2016.

[46] R. Pardo, *The Evaluation and Optimization of Trading Strategies,* vol. 314. Hoboken, NJ, USA: Wiley, 2011.

[47] D. P. Palomar and R. Zhou, *portfolioBacktest: Automated Backtesting of Portfolios Over Multiple Datasets*. 2021, R package version 0.3.1. [Online]. Available: https://CRAN.R-project.org/package=portfolioBacktest

[48] W. F. Sharpe, "Mutual fund performance," *J. Bus.*, vol. 39, no. 1, pp. 119–138, 1966.

[49] S. R. Putta and S. Agrawal, "Regret bounds for optimistic follow the leader: Applications in portfolio selection and linear regression," in *Proc. OPT : Optim. Mach. Learn.*, 2023.

[50] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?" *The Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953, 2009.

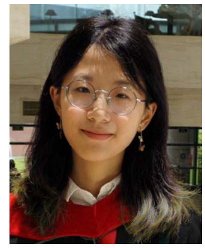
[51] A. Fu, B. Narasimhan, and S. Boyd, "CVXR: An R package for disciplined convex optimization," *J. Statist. Softw.*, vol. 94, no. 14, pp. 1–34, 2020.

[52] B. A. Turlach and A. Weingessel, *quadprog: Functions to Solve Quadratic Programming Problems*. 2019, R package version 1.5-8. [Online]. Available: https://CRAN.R-project.org/package=quadprog

[53] O. Anava, E. Hazan, S. Mannor, and O. Shamir, "Online learning for time series prediction," in *Proc. Conf. Learn. Theory*, PMLR, 2013, pp. 172–184.

[54] C. Liu, S. C. Hoi, P. Zhao, and J. Sun, "Online ARNIMA algorithms for time series prediction," in *Proc. AAAI Conf. Artif. Intell.*, vol. 30, no. 1, 2016, pp. 1867–1873.

[55] A. Tsiamis and G. J. Pappas, "Online learning of the Kalman filter with logarithmic regret," *IEEE Trans. Autom. Control*, vol. 68, no. 5, pp. 2774–2789, May 2023.



Runhao Shi received the bachelor's degree in software engineering from Sun Yat-sen University, Guangzhou, China, in 2021. She is currently working toward the Ph.D. degree in electronic and computer engineering with Hong Kong University of Science and Technology. Her research interests include online convex optimization, portfolio optimization, and financial engineering.

Daniel P. Palomar (Fellow, IEEE) received the Electrical Engineering and Ph.D. degrees from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively. He was a Fulbright Scholar with Princeton University from 2004 to 2006. In 2006, he joined as a Professor with the Department of Electronic & Computer Engineering and in the Department of Industrial Engineering & Decision Analytics, Hong Kong University of Science and Technology (HKUST), Hong Kong. He had previously held several research appointments, namely, at King's College London (KCL), London, U.K.; Stanford University, Stanford, CA; Telecommunications Technological Center of Catalonia (CTTC), Barcelona, Spain; Royal Institute of Technology (KTH), Stockholm, Sweden; University of Rome "La Sapienza", Rome, Italy; and Princeton University, Princeton, NJ. His research interests include data analytics, optimization methods, and deep learning in financial systems. He is a EURASIP Fellow, a recipient of a 2004/06 Fulbright Research Fellowship, the 2004, 2015, and 2020 (co-author) Young Author Best Paper Awards by the IEEE Signal Processing Society, the 2015–16 HKUST Excellence Research Award, the 2002/03 best Ph.D. prize in Information Technologies and Communications by the Technical University of Catalonia (UPC), the 2002/03 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by Epson Foundation, and the 2004 prize for the best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT. He has been a Guest Editor of IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING 2016 Special Issue on Financial Signal Processing and Machine Learning for Electronic Trading, an Associate Editor of IEEE TRANSACTIONS ON INFORMATION THEORY and IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of *IEEE Signal Processing Magazine* 2010 Special Issue on Convex Optimization for Signal Processing, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2008 Special Issue on Game Theory in Communication Systems, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2007 Special Issue on Optimization of MIMO Transceivers for Realistic Communication Networks.