SKETCH-AWARE GENERALIZED LINEAR BANDITS: EFFICIENT ONE-PASS UPDATES VIA FREQUENT DIRECTIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

The nonlinear link function in generalized linear bandits (GLBs) makes it challenging to achieve statistical and computational efficiency simultaneously. Recent online mirror descent methods enable statistically efficient one-pass learning, but their guarantees rely on dense curvature geometry, incurring $O(d^2)$ memory and $O(d^3)$ computation per round in the feature dimension d. We propose a sketch-aware GLBs algorithm based on Frequent Directions, a deterministic sketch that incrementally compresses the adaptively revealed curvature geometry. A single sketch supports both estimator updates and optimistic action selection, reducing memory to $O(md)$ and, for small candidate action sets, amortized per-round computation to $O(md)$ for sketch size $m < d$. We establish a rigorous regret guarantee that quantifies the efficiency–utility trade-off: the approximation penalty is controlled by the discarded spectral tail, while the regret retains the statistical efficiency of its full-matrix counterpart under favorable spectral structure. Across synthetic and real-world benchmarks, our approach closely matches the regret of full-matrix methods while substantially reducing runtime and peak memory.

1 INTRODUCTION

Multi-armed bandits (MAB) provide a general framework for sequential decision-making under partial feedback (Herbert, 1952) and have been widely used in applications such as online recommendation (Li et al., 2010), health surveillance (Bastani et al., 2021), and green security (Xu et al., 2021). In this paper, we study generalized linear bandits (GLBs), an important extension of MAB in which the expected reward follows a generalized linear model (GLM) (McCullagh & Nelder, 1989). Specifically, at round t, the learner selects an action $\boldsymbol{x}_t$ from a time-varying action set $\mathcal{X}_t \subseteq \mathbb{R}^d$ and then observes a stochastic reward $r_t \in \mathbb{R}$ satisfying $\mathbb{E}[r_t \mid \boldsymbol{x}_t] = \mu(\boldsymbol{x}_t^\top \boldsymbol{\theta}_\star)$, where $\mu(\cdot) : \mathbb{R} \to \mathbb{R}$ is a known link function and $\boldsymbol{\theta}_\star \in \mathbb{R}^d$ is an unknown parameter. The objective of the learner is to maximize the cumulative expected reward over the total T rounds. Compared with classical linear bandits (Dani et al., 2007; Abbasi-Yadkori et al., 2011), GLBs capture a broader family of reward models and naturally accommodate distributions such as Gaussian, Bernoulli, and Poisson.

A central challenge in GLBs is to attain statistical and computational efficiency simultaneously, as the nonlinear link function complicates both parameter estimation and confidence construction. The seminal work of Filippi et al. (2010) first formalized the GLBs problem and proposed GLM-UCB, a low-regret algorithm that estimates the unknown parameter $\boldsymbol{\theta}_\star$ via maximum likelihood estimation (MLE) on the historical data $\{(\boldsymbol{x}_s, r_s)\}_{s=1}^{t-1}$ at round t. This yields a nearly optimal regret bound of $\widetilde{O}(\kappa\sqrt{T})$, where the $\widetilde{O}(\cdot)$ notation hides logarithmic factors, and $\kappa = 1/\inf_{\boldsymbol{x}\in\cup_{t=1}^{T}\mathcal{X}_t,\boldsymbol{\theta}\in\Theta} \mu'(\boldsymbol{x}^\top\boldsymbol{\theta})$ is a potentially large problem-dependent constant. Subsequent works (Sawarni et al., 2024; Lee et al., 2024; Liu et al., 2024; Clerico et al., 2025), building on self-concordant analyses of the loss, decouple the regret from this worst-case constant and improve the regret bound to $\widetilde{O}(\sqrt{T/\kappa_\star})$, where $\kappa_\star \leq \kappa$ depends on the local slope at the optimal actions. Despite their statistical efficiency, these methods remain computationally expensive because they rely on MLE over the entire history: they must retain all past data and repeatedly solve an optimization problem over an ever-growing dataset, causing complexity to increase with the horizon t. Recently, Zhang et al. (2025) removed this horizon-dependent bottleneck by replacing MLE with the online mirror descent (OMD) estima-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Table 1: Comparison of GLBs algorithms in terms of regret bounds and computational complexity. Here, $\kappa_\star = T/\sum_{t=1}^{T} \mu'(\boldsymbol{x}_t^{\star\top}\boldsymbol{\theta}_\star)$ and $\boldsymbol{x}_t^\star = \arg\max_{\boldsymbol{x}\in\mathcal{X}_t} \mu(\boldsymbol{x}^\top\boldsymbol{\theta}_\star)$ is the optimal action at round t, with $\kappa_\star \le \kappa$. Define $\widetilde{m} := m + \varepsilon_m d$, where ε_m is the spectral approximation error defined in (3). † reports the leading-order regret term, ‡ indicates the amortized per-round time complexity.

Algorithm	**Regret Bound**†	**Estimator**	**Space**	**Time**‡
Full-history methods				
GLM-UCB (Filippi et al., 2010)	$\widetilde{O}(\kappa d\sqrt{T})$	MLE	$O(t)$	$O(t)$
OFUGLB (Lee et al., 2024)	$\widetilde{O}(d\sqrt{T/\kappa_\star})$	MLE	$O(t)$	$O(t)$
RS-GLinCB (Sawarni et al., 2024)	$\widetilde{O}(d\sqrt{T/\kappa_\star})$	MLE	$O(t)$	$O((\log t)^2)$
One-pass methods				
GLOC (Jun et al., 2017)	$\widetilde{O}(\kappa d\sqrt{T})$	ONS	$O(d^2)$	$O(d^2)$
GLB-OMD (Zhang et al., 2025)	$\widetilde{O}(d\sqrt{T/\kappa_\star})$	OMD	$O(d^2)$	$O(d^3)$
SGLB (this work)	$\widetilde{O}(\sqrt{(1+\varepsilon_m)\, d\widetilde{m}T/\kappa_\star})$	OMD	$O(md)$	$O(md)$

tor for exploration. Their algorithm, GLB-OMD, is jointly efficient: it achieves the nearly optimal regret bound $\widetilde{O}(d\sqrt{T/\kappa_\star})$ while supporting one-pass updates, meaning that each data point is processed only once and past observations need not be stored.

However, one-pass updates alone do not make GLBs scalable in high dimensions. Although GLB-OMD enjoys strong theoretical guarantees and empirical performance, it requires $O(d^2)$ memory and $O(d^3)$ amortized time per round, where d is the feature dimension. The cubic term comes from solving the OMD update over the full Hessian. Such dependence can be prohibitive in large-scale applications with high-dimensional contexts, including online advertising (Zhang et al., 2022), medical decision-making (Bastani & Bayati, 2020), and dynamic pricing (Javanmard & Nazerzadeh, 2019). As summarized in Table 1, existing works on GLBs (Jun et al., 2017; Zhang et al., 2025) have mainly focused on achieving *horizon-efficient* and one-pass learning. A natural next question is whether one can further design *dimension-efficient* GLBs algorithms, while still preserving rigorous guarantees. This is nontrivial because existing one-pass GLB analyses are built on full-matrix second-order geometry, which underlies both parameter update and action selection.

Main Results. To address this gap, we develop a sketch-aware algorithm for GLBs based on Frequent Directions (FD) (Liberty, 2013), a deterministic matrix sketching technique. Our method compresses the full-matrix OMD geometry into an m-dimensional sketch, reducing both the per-round time and the memory cost to $O(md)$. On the statistical side, our algorithm preserves the κ-free confidence-regret structure achieved by recent self-concordant analyses (Russac et al., 2021; Lee et al., 2024; Sawarni et al., 2024; Zhang et al., 2025). Specifically, the dimension-dependent term in the confidence radius is reduced from d to m, up to an additive sketching penalty of $d\varepsilon_m$, which yields the regret bound $\widetilde{O}(\sqrt{(1+\varepsilon_m)\, d\widetilde{m}T/\kappa_\star})$, where $\widetilde{m} := m+\varepsilon_m d$ and ε_m is the spectral error. Moreover, when the curvature-weighted actions effectively lie in an m-dimensional subspace, FD incurs no sketching error and the regret becomes $\widetilde{O}(d\sqrt{T})$, matching the order of the slower non-sketched counterpart. Our result therefore shows that one-pass GLBs need not be inherently tied to full-matrix computation: they can be made dimension-efficient and statistically controlled.

Technical Contribution. Our main technical contribution is to show that matrix sketching can be integrated into the confidence–regret analysis of one-pass GLBs without destroying the statistical structure that underlies optimistic exploration. This is far from immediate. First, the object to be compressed is not a plain covariance matrix, but a curvature-weighted Hessian geometry that simultaneously governs the OMD update, the confidence norm, and the OFU bonus. Second, unlike linear least-squares problems, OMD-based GLBs require solving a parametric projection whose $O(d^3)$ cost dominates the update when the action set is small. We show that this projection can be carried out in sketched form, yielding an efficient, low-rank optimization problem. Finally, our analysis leverages the notion of mix loss (Vovk, 2001; Kirschner et al., 2025), previously used in full-information online learning to obtain fast regret rates, to quantify the discrepancy between the FD-based OMD estimator and the true parameter $\boldsymbol{\theta}_\star$. This discrepancy can be expressed as an explicit sketching penalty, which in turn enables tight confidence sets and a κ-free regret guarantee.

2 PRELIMINARY

This section introduces the GLBs problem and reviews the prior work most closely related to ours.

2.1 GENERALIZED LINEAR BANDITS

We begin by introducing the problem formulation of generalized linear bandits, where a learner interacts with the environment over T rounds. At each round $t \in [T]$, the learner selects an action $\boldsymbol{x}_t \in \mathcal{X}_t \subseteq \mathbb{R}^d$ and then observes a stochastic reward $r_t \in \mathbb{R}$. The finite action set $\mathcal{X}_t$ is allowed to vary with t, capturing practical settings in which the available actions change over time. We denote by $\mathcal{F}_t := \sigma(\boldsymbol{x}_1, r_1, \ldots, \boldsymbol{x}_{t-1}, r_{t-1})$ the filtration generated by the past observations up to round t, and let $\boldsymbol{\theta}_\star \in \mathbb{R}^d$ be the unknown parameter. Conditioned on $\mathcal{F}_t$, the reward r_t follows a canonical exponential-family model with natural parameter $z_t := \boldsymbol{x}_t^\top \boldsymbol{\theta}_\star$, namely

$$\Pr[r_t \mid z_t, \mathcal{F}_t] = \exp\left((r_t z_t - m(z_t))/g(\tau) + h(r_t, \tau)\right),$$

where τ is a known scale parameter, $g(\tau)$ is the dispersion function, $h(r_t, \tau)$ is the base measure, and $m(z)$ is a twice continuously differentiable log-partition function. Standard properties of generalized linear models (McCullagh & Nelder, 1989) imply that $\mathbb{E}[r_t \mid z_t] = \mu(z_t) := m'(z_t)$ and $\mathrm{Var}(r_t \mid z_t) = g(\tau)m''(z_t)$. This framework covers many common reward models, including Gaussian ($\mu(z) = z$), Poisson ($\mu(z) = e^z$), and Bernoulli ($\mu(z) = (1 + e^{-z})^{-1}$).

The learner aims to minimize the regret $\mathrm{Regret}_T = \sum_{t=1}^T \mu(\boldsymbol{x}_t^{\star\top} \boldsymbol{\theta}_\star) - \sum_{t=1}^T \mu(\boldsymbol{x}_t^\top \boldsymbol{\theta}_\star)$, where $\boldsymbol{x}_t^\star = \arg\max_{\boldsymbol{x} \in \mathcal{X}_t} \mu(\boldsymbol{x}^\top \boldsymbol{\theta}_\star)$ denotes the optimal action at round t. Following prior works (Lee et al., 2024; Sawarni et al., 2024; Zhang et al., 2025), we make the following assumptions.

Assumption 1 (Bounded Domain)**.** For all $\boldsymbol{x} \in \mathcal{X}_t$ and $t \in [T]$, we assume $\|\boldsymbol{x}\|_2 \le 1$. We also assume that the unknown parameter satisfies $\|\boldsymbol{\theta}_\star\|_2 \le S$ for constant $S > 0$ known to the learner.

Assumption 2 (Bounded Slope)**.** The link function μ is twice differentiable over its feasible domain. Moreover, there exist constants $c_\mu > 0$ and $L_\mu > 0$ such that $c_\mu \le \mu'(z) \le L_\mu$ for all $z \in [-S, S]$. Consequently, m is strictly convex and μ is strictly increasing on the feasible domain.

Assumption 3 (Self-Concordance)**.** The link function μ satisfies $|\mu''(z)| \le R\mu'(z)$, for all $z \in \mathbb{R}$.

2.2 ONE-PASS GLBS AND GLB-OMD (ZHANG ET AL., 2025)

As summarized in Table 1, a central challenge in GLBs is to maintain an estimator of the unknown $\boldsymbol{\theta}_\star$ that is both statistically and computationally efficient. This estimator is then used to construct the optimistic selection rule. A standard choice is the regularized maximum likelihood estimator,

$$\boldsymbol{\theta}_{t+1}^{\texttt{MLE}} = \arg\min_{\boldsymbol{\theta} \in \Theta} \sum_{s=1}^{t} \ell_s(\boldsymbol{\theta}) + \lambda \|\boldsymbol{\theta}\|_2^2 \quad \text{and} \quad \ell_s(\boldsymbol{\theta}) := (m(\boldsymbol{x}_s^\top \boldsymbol{\theta}) - r_s \boldsymbol{x}_s^\top \boldsymbol{\theta})/g(\tau), \tag{1}$$

where $\lambda > 0$ and Θ denotes a feasible domain.[1] *Full-history* methods rely on MLE estimators and have been shown to enjoy strong statistical guarantees (Filippi et al., 2010; Lee et al., 2024; Sawarni et al., 2024; Liu et al., 2024). However, computing (1) requires repeated access to the entire history $\{(\boldsymbol{x}_s, r_s)\}_{s=1}^{t-1}$. As a result, both the memory usage and the update cost grow with the horizon t.

One-pass GLBs aim to remove this horizon dependence by processing each observation only once, without storing past data. To the best of our knowledge[2], GLB-OMD (Zhang et al., 2025) is the only existing method that is both one-pass and statistically efficient in the sense of achieving a $\kappa_\star$-dependent regret guarantee. Its key idea is to replace MLE with an online mirror descent estimator,

$$\boldsymbol{\theta}_{t+1}^{\texttt{OMD}} = \arg\min_{\boldsymbol{\theta} \in \Theta} \widetilde{\ell}_t(\boldsymbol{\theta}) + \frac{1}{2\eta} \|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|_{\boldsymbol{H}_t}^2 \quad \text{and} \quad \widetilde{\ell}_t(\boldsymbol{\theta}) := \langle \nabla \ell_t(\boldsymbol{\theta}_t), \boldsymbol{\theta} - \boldsymbol{\theta}_t \rangle + \frac{1}{2} \|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|_{\nabla^2 \ell_t(\boldsymbol{\theta}_t)}^2, \tag{2}$$

where $\eta > 0$ is the step size and $\boldsymbol{H}_t = \lambda \boldsymbol{I}_d + \sum_{s=1}^{t-1} \nabla^2 \ell_s(\boldsymbol{\theta}_{s+1})$ is the accumulated local Hessian. Because $\boldsymbol{H}_t$ admits incremental rank-one updates, GLB-OMD supports one-pass, horizon-independent computation and avoids storing the full history. However, GLB-OMD still relies on a full $d \times d$ Hessian geometry. Consequently, it requires $O(d^2)$ memory and $O(d^3)$ time per round, which becomes prohibitive in medium and high dimensions.

[1] Different prior works instantiate the feasible domain differently; we use Θ as a generic notation here.

[2] GLOC (Jun et al., 2017) also admits a one-pass implementation, but does not enjoy a comparably regret.

2.3 Frequent Directions (Ghashami et al., 2016)

Frequent Directions (FD) is a deterministic matrix sketching method for streaming data. Given a streaming matrix $\boldsymbol{A}_t = [\boldsymbol{a}_1^\top, \dots, \boldsymbol{a}_t^\top]^\top \in \mathbb{R}^{t\times d}$, $t \in [T]$, FD maintains a sketch matrix $\boldsymbol{B}_t \in \mathbb{R}^{m\times d}$ with sketch size $m < d$, such that the covariance of $\boldsymbol{B}_t$ approximates that of $\boldsymbol{A}_t$. At a high level, upon receiving a new row $\boldsymbol{a}_t$, FD inserts it into the sketch, computes the singular value decomposition of the current sketch, and then shrinks the singular values so as to retain only the most informative directions. We will provide the pseudo-code of FD in Section 3 later.

FD is well suited to streaming settings: it uses $O(md)$ memory and admits an amortized update time of $O(md)$. Besides, FD provides a spectral approximation guarantee for the covariance matrix of the data stream. The following lemma states the property that will be used throughout our analysis.

Lemma 1 (Theorem 1.3 of Ghashami et al. (2016))**.** *Let $\boldsymbol{A}_T \in \mathbb{R}^{T\times d}$ have singular values $\sigma_1 \geq \cdots \geq \sigma_d \geq 0$, and let $\boldsymbol{B}_T \in \mathbb{R}^{m\times d}$ be the sketch produced by Algorithm 1 and $\lambda > 0$. Then*

$$\frac{\left\|\boldsymbol{A}_T^\top \boldsymbol{A}_T - \boldsymbol{B}_T^\top \boldsymbol{B}_T\right\|_2}{\lambda} \leq \varepsilon_m, \quad \text{where} \quad \varepsilon_m := \min_{0\leq k<m} \frac{\sum_{i=k+1}^{d} \sigma_i^2}{\lambda(m-k)}. \tag{3}$$

3 Sketch-Aware Generalized Linear Bandits

In this section, we present a **S**ketch-aware **G**eneralized **L**inear **B**andit algorithm, termed SGLB. Our goal is to retain the one-pass nature while reducing its dependence on d through matrix sketching. We begin by making explicit the dimension-dependent computational bottleneck in GLB-OMD.

3.1 Dimension-Dependent Computational Bottleneck

Our starting point is GLB-OMD (Zhang et al., 2025), the most efficient existing one-pass method for GLBs. Following the principle of optimism in the face of uncertainty (OFU) (Agrawal, 1995; Auer, 2002), GLB-OMD maintains at each round t a confidence set $\mathcal{C}_t(\delta) \subseteq \mathbb{R}^d$ that contains the true parameter $\boldsymbol{\theta}_\star$ with high probability. The learner then selects an action by maximizing the optimistic reward over $\mathcal{C}_t(\delta)$. Since μ is increasing and $\mathcal{C}_t(\delta)$ is ellipsoidal, this rule reduces to

$$\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}_t} \{\boldsymbol{x}^\top \boldsymbol{\theta}_t^{\texttt{OMD}} + \beta_t(\delta)\|\boldsymbol{x}\|_{\boldsymbol{H}_t^{-1}}\}, \tag{4}$$

where $\beta_t(\delta) > 0$ is the radius of $\mathcal{C}_t(\delta)$ and $\boldsymbol{H}_t = \lambda \boldsymbol{I}_d + \sum_{s=1}^{t-1} \nabla^2 \ell_s(\boldsymbol{\theta}_{s+1}^{\texttt{OMD}})$ is the accumulated local Hessian. Let $\overline{\boldsymbol{H}}_{t-1} = \boldsymbol{H}_{t-1} + \eta \nabla^2 \ell_{t-1}(\boldsymbol{\theta}_{t-1}^{\texttt{OMD}})$. In the above, $\boldsymbol{\theta}_t^{\texttt{OMD}}$ denotes the OMD estimator updated by (2). As shown in Zhang et al. (2025), it admits the equivalent two-step form

$$\boldsymbol{\xi}_t = \boldsymbol{\theta}_{t-1}^{\texttt{OMD}} - \eta \overline{\boldsymbol{H}}_{t-1}^{-1} \nabla \ell_{t-1}(\boldsymbol{\theta}_{t-1}^{\texttt{OMD}}) \quad \text{and} \quad \boldsymbol{\theta}_t^{\texttt{OMD}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \|\boldsymbol{\theta} - \boldsymbol{\xi}_t\|^2_{\overline{\boldsymbol{H}}_{t-1}}. \tag{5}$$

In memory, GLB-OMD needs to maintain the estimator $\boldsymbol{\theta}_t^{\texttt{OMD}}$, the gradient $\nabla \ell_t(\boldsymbol{\theta}_t^{\texttt{OMD}})$, and the inverse matrices $\boldsymbol{H}_t^{-1}$ and $\overline{\boldsymbol{H}}_t^{-1}$, which requires $O(d^2)$ space. In per-round computation, the cost comes from two sources: updating the OMD estimator and evaluating the OFU bonus in (4). For the update in (5), the first step is a gradient update and can be implemented in $O(d^2)$ time via the Sherman–Morrison formula (Sherman & Morrison, 1950), since both $\boldsymbol{H}_t^{-1}$ and $\overline{\boldsymbol{H}}_t^{-1}$ admit incremental rank-one updates. The second step is a projection onto Θ under the metric induced by $\overline{\boldsymbol{H}}_{t-1}$, which can be solved in $O(d^3)$ time using the procedure of Section 4.1 in Mhammedi et al. (2019). For action selection in (4), evaluating the bonus term $\|\boldsymbol{x}\|_{\boldsymbol{H}_t^{-1}}$ costs $O(d^2)$ for each candidate action. Therefore, although GLB-OMD eliminates horizon dependence, it still incurs $O(d^2)$ space and $O(d^3 + |\mathcal{X}_t|d^2)$ per-round time, which can be prohibitive in moderate to high dimensions.

3.2 Proposed Method

In this section, we demonstrate that the one-pass OMD framework can be further transformed into a sketch-aware framework, thereby simultaneously achieving computational improvements in dimensionality and minimizing statistical costs. We first design the geometry to be sketched.

Algorithm 1 One-pass FD Update

Input: Current sketch matrices $(\boldsymbol{B}_t, \boldsymbol{G}_t)$, newly arrived row $\boldsymbol{a}_t \in \mathbb{R}^d$, and $\lambda > 0$

1: Append $\boldsymbol{a}_t$ to the last row of $\boldsymbol{B}_t$
2: Compute $[\boldsymbol{U}, \boldsymbol{\Sigma}, \boldsymbol{V}] \leftarrow \mathrm{SVD}(\boldsymbol{B}_t)$
3: $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \cdots, \sigma_m)$
4: Set $\rho \leftarrow \sigma_m^2$
5: Update sketch matrix $\boldsymbol{B}_t$ by

$$\boldsymbol{B}_{t+1} \leftarrow \sqrt{\boldsymbol{\Sigma}^2 - \rho \boldsymbol{I}_m} \cdot \boldsymbol{V}^\top$$

6: Update sketch matrix $\boldsymbol{G}_t$ by

$$\boldsymbol{G}_{t+1} \leftarrow \mathrm{diag}\left\{ \frac{1}{\lambda + \sigma_1^2 - \rho}, \cdots, \frac{1}{\lambda} \right\}$$

Output: Sketch matrices $\boldsymbol{B}_{t+1}, \boldsymbol{G}_{t+1}$.

Algorithm 2 SGLB

Input: Constant $R, L_\mu, S, \lambda > 0$, $\delta \in (0, 1)$, sketch size $m \in \{1, \ldots, d-1\}$.

1: Initialize $\boldsymbol{\theta}_1 \in \Theta := \{\boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta}\|_2 \le S\}$.
2: Initialize sketch $\boldsymbol{B}_1 \leftarrow \boldsymbol{0}^{m \times d}$, $\boldsymbol{G}_1 \leftarrow \frac{1}{\lambda} \boldsymbol{I}_m$.
3: **for** $t = 1$ **to** T **do**
4: Compute the radius $\widetilde{\beta}_t(\delta)$ by (9).
5: Select action $\boldsymbol{x}_t \in \mathcal{X}_t$ according to rule (10).
6: Receive the reward r_t.
7: Update estimator $\boldsymbol{\theta}_{t+1}$ by (6).
8: Calculate $w_t = \mu'(\boldsymbol{x}_t^\top \boldsymbol{\theta}_{t+1})/g(\tau)$.
9: Update sketch matrices

$$\boldsymbol{B}_{t+1}, \boldsymbol{G}_{t+1} \leftarrow \text{Algorithm 1 } (\boldsymbol{B}_t, \boldsymbol{G}_t, \sqrt{w_t}\boldsymbol{x}_t)$$

10: **end for**

Output: Sketch-based estimator $\boldsymbol{\theta}_{T+1}$.

Sketched Hessian Geometry. Recall from Section 3.1 that the local Hessian in GLB-OMD takes the form $\boldsymbol{H}_t = \lambda \boldsymbol{I}_d + \sum_{s=1}^{t-1} w_s \boldsymbol{x}_s \boldsymbol{x}_s^\top$, where $w_s := \mu'(\boldsymbol{x}_s^\top \boldsymbol{\theta}_{s+1})/g(\tau)$. Let $\boldsymbol{A}_t \in \mathbb{R}^{(t-1) \times d}$ be the curvature-weighted action matrix whose s-th row is $\sqrt{w_s}\boldsymbol{x}_s^\top$, so that $\boldsymbol{H}_t = \lambda \boldsymbol{I}_d + \boldsymbol{A}_t^\top \boldsymbol{A}_t$. Instead of storing $\boldsymbol{A}_t$ or $\boldsymbol{H}_t$ explicitly, we maintain an FD sketch $\boldsymbol{B}_t \in \mathbb{R}^{m \times d}$ with $m \ll d$ and define the sketched Hessian $\widetilde{\boldsymbol{H}}_t := \lambda \boldsymbol{I}_d + \boldsymbol{B}_t^\top \boldsymbol{B}_t$. We presents the pseudo-code of FD updates in Algorithm 1.

Sketch-Based Estimator. Using this sketched geometry, we define the sketch-based estimator by

$$\boldsymbol{\theta}_{t+1} = \arg\min_{\boldsymbol{\theta} \in \Theta} \left\{ \langle \nabla \ell_t(\boldsymbol{\theta}_t), \boldsymbol{\theta} - \boldsymbol{\theta}_t \rangle + \frac{1}{2\eta} \|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|^2_{\widetilde{\boldsymbol{H}}_t + \eta \nabla^2 \ell_t(\boldsymbol{\theta}_t)} \right\} \tag{6}$$

where $\Theta := \{\boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta}\|_2 \le S\}$ with radius $S > 0$, η is the step size. Thus, we retain the same second-order surrogate for the current loss as in GLB-OMD, but replace the full accumulated Hessian by its FD approximation. The optimization problem in (6) is a quadratic program over a Euclidean ball and, following Orabona (2019), can be equivalently rewritten as

$$\boldsymbol{\xi}_{t+1} = \boldsymbol{\theta}_t - \eta \Big(\widetilde{\boldsymbol{H}}_t + \eta \nabla^2 \ell_t(\boldsymbol{\theta}_t) \Big)^{-1} \nabla \ell_t(\boldsymbol{\theta}_t) \ \text{ and } \ \boldsymbol{\theta}_{t+1} = \arg\min_{\boldsymbol{\theta} \in \Theta} \|\boldsymbol{\theta} - \boldsymbol{\xi}_{t+1}\|^2_{\widetilde{\boldsymbol{H}}_t + \eta \nabla^2 \ell_t(\boldsymbol{\theta}_t)} . \tag{7}$$

A naive substitution of $\boldsymbol{H}_t$ by $\widetilde{\boldsymbol{H}}_t$ is not yet sufficient; the computational gain comes from exploiting the low-rank structure in (7). Indeed, $\widetilde{\boldsymbol{H}}_t = \lambda \boldsymbol{I}_d + \boldsymbol{B}_t^\top \boldsymbol{B}_t$ is a rank-m correction of $\lambda \boldsymbol{I}_d$, while the term $\nabla^2 \ell_t(\boldsymbol{\theta}_t) = \mu'(\boldsymbol{x}_t^\top \boldsymbol{\theta}_t)/g(\tau) \boldsymbol{x}_t \boldsymbol{x}_t^\top$ is rank-1 matrix. As stated in Lemma 9, applying the Woodbury identity together with its rank-one special case, the Sherman–Morrison formula, yields

$$\Big(\widetilde{\boldsymbol{H}}_t + \eta \nabla^2 \ell_t(\boldsymbol{\theta}_t) \Big)^{-1} = \widetilde{\boldsymbol{H}}_t^{-1} - \frac{\widetilde{\boldsymbol{H}}_t^{-1} \boldsymbol{x}_t \boldsymbol{x}_t^\top \widetilde{\boldsymbol{H}}_t^{-1}}{\boldsymbol{x}_t^\top \widetilde{\boldsymbol{H}}_t^{-1} \boldsymbol{x}_t + g(\tau)/\big(\eta\, \mu'(\boldsymbol{x}_t^\top \boldsymbol{\theta}_t)\big)} \tag{8}$$

and $\widetilde{\boldsymbol{H}}_t^{-1} = \frac{1}{\lambda}(\boldsymbol{I}_d - \boldsymbol{B}_t^\top \boldsymbol{G}_t \boldsymbol{B}_t)$, where $\boldsymbol{G}_t := (\boldsymbol{B}_t \boldsymbol{B}_t^\top + \lambda \boldsymbol{I}_m)^{-1} \in \mathbb{R}^{m \times m}$. Therefore, there is no need to explicitly store the approximated Hessian matrix. Rather, it is sufficient to maintain only the sketch matrices $\boldsymbol{B}_t$ and $\boldsymbol{G}_t$, which can both be updated efficiently using Algorithm 1.

Using (8), the gradient step in (7) can be implemented using only matrix–vector multiplications involving $\boldsymbol{B}_t$ and $\boldsymbol{G}_t$, and therefore costs $O(md + m^2)$ time. Moreover, the same sketched geometry also simplifies the projection step in (7): instead of solving a full-matrix problem, it reduces to a low-dimensional computation together with a one-dimensional root-finding subproblem. As a result, the projection can be carried out in $O(md)$ time without performing an expensive full-matrix SVD; the details is deferred to Appendix E.1 due to space limitations.

Confidence Ellipsoid and Action Selection. We show that the sketch-based estimator in (6) admits a valid confidence set, which is the key ingredient for obtaining an OFU-style regret guarantee.

Theorem 1 (Confidence Ellipsoid for SGLB). *Under Assumptions 1–3, let $\delta \in (0,1]$, $\eta = 1 + RS$, and $\lambda = \max\{14d\eta R^2, 6\eta RSL_\mu/g(\tau)\}$. Then, for every round $t \in [T]$, with probability at least $1-\delta$, the true parameter $\boldsymbol{\theta}_\star$ belongs to $\widetilde{\mathcal{C}}_t(\delta) := \{\boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|_{\widetilde{\boldsymbol{H}}_t} \le \widetilde{\beta}_t(\delta)\}$, where $\bar{\rho}_t = \sum_{s=1}^{t} \rho_s$ denotes the sum of shrinking values in line 4 of Algorithm 1, and*

$$\widetilde{\beta}_t(\delta) = \sqrt{4\lambda S^2 + 2\eta \ln\frac{1}{\delta} + (6\eta^2+\eta)\, m \ln\Big(1 + \frac{L_\mu t}{m\lambda g(\tau)}\Big) + (6\eta^2+\eta)\,\frac{d\bar{\rho}_t}{\lambda}}. \tag{9}$$

Theorem 1 shows that SGLB retains an ellipsoidal confidence set despite replacing the full Hessian by its sketch. Compared with the non-sketched counterpart (Zhang et al., 2025, Theorem 1), the log term now scales with the sketch size m rather than d, up to the additional penalty $d\bar{\rho}_t/\lambda \le d\varepsilon_m$ by Lemma 1. In particular, when $\varepsilon_m \to 0$, the dominant growth term in (9) is of order $O(\sqrt{m \ln(t/\delta)})$.

Since $\widetilde{\mathcal{C}}_t(\delta)$ is ellipsoidal and μ is increasing, the corresponding OFU-based selection criterion is

$$\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}_t} \max_{\boldsymbol{\theta}_t \in \widetilde{\mathcal{C}}_t(\delta)} \boldsymbol{x}^\top \boldsymbol{\theta} = \arg\max_{\boldsymbol{x}\in\mathcal{X}_t} \Big\{ \boldsymbol{x}^\top \boldsymbol{\theta}_t + \widetilde{\beta}_t(\delta) \|\boldsymbol{x}\|_{\widetilde{\boldsymbol{H}}_t^{-1}} \Big\}. \tag{10}$$

Moreover, by the sketching formulation in (8), the bonus term $\|\boldsymbol{x}\|_{\widetilde{\boldsymbol{H}}_t^{-1}}$ can be evaluated without explicitly forming the approximated inverse Hessian $\widetilde{\boldsymbol{H}}_t^{-1}$, in $O(md + m^2)$ time per action.

Overall Complexity. The full procedure is summarized in Algorithm 2. Its amortized cost comes from three components: updating the FD sketch, computing the sketch-based OMD update in (7), and evaluating the OFU rule in (10). A direct FD compression step requires an $O(m^2 d)$ time for SVD, but this can be reduced to $O(md)$ amortized time either by the standard doubling-space implementation of FD (see Appendix E.2) or via the Gu–Eisenstat procedure (Gu & Eisenstat, 1993). By the discussion above, the sketch-based gradient step costs $O(md+m^2)$, the projection step costs $O(md)$, and each OFU bonus evaluation costs $O(md + m^2)$. Therefore, when the candidate action set is small, SGLB uses $O(md)$ space and admits an $O(md)$ amortized per-round complexity.

4 Analysis and Discussion

In this section, we present the regret guarantee of SGLB and discuss its connections to related topics.

4.1 Regret Analysis

Equipped with the sketched OMD estimator and the sketched confidence ellipsoid in Theorem 1, we now state the regret guarantee of SGLB in Algorithm 2. The regret bound is characterized by two sketch-aware quantities: the sketch size m and the spectral error ε_m at final round T, which is the bound of approximation error induced by matrix sketching, as shown in Lemma 1.

Theorem 2 (Regret Bound of SGLB). *Under Assumptions 1–3, let $\delta \in (0,1]$, $\eta = 1 + RS$, and $\lambda = \max\{14d\eta R^2, 6\eta RSL_\mu/g(\tau)\}$. With probability at least $1-\delta$, Algorithm 2 ensures*

$$\mathrm{Regret}_T \lesssim SR\,(1+\varepsilon_m)^{1/2} \sqrt{(dS^2R + \Phi_m)\,\Phi_m \frac{T}{\kappa_*}} + \kappa S^2 R^3\,(1+\varepsilon_m)^{3/2}(dS^2R + \Phi_m)\,\Phi_m,$$

where $\Phi_m := m\log T + d\varepsilon_m$ and $\lesssim$ hides constants independent of $d, m, \kappa, \kappa_*, S, R, T$.

The proof and a concrete upper bound are deferred to Appendix C. Theorem 2 shows that SGLB preserves the $\kappa_\star$-dependent regret structure of one-pass OMD-based GLBs, while replacing the full-dimensional logarithmic term by the sketch-dependent quantity Φ_m. Equivalently, using $\widetilde{m} := m + d\varepsilon_m$, the leading term can be summarized as $\widetilde{O}((1+\varepsilon_m)^{1/2}\sqrt{d\widetilde{m}T/\kappa_\star})$.

Compared with GLB-OMD (Zhang et al., 2025, Theorem 2), SGLB replaces the log-determinant $d\log T$ by the sketch-dependent quantity $\Phi_m = m\log T + d\varepsilon_m$, with the price of multiplicative

sketching factors $(1+\varepsilon_m)^{1/2}$ in the leading term and $(1+\varepsilon_m)^{3/2}$ in the lower-order term. When $m = d$ and the sketch becomes exact, Φ_m reduces to $d \log T$, and the regret bound recovers the same $\widetilde{O}(d\sqrt{T/\kappa_\star})$ as GLB-OMD. More interestingly, when the curvature-weighted history is effectively rank-m so that $\varepsilon_m = 0$, SGLB achieves the leading regret $\widetilde{O}(\sqrt{dmT/\kappa_\star})$ while still reducing the space and amortized per-round time from $O(d^2)$ and $O(d^3)$ to $O(md)$. This shows that, in favorable regimes such as low-rank or light spectrum, our method yields computational gains without sacrificing the statistical efficiency. We also note that OMD-based one-pass ideas have also been explored in other topics such as logistic bandits (Faury et al., 2022; Zhang & Sugiyama, 2023) and multinomial logit bandits (Lee & Oh, 2025); our contribution is complementary, targeting dimension-dependent computational costs through sketching while retaining the existing one-pass estimation framework.

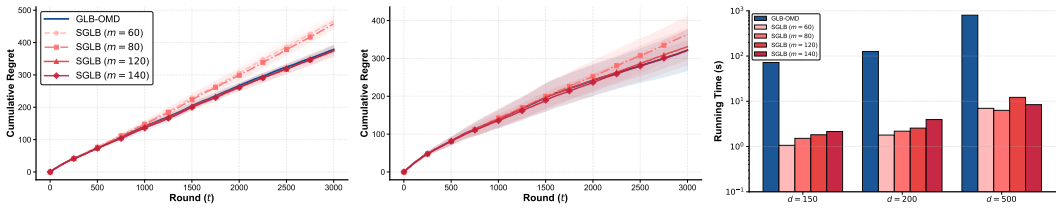
Regret Dependence. In its leading term, SGLB matches the best currently available $\kappa_\star$-dependent for GLBs, while replacing the full-dimensional factor by the sketch-aware quantity $\widetilde{m} = m + d\varepsilon_m$. In particular, recent MLE-based methods (Lee et al., 2024; Sawarni et al., 2024) achieve $\widetilde{O}(d\sqrt{T/\kappa_\star})$ regret, whereas SGLB attains the sketch-aware counterpart $\widetilde{O}(\sqrt{d\widetilde{m}T/\kappa_\star})$. The remaining gap appears in the lower-order term: Sawarni et al. (2024) obtains a slightly sharper dependence through $\kappa_{\mathcal{X}} = 1/\inf_{\boldsymbol{x}\in\cup_{t=1}^{T}\mathcal{X}_t} \mu'(\boldsymbol{x}^\top\boldsymbol{\theta}_\star)$, which can be smaller than $\kappa = 1/\inf_{\boldsymbol{x}\in\cup_{t=1}^{T}\mathcal{X}_t,\boldsymbol{\theta}\in\Theta} \mu'(\boldsymbol{x}^\top\boldsymbol{\theta})$. It is plausible that a similar refinement could also be obtained for SGLB by combining matrix sketching with an estimator warm-up phase as in Sawarni et al. (2024). Besides, unlike recent MLE-based analyses (Lee et al., 2024) that remove dependence on S and R in the leading term , our bound still carries these quantities (e.g., S^2-dependence) because it is built on an OMD-based update; this dependence is inherited from the underlying OMD estimator rather than introduced by sketching.

4.2 Discussion

Comparison with Other Efficient GLBs. Despite GLB-OMD, there are other works pursued computational efficiency on GLBs. RS-GLinCB (Sawarni et al., 2024) reduces update frequency through a rare-update schedule and achieves amortized $O(\log^2 t)$ time, but it remains a full-history MLE and still requires $O(t)$ memory. In contrast, SGLB is one-pass and dimension-efficient, costs only $O(md)$ space and $O(md)$ amortized time. The two ideas are therefore complementary rather than competing, and it would be interesting to combine sketch-based one-pass updates with rare parameter updates in the future work. Another related wrok is Jun et al. (2017), which is based on online-to-confidence-set conversion and instantiated with Online Newton Step (ONS) for GLMs. This approach is one-pass in spirit, but its computation remains dimension-inefficient and its regret bound is weaker than the proposed SGLB, scaling linearly with κ. Related sketch-based ONS techniques (Luo et al., 2016; Feinberg et al., 2023; Yang et al., 2025) may help reduce its dimensional cost, but improving the underlying κ-dependent regret would likely require a different analysis.

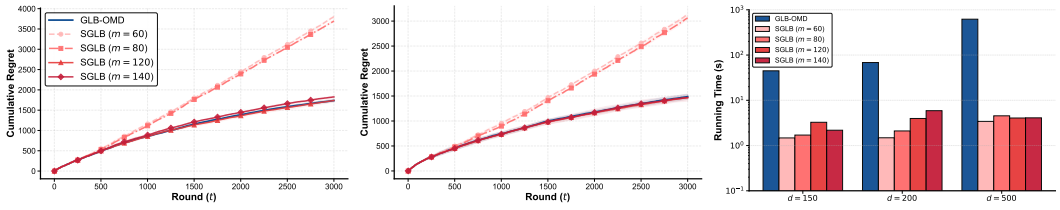
Comparison with Sketched Linear Bandits. In terms of computational and memory complexity, SGLB matches the known guarantees $O(d)$ for sketched linear bandits (Kuzborskij et al., 2019; Chen et al., 2020; Wen et al., 2026).However, extending sketching from linear bandits to SGLB is not a straightforward adaptation. In FD-based linear bandits (Kuzborskij et al., 2019), sketching is applied to the least-squares geometry, where the sketched object is the unweighted correlation matrix under the linear reward model $\mathbb{E}[r_t \mid \boldsymbol{x}_t] = \boldsymbol{x}_t^\top \boldsymbol{\theta}_\star$. By contrast, in SGLB the reward follows a generalized linear model with a non-linear link function $\mu(\cdot)$, and matrix sketching acts on the OMD curvature geometry induced by the historical losses, namely through a curvature-weighted action matrix. Consequently, the analysis must control not only the usual unweighted leverage scores, but also a curvature-dependent weighted elliptical potential that enters both the confidence construction and the regret analysis. From this perspective, linear bandits correspond to the special Gaussian GLM case with $\mu(z) = z$, whereas SGLB requires a substantially more general sketch-aware formulation. To the best of our knowledge, this is the first work to incorporate matrix sketching into generalized linear bandits setting while providing rigorous confidence-set and regret guarantees.

Comparison with Offline PCA. In sketched least-squares regression problems, principal component analysis (PCA) is a natural offline baseline (Shi & Phillips, 2021; Braverman et al., 2023). It computes the top-m principal directions $\boldsymbol{V}_m$ of the design matrix and then performs regression after projecting the data onto $\mathrm{span}(\boldsymbol{V}_m)$. In linear bandit settings, such offline reduction can be competitive and may empirically outperform online sketching methods (Kuzborskij et al., 2019).



(a) Logistic bandits ($d = 150$) (b) Logistic bandits ($d = 200$) (c) Total running time

Figure 1: (a), (b): Cumulative regret of the compared algorithms on synthetic data; (c): Total running time comparison on synthetic data under the logistic bandits model w.r.t. ambient dimension d.



(a) Poisson bandits ($d = 150$) (b) Poisson bandits ($d = 200$) (c) Total running time

Figure 2: (a), (b): Cumulative regret of the compared algorithms on synthetic data; (c): Total running time comparison on synthetic data under the Poisson bandits model w.r.t. ambient dimension d.

However, this offline PCA pipeline does not directly extend to GLBs. In least-squares problems, the approximated object is typically the unweighted covariance matrix $\boldsymbol{\mathcal{V}}_t = \lambda \boldsymbol{I}_d + \sum_{s=1}^{t-1} \boldsymbol{x}_s \boldsymbol{x}_s^\top$, whose geometry is determined solely by the actions. By contrast, in GLBs the relevant object is the curvature-weighted Hessian $\boldsymbol{H}_t = \lambda \boldsymbol{I}_d + \sum_{s=1}^{t-1} w_s \boldsymbol{x}_s \boldsymbol{x}_s^\top$, where $w_s := \mu'(\boldsymbol{x}_s^\top \boldsymbol{\theta}_{s+1})/g(\tau)$. Because these weights depend on both the non-linear link function and the online iterates, they are revealed only as learning unfolds. Consequently, an offline PCA step based solely on the action matrix cannot capture the effective geometry in GLBs, emphasizing the necessity of sketch-aware online reduction. Further experimental results is deferred to Section 5.2.

5 Experiments

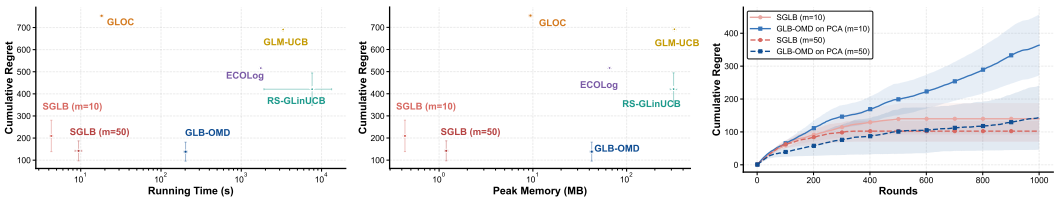
In this section, we evaluate our method on the synthetic dataset and several real-world datasets. The experimental setting, configurations, and additional experiments are available in Appendix F.

5.1 Synthetic Data: SGLB vs. GLB-OMD

We first compare SGLB against its full-matrix counterpart GLB-OMD (Zhang et al., 2025) on a synthetic dataset. We consider ambient dimensions $d = \{150, 200, 500\}$. Unless otherwise stated, we set the regularization parameter to $\lambda = 10$ and $\delta = 0.01$ for all methods, and report results averaged over 10 independent trials. SGLB is instantiated with sketch sizes $m = \{60, 80, 120, 140\}$.

Logistic Bandits. We instantiate the link function as $\mu(z) = 1/(1 + e^{-z})$ and evaluate the algorithms under bounded rewards. Figures 1a and 1b show that SGLB largely preserves the regret behavior of GLB-OMD across all tested sketch sizes. The curves with $m = 120$ and $m = 140$ remain close to GLB-OMD throughout the horizon, whereas smaller sketches $m = 60$ and $m = 80$ incur a moderate increase in regret. Notably, even $m = 60$, which is substantially smaller than d and only 60% of the latent rank, still achieves comparable regret growth over time. This suggests that the FD sketch retains most of the curvature information needed for action selection. In Figure 1c, we report the total running time over different d. As the ambient dimension increases, the total running time of GLB-OMD grows rapidly, while all SGLB variants remain much less sensitive to d.

Poisson Bandits. We next consider Poisson bandits with $\mu(z) = e^z$, which constitute a more challenging GLB setting due to their unbounded rewards. The results are reported in Figure 2. The same qualitative trade-off observed in logistic bandits remains: SGLB substantially improves computational efficiency, while preserving similar regret when the sketch size is sufficiently large.



(a) Regret vs Total Time (b) Regret vs Peak Memory (c) Offline PCA vs SGLB

Figure 3: (a), (b): Comprehensive comparison of different GLB algorithms on `MNIST`; (c): Cumulative regret comparison between offline PCA-based reduction and online sketching-based reduction.

In particular, the curves with $m = 60, 80$ exhibit a noticeable regret increase. Compared with the bounded logistic case, this degradation is more pronounced under the Poisson model, suggesting that the unbounded-reward setting is more sensitive to aggressive compression of the curvature geometry.

5.2 Real-World Data: Comparison with GLBs Baselines

We next evaluate SGLB on the real-world `MNIST` benchmark. The dataset contains 60,000 samples, each represented by a $d = 784$ feature vector and associated with one of 10 digit classes. We follow the setup in Filippi et al. (2010), details in Appendix F.3. We compare representative GLB baselines spanning three update regimes: *full-history* methods, including GLM-UCB (Filippi et al., 2010), ECOLog (Faury et al., 2020), and RS-GLinCB (Sawarni et al., 2024); *one-pass* methods, including GLOC (Jun et al., 2017) and GLB-OMD (Zhang et al., 2025); and the proposed *sketch-aware* method SGLB. For SGLB, we consider sketch sizes $m \in \{10, 50\}$; additional details on the baseline implementations and hyperparameter settings are provided in Appendix F.1.

Regret–Efficiency Trade-off. Figures 3a and 3b summarize the cumulative regret, average running time, and peak memory usage after 1000 rounds. Overall, SGLB exhibits a favorable regret–efficiency trade-off across both computation and memory. Compared with the lightweight baseline GLOC, both SGLB variants achieve substantially lower regret while operating at similarly low resource budgets. Compared with GLB-OMD, which attains the lowest regret among the baselines, SGLB reduces both running time and peak memory by more than one order of magnitude, at the cost of only a moderate increase in regret. Moreover, the two sketch sizes expose a clear operating trade-off: using a larger sketch ($m = 50$) improves regret and moves SGLB closer to GLB-OMD, whereas a smaller sketch ($m = 10$) further reduces resource consumption.

The Necessity of Online Sketching. The favorable regret–efficiency trade-off of SGLB may arise either from online sketching itself or simply from operating in a lower-dimensional space. To disentangle these two effects, we compare SGLB with an offline reduction baseline that first applies PCA (Jolliffe, 2011) to obtain a best m-dimensional subspace and then runs GLB-OMD on the projected features. Figure 3c shows that SGLB consistently outperforms this offline PCA baseline for both $m = 10$ and $m = 50$ throughout the horizon. The gap is particularly pronounced under the more aggressive compression level $m = 10$, where GLB-OMD on PCA accumulates substantially larger regret over time. These results indicate that the gain of SGLB does not come merely from reducing the ambient dimension. Rather, the key benefit comes from maintaining a curvature-aware sketch online, so that the compressed geometry can adapt to the evolving Hessian structure.

6 Conclusion

We introduce a sketch-aware algorithm for GLBs, reducing the memory and amortized per-round time complexities from $O(d^2)$ and $O(d^3)$, respectively, to $O(md)$. We establish a sketch-based regret bound which is explicitly characterized by the sketch size and spectral approximation error. Extensive experiments demonstrate that our method substantially reducing running time and space while remaining competitive performance. Future directions include developing adaptive sketching schemes for GLBs that control approximation error without requiring prior knowledge (Wen et al., 2026), investigating whether complementary techniques such as warm-up (Faury et al., 2022; Sawarni et al., 2024) can further tighten the regret guarantee, and improving the dimension dependence of sketch-aware regret bounds in the finite-arm setting (Jun et al., 2021; Mason et al., 2022).

AI USE STATEMENT

In this work, we used generative AI tools to provide feedback on experimental design and interpretation, check intermediate mathematical derivations and proof steps, and assist with code editing and debugging. We also used generative AI for literature search and summarization, and for language polishing and stylistic editing of author-written text. We did not use generative AI to generate datasets or experimental results, or to independently formulate final mathematical claims or proofs; other required-disclosure tasks were not applicable. All AI-assisted outputs were reviewed and verified by the authors: mathematical arguments were checked by the authors, code was tested against the described algorithms, experimental results were produced by our own implementations, and cited literature was verified against the original sources. We take responsibility for the final content of this work.

ETHICS STATEMENT

This work is theoretical and algorithmic, and does not involve human subjects, private data, or personally identifiable information. Experiments use synthetic data and publicly available benchmarks. Potential risks arise only from downstream deployment of bandit algorithms in sensitive decision-making applications.

REPRODUCIBILITY STATEMENT

Algorithmic and theoretical details are provided in Sections 3–4 and Appendices B–E, including complete proofs and implementation details. Experimental setups, hyperparameters, data processing, and additional results are provided in Section 5 and Appendix F.

REFERENCES

Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems 24 (NeurIPS)*, pp. 2312–2320, 2011.

Marc Abeille and Alessandro Lazaric. Linear thompson sampling revisited. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 176–184, 2017.

Marc Abeille, Louis Faury, and Clément Calauzènes. Instance-wise minimax-optimal algorithms for logistic bandits. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 3691–3699, 2021.

Dimitris Achlioptas. Database-friendly random projections. In *Proceedings of the 29th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS)*, pp. 274–281, 2001.

Rajeev Agrawal. Sample mean based index policies by $o(\log n)$ regret for the multi-armed bandit problem. *Advances in Applied Probability*, 27(4):1054–1078, 1995.

Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pp. 127–135, 2013.

Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3:397–422, 2002.

Hamsa Bastani and Mohsen Bayati. Online decision making with high-dimensional covariates. *Oper. Res.*, 68:276–294, 2020.

Hamsa Bastani, Kimon Drakopoulos, Vishal Gupta, Ioannis Vlachogiannis, Christos Hadjichristodoulou, Pagona Lagiou, Gkikas Magiorkinis, Dimitrios Paraskevis, and Sotirios Tsiodras. Efficient and targeted COVID-19 border testing via reinforcement learning. *Nature*, 599(7883): 108–113, 2021.

Vladimir Braverman, Dan Feldman, Harry Lang, Daniela Rus, and Adiel Statman. Least-mean-squares coresets for infinite streams. *IEEE Trans. Knowl. Data Eng.*, 35(9):8699–8712, 2023.

Daniele Calandriello, Alessandro Lazaric, and Michal Valko. Efficient second-order online kernel learning with adaptive embedding. In *Advances in Neural Information Processing Systems 30 (NeurIPS)*, pp. 6140–6150, 2017.

Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, and Lorenzo Rosasco. Gaussian process optimization with adaptive sketching: Scalable and no regret. In *Proceedings of the 32nd Annual Conference on Learning Theory (COLT)*, pp. 533–557, 2019.

Cheng Chen, Luo Luo, Weinan Zhang, Yong Yu, and Yijiang Lian. Efficient and robust high-dimensional linear contextual bandits. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 4259–4265, 2020.

Eugenio Clerico, Hamish Flynn, Wojciech Kotlowski, and Gergely Neu. Confidence sequences for generalized linear models via regret analysis. *arXiv preprint arXiv:2504.16555*, 2025.

Varsha Dani, Thomas P. Hayes, and Sham M. Kakade. The price of bandit information for online optimization. In *Advances in Neural Information Processing Systems 20 (NeurIPS)*, pp. 345–352, 2007.

Amit Deshpande and Luis Rademacher. Efficient volume sampling for row/column subset selection. In *Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 329–338, 2010.

Shi Dong, Tengyu Ma, and Benjamin Van Roy. On the performance of thompson sampling on logistic bandits. In *Proceedings of the 32nd Annual Conference on Learning Theory (COLT)*, pp. 1158–1160, 2019.

Louis Faury, Marc Abeille, Clément Calauzènes, and Olivier Fercoq. Improved optimistic algorithms for logistic bandits. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pp. 3052–3060, 2020.

Louis Faury, Marc Abeille, Kwang-Sung Jun, and Clément Calauzènes. Jointly efficient and optimal algorithms for logistic bandits. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 546–580, 2022.

Vladimir Feinberg, Xinyi Chen, Y. Jennifer Sun, Rohan Anil, and Elad Hazan. Sketchy: Memory-efficient adaptive regularization with frequent directions. In *Advances in Neural Information Processing Systems 36 (NeurIPS)*, 2023.

Sarah Filippi, Olivier Cappé, Aurélien Garivier, and Csaba Szepesvári. Parametric bandits: The generalized linear case. In *Advances in Neural Information Processing Systems 23 (NeurIPS)*, pp. 586–594, 2010.

Alan M. Frieze, Ravi Kannan, and Santosh S. Vempala. Fast monte-carlo algorithms for finding low-rank approximations. *Journal of the ACM*, 51(6):1025–1041, 2004.

Mina Ghashami, Edo Liberty, Jeff M. Phillips, and David P. Woodruff. Frequent directions: Simple and deterministic matrix sketching. *SIAM Journal on Computing*, 45(5):1762–1792, 2016.

Alon Gonen, Francesco Orabona, and Shai Shalev-Shwartz. Solving ridge regression using sketched preconditioned SVRG. In *Proceedings of the 33nd International Conference on Machine Learning (ICML)*, pp. 1397–1405, 2016.

Ming Gu and Stanley C Eisenstat. A stable and fast algorithm for updating the singular value decomposition, 1993.

Robbins Herbert. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, pp. 527–535, 1952.

Shunsuke Ihara. *Information theory for continuous systems*. World Scientific, 1993.

Adel Javanmard and Hamid Nazerzadeh. Dynamic pricing in high-dimensions. *J. Mach. Learn. Res.*, 20:9:1–9:49, 2019.

Ian T. Jolliffe. Principal component analysis. In *International Encyclopedia of Statistical Science*, pp. 1094–1096. Springer, 2011.

Kwang-Sung Jun, Aniruddha Bhargava, Robert D. Nowak, and Rebecca Willett. Scalable generalized linear bandits: Online computation and hashing. In *Advances in Neural Information Processing Systems 30 (NeurIPS)*, pp. 99–109, 2017.

Kwang-Sung Jun, Lalit Jain, Houssam Nassif, and Blake Mason. Improved confidence bounds for the linear logistic model and applications to bandits. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pp. 5148–5157, 2021.

Johannes Kirschner, Andreas Krause, Michele Meziu, and Mojmir Mutny. Confidence estimation via sequential likelihood mixing. *arXiv preprint arXiv:2502.14689*, 2025.

Gabriel Krummenacher, Brian McWilliams, Yannic Kilcher, Joachim M. Buhmann, and Nicolai Meinshausen. Scalable adaptive stochastic optimization using random projections. In *Advances in Neural Information Processing Systems 29 (NeurIPS)*, pp. 1750–1758, 2016.

Ilja Kuzborskij, Leonardo Cella, and Nicolò Cesa-Bianchi. Efficient linear bandits through matrix sketching. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 177–185, 2019.

Joongkyu Lee and Min-hwan Oh. Improved online confidence bounds for multinomial logistic bandits. *arXiv preprint arXiv:2502.10020*, 2025.

Junghyun Lee, Se-Young Yun, and Kwang-Sung Jun. A unified confidence sequence for generalized linear models, with applications to bandits. In *Advances in Neural Information Processing Systems 38 (NeurIPS)*, 2024.

Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web (WWW)*, pp. 661–670. ACM, 2010.

Edo Liberty. Simple and deterministic matrix sketching. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 581–588, 2013.

Shuai Liu, Alex Ayoub, Flore Sentenac, Xiaoqi Tan, and Csaba Szepesvári. Almost free: Self-concordance in natural exponential families and an application to bandits. In *Advances in Neural Information Processing Systems 38 (NeurIPS)*, 2024.

Haipeng Luo, Alekh Agarwal, Nicolò Cesa-Bianchi, and John Langford. Efficient second order online learning by sketching. In *Advances in Neural Information Processing Systems 29 (NeurIPS)*, pp. 902–910, 2016.

Luo Luo, Cheng Chen, Zhihua Zhang, Wu-Jun Li, and Tong Zhang. Robust frequent directions with application in online learning. *Journal of Machine Learning Research*, 20:45:1–45:41, 2019.

Blake Mason, Kwang-Sung Jun, and Lalit Jain. An experimental design approach for regret minimization in logistic bandits. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 7736–7743, 2022.

Peter McCullagh and John A. Nelder. *Generalized Linear Models*. Springer, 1989.

Zakaria Mhammedi, Wouter M. Koolen, and Tim van Erven. Lipschitz adaptivity with multiple learning rates in online learning. In *Proceedings of the 32nd Annual Conference on Learning Theory (COLT)*, pp. 2490–2511, 2019.

Jayadev Misra and David Gries. Finding repeated elements. *Science of computer programming*, 2 (2):143–152, 1982.

Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.

Mark D. Reid, Rafael M. Frongillo, Robert C. Williamson, and Nishant A. Mehta. Generalized mixability via entropic duality. In *Proceedings of the 28th Annual Conference on Learning Theory (COLT)*, pp. 1501–1522, 2015.

Yoan Russac, Louis Faury, Olivier Cappé, and Aurélien Garivier. Self-concordant analysis of generalized linear bandits with forgetting. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 658–666, 2021.

Tamás Sarlós. Improved approximation algorithms for large matrices via random projections. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 143–152, 2006.

Ayush Sawarni, Nirjhar Das, Siddharth Barman, and Gaurav Sinha. Generalized linear bandits with limited adaptivity. In *Advances in Neural Information Processing Systems 38 (NeurIPS)*, 2024.

Jack Sherman and Winifred J Morrison. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1):124–127, 1950.

Benwei Shi and Jeff M. Phillips. A deterministic streaming sketch for ridge regression. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 586–594, 2021.

Joaquin Vanschoren, Jan N. van Rijn, Bernd Bischl, and Luís Torgo. Openml: networked science in machine learning. *SIGKDD Explor.*, 15(2):49–60, 2013.

Vladimir Vovk. Competitive on-line statistics. *International Statistical Review*, 69(2):213–248, 2001.

Jing Wang, Miao Yu, Peng Zhao, and Zhi-Hua Zhou. Learning with adaptive resource allocation. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, pp. 52099–52116, 2024.

Jing Wang, Xi-Tong Liu, and Zhi-Hua Zhou. Core-learning with look-ahead and immediate resource allocation. In *Proceedings of the 40th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 26294–26301, 2026.

Dongxie Wen, Xiao Zhang, Zhewei Wei, Chenping Hou, Shuai Li, and Weinan Zhang. Fast second-order online kernel learning through incremental matrix sketching and decomposition. In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 6552–6560, 2025.

Dongxie Wen, Hanyan Yin, Xiao Zhang, Peng Zhao, Lijun Zhang, and Zhewei Wei. Revisiting matrix sketching in linear bandits: Achieving sublinear regret via dyadic block sketching. In *Proceedings of the the 14th International Conference on Learning Representations (ICLR)*, 2026.

David P. Woodruff. Sketching as a tool for numerical linear algebra. *Foundations and Trends in Theoretical Computer Science*, 10(1-2):1–157, 2014.

Lily Xu, Elizabeth Bondi, Fei Fang, Andrew Perrault, Kai Wang, and Milind Tambe. Dual-mandate patrols: Multi-armed bandits for green security. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 14974–14982, 2021.

Sifan Yang, Yuanyu Wan, Peijia Li, Yibo Wang, Xiao Zhang, Zhewei Wei, and Lijun Zhang. Dimension-free adaptive subgradient methods with frequent directions. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, 2025.

Xiaotian Yu, Michael R. Lyu, and Irwin King. CBRAP: contextual bandits with random projection. In *Proceedings of the 31th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 2859–2866, 2017.

Lijun Zhang, Tianbao Yang, Rong Jin, Yichi Xiao, and Zhi-Hua Zhou. Online stochastic linear optimization under one-bit feedback. In *Proceedings of the 33nd International Conference on Machine Learning (ICML)*, pp. 392–401, 2016.

Xiao Zhang, Sunhao Dai, Jun Xu, Zhenhua Dong, Quanyu Dai, and Ji-Rong Wen. Counteracting user attention bias in music streaming recommendation via reward modification. In *Proceedings of the 28th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 2504–2514, 2022.

Xiao Zhang, Ninglu Shao, Zihua Si, Jun Xu, Wenhan Wang, Hanjing Su, and Ji-Rong Wen. Reward imputation with sketching for contextual batched bandits. In *Advances in Neural Information Processing Systems 36 (NeurIPS)*, 2023.

Yu-Jie Zhang and Masashi Sugiyama. Online (multinomial) logistic bandit: Improved regret and constant computation cost. In *Advances in Neural Information Processing Systems 36 (NeurIPS)*, 2023.

Yu-Jie Zhang, Sheng-An Xu, Peng Zhao, and Masashi Sugiyama. Generalized linear bandits: Almost optimal regret with one-pass update. In *Advances in Neural Information Processing Systems 39 (NeurIPS)*, 2025.

Zhi-Hua Zhou. Learnability with time-sharing computational resource concerns. *National Science Review*, 11(10):204, 2024.

Appendix: Contents

A More Related Works **16**
A.1 Matrix Sketching 16
A.2 Sketch-Based Online Learning 16
A.3 Related Topics of GLBs 17

B Proof of Theorem 1 **17**
B.1 Main Proof 17
B.2 Supporting Lemmas 19

C Proof of Theorem 2 **24**
C.1 Main Proof 24
C.2 Supporting Lemmas 27

D Omitted Standard Lemmas **29**

E Omitted Algorithms **30**
E.1 Efficient Projection by Sketching 30
E.2 Fast Algorithm of SGLB 32

F Omitted Details for Section 5 **33**
F.1 Experimental Setups 33
F.2 Additional Results on Synthetic Data 34
F.3 Additional Results on Real-World Data 34

A More Related Works

In this section, we review several lines of work that are closely related to our approach, including matrix sketching in Appendix A.1, sketch-based online learning in Appendix A.2, and related topics in generalized linear bandits in Appendix A.3. We first provide omitted notations in the main paper.

Notations. Let $[n] = \{1, 2, \ldots, n\}$, upper-case bold letters (e.g., $\boldsymbol{A}$) represent matrix and lower-case bold letters (e.g., $\boldsymbol{a}$) represent vectors. We denote by $\|\boldsymbol{A}\|_2$ and $\|\boldsymbol{A}\|_F$ the spectral and Frobenius norms of $\boldsymbol{A}$. We define $|\boldsymbol{A}|$ and $\mathrm{Tr}(\boldsymbol{A})$ as the determinant and trace of matrix $\boldsymbol{A}$. For a positive semi-definite matrix $\boldsymbol{A}$, the matrix norm of vector $\boldsymbol{x}$ is defined by $\|\boldsymbol{x}\|_{\boldsymbol{A}} = \sqrt{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}$. For two positive semi-definite matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we use $\boldsymbol{A} \succeq \boldsymbol{B}$ to represent that $\boldsymbol{A} - \boldsymbol{B}$ is positive semi-definite. We use $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$ to represent the singular value decomposition (SVD) of $\boldsymbol{A}$, where $\boldsymbol{U}, \boldsymbol{V}$ denote the left and right matrices of singular vectors and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, ..., \sigma_n)$ is the diagonal matrix of singular values in the descending order. For the function $f(\cdot) : \mathbb{R} \to \mathbb{R}$, its first and second derivatives are denoted by f' and f'', respectively.

A.1 Matrix Sketching

Matrix sketching studies how to maintain a compact surrogate for a matrix observed sequentially in a streaming, row-arrival setting. Given the rows of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times d}$ arriving one by one, the goal is to maintain a sketch $\boldsymbol{B} \in \mathbb{R}^{m\times d}$ with $m \ll n$ such that $\boldsymbol{B}^\top \boldsymbol{B}$ well approximates $\boldsymbol{A}^\top \boldsymbol{A}$, thereby preserving the dominant covariance structure of the stream.

Existing methods can be grouped into three broad families. The first family is *sampling-based*, which selects a small subset of rows or columns that approximately preserves the geometry of the full matrix (Frieze et al., 2004; Deshpande & Rademacher, 2010). The second family is *projection-based*, which forms randomized linear combinations of the input rows, typically via random projections or hashing, to obtain a low-dimensional embedding with provable approximation guarantees (Achlioptas, 2001; Sarlós, 2006). The third family is *deterministic shrinkage-based*. A representative example is Frequent Directions (FD) (Liberty, 2013; Ghashami et al., 2016; Luo et al., 2019), which can be viewed as a matrix analogue of the Misra–Gries heavy-hitters algorithm (Misra & Gries, 1982): it repeatedly shrinks less informative directions while retaining the dominant ones.

Among these approaches, Frequent Directions is particularly attractive for streaming learning. Unlike sampling- or projection-based alternatives, FD supports one-pass updates, uses only $O(md)$ memory, and provides a deterministic spectral approximation guarantee; by contrast, some random projection-based approaches do not offer the same reduction in memory footprint in the streaming setting (Krummenacher et al., 2016; Yu et al., 2017). These properties are especially appealing in our setting, where the object to be compressed is not merely a data matrix, but a streaming second-order geometry that underlies both the estimator update and the confidence construction. In particular, FD aligns well with our goal of replacing the full-matrix geometry in one-pass GLBs with a compact sketch while retaining rigorous control of the induced approximation error. For broader background on matrix sketching, we refer the reader to the survey of Woodruff (2014).

A.2 Sketch-Based Online Learning

A closely related line of work studies how matrix sketching can be used to accelerate online learning by compressing second-order or kernel-dependent structures while retaining the information most relevant for prediction and optimization. Existing approaches include both probabilistic sketches, such as random projections and sampling, and deterministic sketches, such as Frequent Directions. Beyond linear bandits, sketching has been used to improve the computational efficiency of second-order online optimization (Luo et al., 2016; Feinberg et al., 2023; Yang et al., 2025), online kernel learning (Calandriello et al., 2017; Wen et al., 2025), stochastic optimization (Gonen et al., 2016), and contextual batched bandits (Zhang et al., 2023).

Most existing methods in this literature employ *single-scale* sketches, namely, a fixed sketch size m maintained throughout learning. Recent work has begun to explore *multi-scale* sketching strategies, in which sketches of different sizes are used to guide the learning process more adaptively (Calandriello et al., 2019; Wen et al., 2026). In particular, Wen et al. (2026) showed in linear bandits (Kuzborskij et al., 2019; Chen et al., 2020) that an undersized sketch can fundamentally alter

the statistical behavior of the algorithm and may even lead to linear regret. In their analysis, the spectral error ε_m introduced by sketching can enter the leading regret term unless it is properly controlled. To address this issue, they maintain multiple sketches of different sizes and dynamically adjust the effective approximation scale during learning, thereby preserving computational gains while avoiding the need for prior knowledge of a suitable sketch size. Our work is closer to the single-scale setting: we also maintain a single FD sketch of size m throughout learning. As a result, our method does not resolve the sketch-size selection problem, and analogous regret pitfalls may arise when the spectral error is too large. Establishing error-controlled or adaptive multi-scale sketching schemes for GLBs therefore remains a open problem.

Another conceptually related direction is CoRE-Learning (Zhou, 2024; Wang et al., 2024; 2026), which studies how computational resources should be allocated during learning in order to balance efficiency and learnability. From this perspective, the proposed sketch-aware GLBs method can be viewed as one form of computational resource allocation in single thread.

A.3 Related Topics of GLBs

Logistic Bandits. Logistic bandits (Zhang et al., 2016; Faury et al., 2020; Abeille et al., 2021; Faury et al., 2022; Zhang & Sugiyama, 2023) form an important special case of generalized linear bandits, in which the specific structure of the logistic link can be exploited to obtain sharper statistical regret guarantees than those available for general GLBs. For example, in the finite-N arms setting, Mason et al. (2022) proposed an experimental-design-based method with regret $\widetilde{O}(\sqrt{d \log N \cdot T})$, improving the dependence on the ambient dimension. Likewise, Jun et al. (2021) combined fixed-design ideas with variance-sensitive confidence bounds to obtain improved d dependence in logistic bandits. These results are complementary to our perspective. They exploit additional structure specific to the logistic model, whereas our goal is to reduce the full-matrix computational burden of one-pass GLBs through sketching while remaining applicable to general link functions. Since the performance of sketch-aware GLBs is sensitive to the effective dimension and the spectral error, an interesting open direction is to investigate whether such logistic-specific ideas can be incorporated into our framework to further mitigate the statistical cost induced by sketching.

Thompson Sampling and Large Action Sets. Another relevant direction is Thompson sampling (Agrawal & Goyal, 2013; Abeille & Lazaric, 2017), which provides an alternative to UCB-based exploration in generalized linear and contextual bandits. This perspective is particularly relevant when the action set is very large or infinite. In such settings, even if the estimator update is made efficient, the action selection step in (10) can remain the dominant computational bottleneck because it requires optimizing an optimism bonus over the entire action set. By contrast, a Bayesian-style strategy replaces this step by optimization with respect to a sampled parameter, which is often algorithmically simpler. From this viewpoint, our sketch-aware OMD estimator can be regarded as a modular plug-in component rather than as being tied exclusively to OFU exploration. In particular, the proposed SGLB can potentially be combined with sketch-aware Thompson sampling (Kuzborskij et al., 2019; Dong et al., 2019; Faury et al., 2022), where arm selection reduces to a linear or convex optimization problem for suitable arm classes, such as convex action sets. Exploring such combinations in the GLB setting is another promising direction for future work.

B Proof of Theorem 1

This section is devoted to the proof of Theorem 1. We first present the main proof in Appendix B.1, and then state and prove the supporting technical lemmas in Appendix B.2.

B.1 Main Proof

Roadmap. We prove Theorem 1, namely, that with probability at least $1-\delta$, the confidence set

$$\widetilde{\mathcal{C}}_t(\delta) := \left\{ \boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|_{\widetilde{\boldsymbol{H}}_t} \le \widetilde{\beta}_t(\delta) \right\}$$

contains $\boldsymbol{\theta}_\star$ simultaneously for all rounds t. Following Zhang et al. (2025), the proof has three ingredients. First, an inverse-regret inequality reduces the estimation error $\|\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t}$ to a cumulative excess-loss term. Second, we decompose this excess loss via the mix loss m_s. Third,

we bound the resulting stochastic and deterministic terms separately: the stochastic term via Ville's inequality, and the deterministic term via FD-aware trace/KL bounds together with a sketched log-determinant estimate. For convenience, we restate the sketch-based OMD estimator:

$$\boldsymbol{\theta}_{t+1} = \arg\min_{\boldsymbol{\theta}\in\Theta} \left\{ \widetilde{\ell}_t(\boldsymbol{\theta}) + \frac{1}{2\eta}\|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|^2_{\widetilde{\boldsymbol{H}}_t} \right\} \quad \text{where} \quad \widetilde{\boldsymbol{H}}_t := \lambda \boldsymbol{I}_d + \boldsymbol{B}_t^\top \boldsymbol{B}_t.$$

Proof of Theorem 1. We begin with the relation of the estimation error to the so-called "inverse-regret". Under Assumptions 1–3, with $\eta = 1 + RS$ and $\lambda \geq 6\eta RSL_\mu / g(\tau)$, Lemma 2 yields

$$\|\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t} \leq 2\eta \left(\sum_{s=1}^t \ell_s(\boldsymbol{\theta}_\star) - \sum_{s=1}^t \ell_s(\boldsymbol{\theta}_{s+1}) \right) + 4\lambda S^2 - \frac{2}{3}\sum_{s=1}^t \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}. \tag{11}$$

The first term on the right-hand side is the inverse-regret term. Let $\alpha = 3\eta/2$, define the mix loss by

$$m_s(P_s) := -\ln \mathbb{E}_{\boldsymbol{\theta}\sim P_s}\left[e^{-\ell_s(\boldsymbol{\theta})} \right], \quad \text{where} \quad P_s = \mathcal{N}(\boldsymbol{\theta}_s, \alpha \widetilde{\boldsymbol{H}}_s^{-1}).$$

We refer to m_s as the mix loss because it averages the loss under the distribution P_s. It has been been found useful for achieving fast rates in prediction with expert advice and online optimization problems (Vovk, 2001; Kirschner et al., 2025). We then decompose the inverse-regret term as

$$\sum_{s=1}^t \ell_s(\boldsymbol{\theta}_\star) - \sum_{s=1}^t \ell_s(\boldsymbol{\theta}_{s+1}) = \underbrace{\sum_{s=1}^t [\ell_s(\boldsymbol{\theta}_\star) - m_s(P_s)]}_{\text{term (a)}} + \underbrace{\sum_{s=1}^t [m_s(P_s) - \ell_s(\boldsymbol{\theta}_{s+1})]}_{\text{term (b)}}. \tag{12}$$

Bounding Term (a). By construction, P_s is a Gaussian distribution and $\mathcal{F}_s$-measurable. Hence, Lemma 12 in Appendix D together implies that, with probability at least $1-\delta$, for all $t \geq 1$,

$$\text{term (a)} \leq \ln \frac{1}{\delta}. \tag{13}$$

Bounding Term (b). Since the mix loss is the convex conjugate of the KL divergence, according to Lemma 13, we can decompose the mix loss term as

$$m_s(P_s) = \mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\ell_s(\boldsymbol{\theta})] + \mathrm{KL}(Q_s \,\|\, P_s) - \mathrm{KL}(Q_s \,\|\, P_s^\star),$$

where

$$Q_s = \mathcal{N}(\boldsymbol{\theta}_{s+1}, \alpha \widetilde{\boldsymbol{H}}_{s+1}^{-1}), \qquad P_s^\star(\boldsymbol{\theta}) \propto P_s(\boldsymbol{\theta}) e^{-\ell_s(\boldsymbol{\theta})}, \quad \forall \boldsymbol{\theta} \in \mathbb{R}^d.$$

Since $-\mathrm{KL}(Q_s \,\|\, P_s^\star) \leq 0$, we obtain

$$m_s(P_s) - \ell_s(\boldsymbol{\theta}_{s+1}) \leq (\mathbb{E}_{Q_s}[\ell_s(\boldsymbol{\theta})] - \ell_s(\boldsymbol{\theta}_{s+1})) + \mathrm{KL}(Q_s \,\|\, P_s). \tag{14}$$

For any $s \in [T]$, let ρ_s denote the shrinkage incurred by the FD update at round s in Algorithm 1. Given $\lambda \geq 14 d\eta R^2$ and applying Lemma 3, we bound the first term in (14) as

$$\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\ell_s(\boldsymbol{\theta})] - \ell_s(\boldsymbol{\theta}_{s+1}) \leq 2\alpha \left(\ln \frac{\det(\widetilde{\boldsymbol{H}}_{s+1})}{\det(\widetilde{\boldsymbol{H}}_s)} + \frac{\rho_s d}{\lambda} \right).$$

Next, applying Lemma 4, we bound the second term in (14) as

$$\mathrm{KL}(Q_s \,\|\, P_s) \leq \frac{1}{2} \left(\ln \frac{\det(\widetilde{\boldsymbol{H}}_{s+1})}{\det(\widetilde{\boldsymbol{H}}_s)} + \frac{\rho_s d}{\lambda} + \frac{\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}}{\alpha} \right).$$

Summing over $s = 1, \ldots, t$ and writing $\bar{\rho}_t = \sum_{s=1}^t \rho_s$, we obtain

$$\begin{aligned}
\text{term (b)} &\leq \left(2\alpha + \frac{1}{2}\right) \sum_{s=1}^t \ln \frac{\det(\widetilde{\boldsymbol{H}}_{s+1})}{\det(\widetilde{\boldsymbol{H}}_s)} + \left(2\alpha + \frac{1}{2}\right) \sum_{s=1}^t \frac{\rho_s d}{\lambda} + \frac{1}{2\alpha} \sum_{s=1}^t \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s} \\
&= \left(2\alpha + \frac{1}{2}\right) \ln \frac{\det(\widetilde{\boldsymbol{H}}_{t+1})}{\det(\widetilde{\boldsymbol{H}}_1)} + \left(2\alpha + \frac{1}{2}\right) \frac{d\bar{\rho}_t}{\lambda} + \frac{1}{2\alpha} \sum_{s=1}^t \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s} \qquad (15) \\
&\leq \left(2\alpha + \frac{1}{2}\right) \left[m \ln \left(1 + \frac{L_\mu t}{m\lambda g(\tau)} \right) + \frac{d\bar{\rho}_t}{\lambda} \right] + \frac{1}{2\alpha} \sum_{s=1}^t \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s},
\end{aligned}$$

where the last inequality follows from Lemma 5.

Combining the bounds. Substituting (13) and (15) into (12), we obtain

$$\sum_{s=1}^{t} \ell_s(\boldsymbol{\theta}_\star) - \sum_{s=1}^{t} \ell_s(\boldsymbol{\theta}_{s+1}) \le \ln \frac{1}{\delta} + \left(2\alpha + \frac{1}{2}\right)\left[m \ln\left(1 + \frac{L_\mu t}{m\lambda g(\tau)}\right) + \frac{d\bar{\rho}_t}{\lambda}\right] + \frac{1}{2\alpha}\sum_{s=1}^{t} \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}. \tag{16}$$

Substituting (16) back into (11) and recalling that $\alpha = 3\eta/2$, we have

$$\begin{aligned}
\|\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t} &\le 4\lambda S^2 + 2\eta \ln \frac{1}{\delta} + 2\eta\left(3\eta + \frac{1}{2}\right)\left[m \ln\left(1 + \frac{L_\mu t}{m\lambda g(\tau)}\right) + \frac{d\bar{\rho}_t}{\lambda}\right] \\
&\quad + \frac{2}{3}\sum_{s=1}^{t} \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s} - \frac{2}{3}\sum_{s=1}^{t} \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s} \\
&= 4\lambda S^2 + 2\eta \ln \frac{1}{\delta} + 2\eta\left(3\eta + \frac{1}{2}\right)\left[m \ln\left(1 + \frac{L_\mu t}{m\lambda g(\tau)}\right) + \frac{d\bar{\rho}_t}{\lambda}\right] \\
&:= \widetilde{\beta}_t^2(\delta) = O\left(S^2R^2\left(m \ln \frac{t}{\delta} + d\left(S^2R + \varepsilon_m\right)\right)\right),
\end{aligned}$$

where the last equality holds since $\bar{\rho}_t/\lambda \le \varepsilon_m$ by Lemma 1. This completes the proof. □

B.2 Supporting Lemmas

This section presents several useful lemmas used in the proof of Theorem 1.

Lemma 2 (Pathway of Sketched Estimator to Sketched Inverse Regret)**.** *Under Assumptions 1–3, with $\eta = 1 + RS$ and $\lambda \ge 6\eta RSL_\mu/g(\tau)$, the FD-sketched OMD estimator satisfies*

$$\|\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t} \le 2\eta\left(\sum_{s=1}^{t} \ell_s(\boldsymbol{\theta}_\star) - \sum_{s=1}^{t} \ell_s(\boldsymbol{\theta}_{s+1})\right) + 4\lambda S^2 - \frac{2}{3}\sum_{s=1}^{t} \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}.$$

Proof of Lemma 2. We start from the integral form of Taylor's theorem. Since μ is twice differentiable, we have

$$\ell_s(\boldsymbol{\theta}_{s+1}) - \ell_s(\boldsymbol{\theta}_\star) = \langle \nabla \ell_s(\boldsymbol{\theta}_{s+1}), \boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star \rangle - \|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{h}_s}, \tag{17}$$

where

$$\widetilde{h}_s = \int_0^1 (1-v)\, \nabla^2 \ell_s(\boldsymbol{\theta}_{s+1} + v(\boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}))\, dv.$$

By the definition of the loss function, we can rewrite $\widetilde{h}_s$ as

$$\widetilde{h}_s = \frac{\widetilde{\alpha}(\boldsymbol{\theta}_{s+1}, \boldsymbol{\theta}_\star, \boldsymbol{x}_s)}{g(\tau)}\, \boldsymbol{x}_s \boldsymbol{x}_s^\top \quad \text{where} \quad \widetilde{\alpha}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{x}_s) = \int_0^1 (1-v)\, \mu'\left(\boldsymbol{x}_s^\top \boldsymbol{\theta}_1 + v\, \boldsymbol{x}_s^\top (\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)\right) dv.$$

Under Assumption 3, we have

$$|\mu''(z)| \le R \cdot \mu'(z), \qquad \forall z \in [-S, S].$$

Therefore, Lemma 11 in Appendix D implies that for any $\boldsymbol{\theta}_\star \in \Theta := \{\boldsymbol{\theta} \in \mathbb{R}^d \mid \|\boldsymbol{\theta}\|_2 \le S\}$,

$$\widetilde{\alpha}(\boldsymbol{\theta}_{s+1}, \boldsymbol{\theta}_\star, \boldsymbol{x}_s) \ge \frac{\mu'(\boldsymbol{x}_s^\top \boldsymbol{\theta}_{s+1})}{2 + 2RS}. \tag{18}$$

Substituting (18) into (17) gives

$$\begin{aligned}
\ell_s(\boldsymbol{\theta}_{s+1}) - \ell_s(\boldsymbol{\theta}_\star) &\le \langle \nabla \ell_s(\boldsymbol{\theta}_{s+1}), \boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star \rangle - \frac{1}{2 + 2RS} \|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\nabla^2 \ell_s(\boldsymbol{\theta}_{s+1})} \\
&= \langle \nabla \ell_s(\boldsymbol{\theta}_{s+1}), \boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star \rangle - \frac{1}{2\eta} \|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\nabla^2 \ell_s(\boldsymbol{\theta}_{s+1})},
\end{aligned} \tag{19}$$

where the equality uses $\eta = 1 + RS$.

Since $\boldsymbol{\theta}_{s+1}$ is the minimizer in (6), Lemma 10, applied with $f = \widetilde{\ell}_s$ and $\mathbf{u} = \boldsymbol{\theta}_\star$, yields

$$\|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s} \le 2\eta\langle\nabla\widetilde{\ell}_s(\boldsymbol{\theta}_{s+1}),\, \boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}\rangle + \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s} - \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}. \tag{20}$$

We next decompose the gradient term:

$$\langle\nabla\widetilde{\ell}_s(\boldsymbol{\theta}_{s+1}),\, \boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}\rangle = \langle\nabla\ell_s(\boldsymbol{\theta}_{s+1}),\, \boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}\rangle + \langle\nabla\widetilde{\ell}_s(\boldsymbol{\theta}_{s+1}) - \nabla\ell_s(\boldsymbol{\theta}_{s+1}),\, \boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}\rangle.$$

Substituting (19) into the above identity gives

$$\begin{aligned}\langle\nabla\widetilde{\ell}_s(\boldsymbol{\theta}_{s+1}),\, \boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}\rangle \le \ell_s(\boldsymbol{\theta}_\star) - \ell_s(\boldsymbol{\theta}_{s+1}) - \frac{1}{2\eta}\|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})} \\ + \langle\nabla\widetilde{\ell}_s(\boldsymbol{\theta}_{s+1}) - \nabla\ell_s(\boldsymbol{\theta}_{s+1}),\, \boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}\rangle.\end{aligned} \tag{21}$$

By the definitions of the true and surrogate losses, the gradient difference is

$$\begin{aligned}\nabla\widetilde{\ell}_s(\boldsymbol{\theta}_{s+1}) - \nabla\ell_s(\boldsymbol{\theta}_{s+1}) &= \frac{\mu(\boldsymbol{x}_s^\top\boldsymbol{\theta}_s) - \mu(\boldsymbol{x}_s^\top\boldsymbol{\theta}_{s+1})}{g(\tau)}\boldsymbol{x}_s + \nabla^2\ell_s(\boldsymbol{\theta}_s)(\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_s) \\ &= \frac{\mu(\boldsymbol{x}_s^\top\boldsymbol{\theta}_s)\boldsymbol{x}_s - \mu(\boldsymbol{x}_s^\top\boldsymbol{\theta}_{s+1})\boldsymbol{x}_s + \mu'(\boldsymbol{x}_s^\top\boldsymbol{\theta}_s)\cdot\boldsymbol{x}_s\boldsymbol{x}_s^\top(\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_s)}{g(\tau)}.\end{aligned}$$

By Taylor's theorem with Lagrange remainder, for some $\boldsymbol{\zeta} \in \Theta$ lying on the line segment connecting $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_{s+1}$, we have

$$\begin{aligned}\langle\nabla\widetilde{\ell}_s(\boldsymbol{\theta}_{s+1}) - \nabla\ell_s(\boldsymbol{\theta}_{s+1}),\, \boldsymbol{\theta}_\star - \boldsymbol{\theta}_{s+1}\rangle &= \frac{\mu''(\boldsymbol{x}_s^\top\boldsymbol{\zeta})}{2g(\tau)}\cdot\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\boldsymbol{x}_s\boldsymbol{x}_s^\top}\cdot\boldsymbol{x}_s^\top(\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star) \\ &\le \frac{R}{2g(\tau)}\cdot\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\mu'(\boldsymbol{x}_s^\top\boldsymbol{\zeta})\boldsymbol{x}_s\boldsymbol{x}_s^\top}\cdot\boldsymbol{x}_s^\top(\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star) \\ &\le RS\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\nabla^2\ell_s(\boldsymbol{\zeta})} \\ &\le \frac{RSL_\mu}{g(\tau)}\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|_2^2,\end{aligned} \tag{22}$$

where the first inequality uses Assumption 3, the second uses $\|\boldsymbol{x}_s\|_2 \le 1$ together with the fact that both $\boldsymbol{\theta}_{s+1}$ and $\boldsymbol{\theta}_\star$ belong to Θ, and the last uses Assumption 2.

Substituting (21) and (22) into (20) gives

$$\begin{aligned}\|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s} \le 2\eta\left(\ell_s(\boldsymbol{\theta}_\star) - \ell_s(\boldsymbol{\theta}_{s+1})\right) + \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s} \\ + \frac{2\eta RSL_\mu}{g(\tau)}\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|_2^2 - \|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})} - \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}.\end{aligned}$$

Summing over $s = 1, \ldots, t$ yields

$$\begin{aligned}&\sum_{s=1}^t\left(\|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s} - \|\boldsymbol{\theta}_s - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s}\right) \\ &\le 2\eta\sum_{s=1}^t\left(\ell_s(\boldsymbol{\theta}_\star) - \ell_s(\boldsymbol{\theta}_{s+1})\right) - \sum_{s=1}^t\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s} - \sum_{s=1}^t\|\boldsymbol{\theta}_{s+1} - \boldsymbol{\theta}_\star\|^2_{\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})} \\ &\quad + \frac{2\eta RSL_\mu}{g(\tau)}\sum_{s=1}^t\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|_2^2.\end{aligned} \tag{23}$$

The left-hand side telescopes, with an additional correction term due to the time-varying metric:

$$\begin{aligned}
&\sum_{s=1}^{t}\left(\|\boldsymbol{\theta}_{s+1}-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s}-\|\boldsymbol{\theta}_s-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s}\right)\\
&=\|\boldsymbol{\theta}_{t+1}-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t}-\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_1}-\sum_{s=2}^{t}\|\boldsymbol{\theta}_s-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_s-\widetilde{\boldsymbol{H}}_{s-1}}\\
&\geq\|\boldsymbol{\theta}_{t+1}-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t}-\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_1}-\sum_{s=1}^{t-1}\|\boldsymbol{\theta}_{s+1}-\boldsymbol{\theta}_\star\|^2_{\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})},
\end{aligned}$$

where the inequality holds since $\widetilde{\boldsymbol{H}}_s \preceq \boldsymbol{H}_s$ for all $s \in [T]$. Substituting the inequality into the left hand side of (23), we obtain

$$\begin{aligned}
\|\boldsymbol{\theta}_{t+1}-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t} &\leq \|\boldsymbol{\theta}_1-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_1}+2\eta\sum_{s=1}^{t}\left(\ell_s(\boldsymbol{\theta}_\star)-\ell_s(\boldsymbol{\theta}_{s+1})\right)\\
&\quad+\frac{2\eta RSL_\mu}{g(\tau)}\sum_{s=1}^{t}\|\boldsymbol{\theta}_s-\boldsymbol{\theta}_{s+1}\|_2^2-\sum_{s=1}^{t}\|\boldsymbol{\theta}_s-\boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}\\
&\leq \|\boldsymbol{\theta}_1-\boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_1}+2\eta\sum_{s=1}^{t}\left(\ell_s(\boldsymbol{\theta}_\star)-\ell_s(\boldsymbol{\theta}_{s+1})\right)-\frac{2}{3}\sum_{s=1}^{t}\|\boldsymbol{\theta}_s-\boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}\\
&\leq 2\eta\sum_{s=1}^{t}\left(\ell_s(\boldsymbol{\theta}_\star)-\ell_s(\boldsymbol{\theta}_{s+1})\right)+4\lambda S^2-\frac{2}{3}\sum_{s=1}^{t}\|\boldsymbol{\theta}_s-\boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}.
\end{aligned}$$

The second inequality uses $\lambda \geq 6\eta RSL_\mu/g(\tau)$ and $\widetilde{\boldsymbol{H}}_s \succeq \lambda \boldsymbol{I}_d$. The last inequality uses $\widetilde{\boldsymbol{H}}_1 = \lambda \boldsymbol{I}_d$ together with $\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_\star\|_2 \leq 2S$. This completes the proof. □

Lemma 3 (Sketched Bregman Divergence Bound). *Let $Q_s = \mathcal{N}(\boldsymbol{\theta}_{s+1}, \alpha\widetilde{\boldsymbol{H}}_{s+1}^{-1})$ be a Gaussian distribution with mean $\boldsymbol{\theta}_{s+1}$ and covariance matrix $\alpha\widetilde{\boldsymbol{H}}_{s+1}^{-1}$, where $\widetilde{\boldsymbol{H}}_{s+1}$ is the approximate Hessian with shrinking value ρ_s and $\boldsymbol{\theta}_{s+1}$ is the FD-sketched OMD estimator at round $s+1$. Then:*

$$\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\ell_s(\boldsymbol{\theta})]-\ell_s(\boldsymbol{\theta}_{s+1}) \leq 2\alpha\left(\ln\frac{\det(\widetilde{\boldsymbol{H}}_{s+1})}{\det(\widetilde{\boldsymbol{H}}_s)}+\frac{\rho_s d}{\lambda}\right).$$

Proof of Lemma 3. Since Q_s is symmetric around $\boldsymbol{\theta}_{s+1}$, we can rewrite

$$\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\ell_s(\boldsymbol{\theta})]-\ell_s(\boldsymbol{\theta}_{s+1}) = \mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\mathcal{B}_{\ell_s}(\boldsymbol{\theta},\boldsymbol{\theta}_{s+1})],$$

where

$$\mathcal{B}_{\ell_s}(\boldsymbol{\theta},\boldsymbol{\theta}_{s+1}) := \ell_s(\boldsymbol{\theta})-\ell_s(\boldsymbol{\theta}_{s+1})-\langle\nabla\ell_s(\boldsymbol{\theta}_{s+1}),\,\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\rangle$$

is the Bregman divergence of ℓ_s between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_{s+1}$.

By the integral form of Taylor's theorem and the definition of ℓ_s, for which

$$\nabla^2\ell_s(\boldsymbol{\theta}) = \frac{\mu'(\boldsymbol{x}_s^\top\boldsymbol{\theta})}{g(\tau)}\boldsymbol{x}_s\boldsymbol{x}_s^\top, \qquad \forall\boldsymbol{\theta}\in\mathbb{R}^d,$$

we have

$$\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\mathcal{B}_{\ell_s}(\boldsymbol{\theta},\boldsymbol{\theta}_{s+1})] = \mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|^2_{\widetilde{h}_s(\boldsymbol{\theta})}\right],$$

where

$$\widetilde{h}_s(\boldsymbol{\theta}) = \int_0^1(1-v)\,\mu'(\boldsymbol{x}_s^\top\boldsymbol{\theta}_{s+1}+v\,\boldsymbol{x}_s^\top(\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}))\,dv\cdot\frac{\boldsymbol{x}_s\boldsymbol{x}_s^\top}{g(\tau)}.$$

According to Lemma 11 and $\|\boldsymbol{x}_s\|_2 \leq 1$, we can bound the Hessian as

$$\widetilde{h}_s(\boldsymbol{\theta}) \leq \exp(R^2\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|_2^2)\cdot\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}).$$

Therefore,

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\mathcal{B}_{\ell_s}(\boldsymbol{\theta},\boldsymbol{\theta}_{s+1})] &\le \mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\exp(R^2\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|_2^2)\cdot\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|^2_{\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})}\right]\\ &\le \sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\exp(2R^2\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|_2^2)\right]}\cdot\sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|^4_{\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})}\right]}\\ &= \sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\exp(2R^2\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|_2^2)\right]}\cdot\sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\left\|(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}))^{\frac{1}{2}}(\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1})\right\|_2^4\right]}\end{aligned} \tag{24}$$

where the inequality follows from the Cauchy–Schwarz inequality.

To bound the former term in (24), let $\boldsymbol{e}_i$ be an orthogonal basis and λ_i be the i-th largest eigenvalue of $\widetilde{\boldsymbol{H}}_{s+1}^{-1}$, the random variable $\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}$ follows the same distribution as

$$\sum_{i=1}^{d}\sqrt{\alpha\lambda_i}n_i\boldsymbol{e}_i, \qquad n_i\sim\mathcal{N}(0,1),\ \forall i\in[d].$$

Then, the former term in (24) is equivalent to

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\exp(2R^2\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|_2^2)\right] &= \mathbb{E}_{n_i\sim\mathcal{N}(0,1)}\left[\prod_{i=1}^{d}\exp(2R^2\alpha\lambda_i n_i^2)\right]\\ &\le \left(\mathbb{E}_{n\sim\mathcal{N}(0,1)}\left[\exp(2R^2\alpha n^2/\lambda)\right]\right)^d,\end{aligned}$$

where the inequality holds because $\widetilde{\boldsymbol{H}}_{s+1}\succeq\lambda\boldsymbol{I}_d$ and $\max_{i\in[d]}\lambda_i\le 1/\lambda$. Let χ^2 denotes the chi-squared distribution with degree of freedom 1. By the Jensen's inequality, we can rewrite the above equation as

$$\begin{aligned}\left(\mathbb{E}_{n\sim\mathcal{N}(0,1)}\left[\exp(2R^2n^2/\lambda)\right]\right)^d &= \left(\mathbb{E}_{Z\sim\chi^2}\left[\exp(2R^2\alpha Z/\lambda)\right]\right)^d\\ &\le \mathbb{E}_{Z\sim\chi^2}\left[\exp\left(\frac{2R^2\alpha d}{\lambda}Z\right)\right]\le\frac{4}{3},\end{aligned} \tag{25}$$

where the last inequality is due to the condition $\lambda\ge 14d\eta R^2$ and the fact that the moment-generating function of the chi-squared distribution $\mathbb{E}[\exp(cZ)]\le(1-2c)^{-1/2}$ for all $c<1/2$.

To bound the latter term in (24), let $\boldsymbol{e}_i$ be an orthogonal basis and λ_i be the i-th largest eigenvalue of $(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}))^{1/2}\widetilde{\boldsymbol{H}}_{s+1}^{-1}(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}))^{1/2}$. Then the random variable $(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}))^{1/2}(\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1})$ has the same distribution as

$$\sum_{i=1}^{d}\sqrt{\alpha\lambda_i}n_i\boldsymbol{e}_i, \qquad n_i\sim\mathcal{N}(0,1),\ \forall i\in[d].$$

Hence,

$$\begin{aligned}\sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\left\|(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}))^{\frac{1}{2}}(\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1})\right\|_2^4\right]} &= \sqrt{\mathbb{E}_{n_i\sim\mathcal{N}(0,1)}\left[\left(\sum_{i=1}^{d}\alpha\lambda_i n_i^2\right)^2\right]}\\ &= \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d}\alpha^2\lambda_i\lambda_j\mathbb{E}_{n_i,n_j\sim\mathcal{N}(0,1)}[n_i^2n_j^2]}\\ &= \sqrt{3}\alpha\mathrm{Tr}\left(\widetilde{\boldsymbol{H}}_{s+1}^{-1}(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}))\right),\end{aligned} \tag{26}$$

where the last equality uses $\mathrm{Tr}(\boldsymbol{AB})=\mathrm{Tr}(\boldsymbol{BA})$ for matrices $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{d\times d}$.

Substituting (25) and (26) to (24), we have

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}[\mathcal{B}_{\ell_s}(\boldsymbol{\theta},\boldsymbol{\theta}_{s+1})] &\le \sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\exp(2R^2\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|_2^2)\right]}\cdot\sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim Q_s}\left[\|\boldsymbol{\theta}-\boldsymbol{\theta}_{s+1}\|^4_{\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})}\right]}\\ &\le \frac{2}{\sqrt{3}}\sqrt{3}\alpha\mathrm{Tr}\left(\widetilde{\boldsymbol{H}}_{s+1}^{-1}(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}))\right).\end{aligned}$$

According to the definition of Bregman divergence, we obtain

$$\begin{aligned}\mathbb{E}_{Q_s}[\ell_s(\boldsymbol{\theta})] - \ell_s(\boldsymbol{\theta}_{s+1}) &\le 2\alpha \mathrm{Tr}\left(\widetilde{\boldsymbol{H}}_{s+1}^{-1}\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})\right)\\ &\le 2\alpha\left(\mathrm{Tr}\left(\widetilde{\boldsymbol{H}}_{s+1}^{-1}(\widetilde{\boldsymbol{H}}_{s+1}-\widetilde{\boldsymbol{H}}_s)\right) + \rho_s \mathrm{Tr}(\widetilde{\boldsymbol{H}}_{s+1}^{-1})\right)\\ &\le 2\alpha \ln\det(\widetilde{\boldsymbol{H}}_{s+1}) - 2\alpha\ln\det(\widetilde{\boldsymbol{H}}_s) + 2\alpha\rho_s\mathrm{Tr}(\widetilde{\boldsymbol{H}}_{s+1}^{-1})\\ &\le 2\alpha \ln\det(\widetilde{\boldsymbol{H}}_{s+1}) - 2\alpha\ln\det(\widetilde{\boldsymbol{H}}_s) + \frac{2\alpha d\rho_s}{\lambda},\end{aligned}$$

where the second inequality uses the FD recursion

$$\widetilde{\boldsymbol{H}}_{s+1} = \widetilde{\boldsymbol{H}}_s + \nabla^2\ell_s(\boldsymbol{\theta}_{s+1}) - \rho_s\boldsymbol{I}_d,$$

and the last inequality uses $\widetilde{\boldsymbol{H}}_{s+1} \succeq \lambda\boldsymbol{I}_d$. □

Lemma 4 (Sketched KL Divergence Bound)**.** *Given a constant $\alpha > 0$, let*

$$P_s = \mathcal{N}(\boldsymbol{\theta}_s, \alpha\widetilde{\boldsymbol{H}}_s^{-1}) \quad \textit{and} \quad Q_s = \mathcal{N}(\boldsymbol{\theta}_{s+1}, \alpha\widetilde{\boldsymbol{H}}_{s+1}^{-1})$$

be Gaussian distributions with means $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_{s+1}$, and covariance matrices $\alpha\widetilde{\boldsymbol{H}}_s^{-1}$ and $\alpha\widetilde{\boldsymbol{H}}_{s+1}^{-1}$, respectively. Then

$$\mathrm{KL}(Q_s \,\|\, P_s) \le \frac{1}{2}\left(\ln\frac{\det(\widetilde{\boldsymbol{H}}_{s+1})}{\det(\widetilde{\boldsymbol{H}}_s)} + \frac{\rho_s d}{\lambda} + \frac{\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}}{\alpha}\right).$$

Proof. By Lemma 14 in Appendix D, the exact KL divergence between Q_s and P_s is

$$\mathrm{KL}(Q_s \,\|\, P_s) = \frac{1}{2}\left(\ln\frac{\det(\alpha\widetilde{\boldsymbol{H}}_s^{-1})}{\det(\alpha\widetilde{\boldsymbol{H}}_{s+1}^{-1})} + \mathrm{Tr}(\widetilde{\boldsymbol{H}}_s\widetilde{\boldsymbol{H}}_{s+1}^{-1}) - d + \frac{\|\boldsymbol{\theta}_s - \boldsymbol{\theta}_{s+1}\|^2_{\widetilde{\boldsymbol{H}}_s}}{\alpha}\right).$$

We first simplify the determinant ratio:

$$\ln\frac{\det(\alpha\widetilde{\boldsymbol{H}}_s^{-1})}{\det(\alpha\widetilde{\boldsymbol{H}}_{s+1}^{-1})} = \ln\frac{\det(\widetilde{\boldsymbol{H}}_{s+1})}{\det(\widetilde{\boldsymbol{H}}_s)}.$$

Next, for the trace term, we have

$$\begin{aligned}\mathrm{Tr}(\widetilde{\boldsymbol{H}}_s\widetilde{\boldsymbol{H}}_{s+1}^{-1}) - d &= \mathrm{Tr}\Big((\widetilde{\boldsymbol{H}}_s - \widetilde{\boldsymbol{H}}_{s+1})\widetilde{\boldsymbol{H}}_{s+1}^{-1}\Big)\\ &= \mathrm{Tr}\Big((-\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}) + \rho_s\boldsymbol{I}_d)\widetilde{\boldsymbol{H}}_{s+1}^{-1}\Big)\\ &= -\mathrm{Tr}\Big(\nabla^2\ell_s(\boldsymbol{\theta}_{s+1})\widetilde{\boldsymbol{H}}_{s+1}^{-1}\Big) + \rho_s\mathrm{Tr}(\widetilde{\boldsymbol{H}}_{s+1}^{-1}).\end{aligned}$$

Since $\nabla^2\ell_s(\boldsymbol{\theta}_{s+1}) \succeq 0$ by Assumption 2 and $\widetilde{\boldsymbol{H}}_{s+1}^{-1} \succ 0$, the first term is non-positive. Therefore,

$$\mathrm{Tr}(\widetilde{\boldsymbol{H}}_s\widetilde{\boldsymbol{H}}_{s+1}^{-1}) - d \le \rho_s\mathrm{Tr}(\widetilde{\boldsymbol{H}}_{s+1}^{-1}) \le \frac{\rho_s d}{\lambda},$$

where the last inequality uses $\widetilde{\boldsymbol{H}}_{s+1} \succeq \lambda\boldsymbol{I}_d$. Substituting the above bounds into the KL divergence between Q_s and P_s yields the claimed result. □

Lemma 5 (Sketched Hessian Determinant-trace Inequality)**.** *The log-determinant ratio satisfies*

$$\ln\frac{\det(\widetilde{\boldsymbol{H}}_{T+1})}{\det(\lambda\boldsymbol{I}_d)} \le m\ln\left(1 + \frac{L_\mu T}{m\lambda g(\tau)}\right).$$

Proof. Since

$$\widetilde{\boldsymbol{H}}_{T+1} = \lambda \boldsymbol{I}_d + \boldsymbol{B}_T^\top \boldsymbol{B}_T \qquad \text{and} \qquad \operatorname{rank}(\boldsymbol{B}_T) \le m,$$

all but at most m eigenvalues of $\widetilde{\boldsymbol{H}}_{T+1}$ are equal to λ. Let $\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_m \ge 0$ denote the eigenvalues of $\boldsymbol{B}_T^\top \boldsymbol{B}_T$, padded with zeros if necessary. Then

$$\begin{aligned} \ln \frac{\det(\widetilde{\boldsymbol{H}}_{T+1})}{\det(\lambda \boldsymbol{I}_d)} &= \sum_{i=1}^{m} \ln\left(\frac{\widetilde{\lambda}_i + \lambda}{\lambda}\right) \\ &= \sum_{i=1}^{m} \ln\left(1 + \frac{\widetilde{\lambda}_i}{\lambda}\right). \end{aligned}$$

By the AM–GM inequality,

$$\sum_{i=1}^{m} \ln\left(1 + \frac{\widetilde{\lambda}_i}{\lambda}\right) \le m \ln\left(1 + \frac{1}{m}\sum_{i=1}^{m} \frac{\widetilde{\lambda}_i}{\lambda}\right) = m \ln\left(1 + \frac{\operatorname{Tr}(\boldsymbol{B}_T^\top \boldsymbol{B}_T)}{m\lambda}\right).$$

Since FD preserves the trace up to sketching,

$$\operatorname{Tr}(\boldsymbol{B}_T^\top \boldsymbol{B}_T) \le \operatorname{Tr}(\boldsymbol{A}_T^\top \boldsymbol{A}_T) = \sum_{s=1}^{T} w_s(\boldsymbol{\theta}_{s+1}) \, \|\boldsymbol{x}_s\|_2^2 \le \frac{L_\mu T}{g(\tau)},$$

where the last inequality follows from Assumptions 1 and 2. Substituting this bound into the previous display finishes the proof. □

C Proof of Theorem 2

In this section, we establish Theorem 2 in full detail. The core argument is developed in Appendix C.1, while the auxiliary technical lemmas and their proofs are deferred to Appendix C.2.

C.1 Main Proof

Roadmap. With the confidence set in Theorem 1 at hand, the regret analysis of Algorithm 2 can in principle be carried out within the standard OFU framework (Agrawal, 1995; Abbasi-Yadkori et al., 2011; Zhang et al., 2025). However, doing so in our setting requires several new ingredients. The first difficulty is to connect the reward approximation error of the OMD estimator (Zhang et al., 2025) with the error induced by sketching the Hessian geometry. In particular, we must propagate the spectral approximation error into the regret analysis while accounting for the fact that the FD-based Hessian approximation no longer preserves the Loewner order, which necessitates additional error terms. A further challenge, absent from existing sketched bandit analyses (Kuzborskij et al., 2019; Chen et al., 2020; Wen et al., 2026), is that the nonlinear link function μ in GLBs forces us to control both curvature-weighted elliptical potentials and unweighted leverage scores.

Proof of Theorem 2. Let the optimal action at round t be $\boldsymbol{x}_t^\star = \arg\max_{\boldsymbol{x} \in \mathcal{X}_t} \mu(\boldsymbol{x}^\top \boldsymbol{\theta}_\star)$, and let

$$(\boldsymbol{x}_t, \boldsymbol{\theta}_t^{\texttt{OFU}}) = \operatorname*{arg\,max}_{\boldsymbol{x} \in \mathcal{X}_t, \boldsymbol{\theta} \in \widetilde{\mathcal{C}}_t(\delta)} \mu(\boldsymbol{x}^\top \boldsymbol{\theta}).$$

Throughout the proof, we work on the event $\mathcal{E} = \{\forall t \ge 1,\ \boldsymbol{\theta}_\star \in \widetilde{\mathcal{C}}_t(\delta)\}$, which holds with probability at least $1 - \delta$ by Theorem 1. On this event, the OFU-based selection rule in (10) guarantees that, for each round t, there exists some $\boldsymbol{\theta}_t^{\texttt{OFU}} \in \widetilde{\mathcal{C}}_t(\delta)$ such that $\mu(\boldsymbol{x}_t^\top \boldsymbol{\theta}_t^{\texttt{OFU}}) \ge \mu(\boldsymbol{x}_t^{\star\top} \boldsymbol{\theta}_\star)$.

We now decompose the regret. Since μ is increasing, the above inequality allows us to upper bound the instantaneous regret by an optimistic gap evaluated at the played action. Applying Taylor expan-

sion with integral remainder gives

$$\begin{aligned}\mathrm{Regret}_T &= \sum_{t=1}^{T}\big[\mu(\boldsymbol{x}_t^{\star\top}\boldsymbol{\theta}_\star) - \mu(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\big] \\ &\le \sum_{t=1}^{T}\big[\mu(\boldsymbol{x}_t^\top\boldsymbol{\theta}_t^{\texttt{OFU}}) - \mu(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\big] \\ &= \underbrace{\sum_{t=1}^{T}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\boldsymbol{x}_t^\top(\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star)}_{\text{term (a)}} + \underbrace{\frac{1}{2}\sum_{t=1}^{T}\widetilde{\alpha}_t\left(\boldsymbol{x}_t^\top(\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star)\right)^2}_{\text{term (b)}},\end{aligned} \tag{27}$$

where $\widetilde{\alpha}_t = \int_0^1 (1-v)\,\mu''\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star + v\,\boldsymbol{x}_t^\top(\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star)\big)\,dv$.

Bounding Term (a). We first control the former term in (27). Since both $\boldsymbol{\theta}_\star$ and $\boldsymbol{\theta}_t^{\texttt{OFU}}$ belong to $\widetilde{\mathcal{C}}_t(\delta)$, Hölder's inequality for dual norms yields

$$\left|\boldsymbol{x}_t^\top(\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star)\right| \le \|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}}\left\|\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star\right\|_{\widetilde{\boldsymbol{H}}_t} \le 2\widetilde{\beta}_t(\delta)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}}.$$

Substituting this bound into term (a), and then using $\widetilde{\beta}_t(\delta) \le \widetilde{\beta}_T(\delta)$, we obtain

$$\begin{aligned}\text{term (a)} &\le 2\sum_{t=1}^{T}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\widetilde{\beta}_t(\delta)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} \\ &\le 2\widetilde{\beta}_T(\delta)\sum_{t=1}^{T}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} \\ &= 2\widetilde{\beta}_T(\delta)\sum_{t\in\mathcal{T}_1}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} + 2\widetilde{\beta}_T(\delta)\sum_{t\in\mathcal{T}_2}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}},\end{aligned} \tag{28}$$

where

$$\mathcal{T}_1 = \{t \in [T] : \mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star) \ge \mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1})\}, \qquad \mathcal{T}_2 = [T]\setminus\mathcal{T}_1.$$

We next bound the two parts separately.

For the contribution over $\mathcal{T}_1$, we compare these two quantities by expanding μ' around $\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}$. By Taylor expansion and the self-concordance conditions in Assumptions 2 and 3,

$$\begin{aligned}\mu'\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star\big) &= \mu'\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}\big) + \int_0^1 \mu''\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1} + v\boldsymbol{x}_t^\top(\boldsymbol{\theta}_\star - \boldsymbol{\theta}_{t+1})\big)\,dv\cdot\boldsymbol{x}_t^\top(\boldsymbol{\theta}_\star - \boldsymbol{\theta}_{t+1}) \\ &\le \mu'\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}\big) + R\int_0^1 \mu'\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1} + v\boldsymbol{x}_t^\top(\boldsymbol{\theta}_\star - \boldsymbol{\theta}_{t+1})\big)\,dv\cdot\left|\boldsymbol{x}_t^\top(\boldsymbol{\theta}_\star - \boldsymbol{\theta}_{t+1})\right| \\ &\le \mu'\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}\big) + RL_\mu\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_{t+1}^{-1}}\,\|\boldsymbol{\theta}_\star - \boldsymbol{\theta}_{t+1}\|_{\widetilde{\boldsymbol{H}}_{t+1}}.\end{aligned}$$

Here, the second line uses self-concordance, and the third line uses Hölder's inequality.

We now invoke the confidence bound from Theorem 1. Since $\boldsymbol{\theta}_\star \in \widetilde{\mathcal{C}}_{t+1}(\delta)$ on $\mathcal{E}$,

$$\begin{aligned}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star) &\le \mu'\big(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}\big) + RL_\mu\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_{t+1}^{-1}}\,\|\boldsymbol{\theta}_\star - \boldsymbol{\theta}_{t+1}\|_{\widetilde{\boldsymbol{H}}_{t+1}} \\ &\le \mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}) + RL_\mu\,\widetilde{\beta}_{t+1}(\delta)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_{t+1}^{-1}} \\ &\le \mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}) + RL_\mu\,\widetilde{\beta}_{T+1}(\delta)\,\sqrt{1+\varepsilon_m}\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}},\end{aligned}$$

where the last inequality follows from Lemma 6. Substituting it into the $\mathcal{T}_1$ part of (28), we obtain

$$\begin{aligned}\sum_{t\in\mathcal{T}_1}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} &\le \sum_{t\in\mathcal{T}_1}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1})\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} + RL_\mu\,\widetilde{\beta}_{T+1}\sqrt{1+\varepsilon_m}\sum_{t=1}^{T}\|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}}\\ &= \sum_{t\in\mathcal{T}_1}\sqrt{\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1})}\;\|\boldsymbol{z}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} + RL_\mu\,\widetilde{\beta}_{T+1}\sqrt{1+\varepsilon_m}\sum_{t=1}^{T}\|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}}\\ &\le \sqrt{\sum_{t\in\mathcal{T}_1}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1})}\cdot\sqrt{\sum_{t=1}^{T}\|\boldsymbol{z}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} + RL_\mu\,\widetilde{\beta}_{T+1}\sqrt{1+\varepsilon_m}\sum_{t=1}^{T}\|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}}}\,,\end{aligned} \tag{29}$$

where the second line uses $\boldsymbol{z}_t := \sqrt{w_s(\boldsymbol{\theta}_{s+1})}\boldsymbol{x}_s$, and the third line is Cauchy–Schwarz inequality.

We next control the two sums appearing on the right-hand side. By Lemma 7,

$$\sum_{t=1}^{T}\|\boldsymbol{z}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \le 2(1+\varepsilon_m)(1+L_\mu/g(\tau))\,\Lambda_m^T.$$

where Λ_m^T is defined in (32) for simplicity. By Lemma 8,

$$\sum_{t=1}^{T}\|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \le 2\kappa g(\tau)\,(1+\varepsilon_m)(1+L_\mu/g(\tau))\,\Lambda_m^T.$$

Plugging these bounds into (29) yields

$$\begin{aligned}2\widetilde{\beta}_T(\delta)\sum_{t\in\mathcal{T}_1}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} &\le \widetilde{\gamma}_T^{(1)}\sqrt{\sum_{t\in\mathcal{T}_1}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)} + \widetilde{\gamma}_T^{(2)}\\ &\le \widetilde{\gamma}_T^{(1)}\sqrt{T/\kappa_* + R\cdot\mathrm{Regret}_T} + \widetilde{\gamma}_T^{(2)},\end{aligned} \tag{30}$$

where the last inequality uses the same argument as in the proof of Theorem 1 of Abeille et al. (2021) to bound $\sum_{t\in\mathcal{T}_1}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)$, and where, for notational convenience,

$$\begin{aligned}\widetilde{\gamma}_T^{(1)}(\delta) &= \frac{2\widetilde{\beta}_T(\delta)}{\sqrt{g(\tau)}}\sqrt{2(1+\varepsilon_m)(1+L_\mu/g(\tau))\,\Lambda_m^T},\\ \widetilde{\gamma}_T^{(2)}(\delta) &= 4RL_\mu\,\kappa g(\tau)\,(1+\varepsilon_m)^{3/2}\,(1+L_\mu/g(\tau))\,\widetilde{\beta}_T(\delta)\,\widetilde{\beta}_{T+1}(\delta)\,\Lambda_m^T.\end{aligned}$$

We now turn to the contribution over $\mathcal{T}_2$. In this case,

$$\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star) \le \mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1}).$$

Therefore, the same Cauchy–Schwarz and elliptical-potential argument used above gives

$$\begin{aligned}2\widetilde{\beta}_T(\delta)\sum_{t\in\mathcal{T}_2}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star)\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}} &\le 2\widetilde{\beta}_T(\delta)\sum_{t\in\mathcal{T}_2}\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_{t+1})\,\|\boldsymbol{x}_t\|_{\widetilde{\boldsymbol{H}}_t^{-1}}\\ &\le \widetilde{\gamma}_T^{(1)}\sqrt{T/\kappa_* + R\cdot\mathrm{Regret}_T}.\end{aligned} \tag{31}$$

Combining (30), (31), and (28), we conclude that

$$\text{term (a)} \le 2\widetilde{\gamma}_T^{(1)}\sqrt{T/\kappa_* + R\cdot\mathrm{Regret}_T} + \widetilde{\gamma}_T^{(2)}.$$

Bounding Term (b). We next bound the quadratic remainder term in (27). By the self-concordant property of μ and Assumption 2,

$$\begin{aligned}\text{term (b)} &= \frac{1}{2}\sum_{t=1}^{T}\widetilde{\alpha}_t\left(\boldsymbol{x}_t^\top(\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star)\right)^2\\ &\le \frac{R}{2}\sum_{t=1}^{T}\int_0^1\mu'\left(\boldsymbol{x}_t^\top\boldsymbol{\theta}_\star + v\,\boldsymbol{x}_t^\top(\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star)\right)dv\,\left(\boldsymbol{x}_t^\top(\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star)\right)^2\\ &\le \frac{RL_\mu}{2}\sum_{t=1}^{T}\left\|\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star\right\|^2_{\widetilde{\boldsymbol{H}}_t}\|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}}\,,\end{aligned}$$

where the first inequality uses self-concordance, and the second uses Cauchy–Schwarz inequality. Since $\|\boldsymbol{\theta}_t^{\texttt{OFU}} - \boldsymbol{\theta}_\star\|^2_{\widetilde{\boldsymbol{H}}_t} \le 4\widetilde{\beta}_T^2(\delta)$ by Theorem 1, we further obtain

$$\text{term (b)} \le 2RL_\mu\, \widetilde{\beta}_T^2(\delta) \sum_{t=1}^{T} \|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \le \widetilde{\gamma}_T^{(2)},$$

where the last inequality follows from Lemma 8 and $\varepsilon_m > 0$.

Combining Together. Substituting the bounds of term (a) and term (b) into (27), we obtain

$$\mathrm{Regret}_T \le 2\widetilde{\gamma}_T^{(1)} \sqrt{T/\kappa_* + R \cdot \mathrm{Regret}_T} + 2\widetilde{\gamma}_T^{(2)}.$$

For brevity, write $\gamma_1 = \widetilde{\gamma}_T^{(1)}$ and $\gamma_2 = \widetilde{\gamma}_T^{(2)}$. Rearranging the above inequality gives

$$(\mathrm{Regret}_T - 2\gamma_2)^2 \le 4\gamma_1^2(T/\kappa_* + R \cdot \mathrm{Regret}_T).$$

Expanding the right-hand side and solving the resulting quadratic inequality yield

$$\mathrm{Regret}_T^2 - 4\gamma_2 \mathrm{Regret}_T + 4\gamma_2^2 \le 4\gamma_1^2 T/\kappa_* + 4\gamma_1^2 R \mathrm{Regret}_T,$$
$$\mathrm{Regret}_T^2 - (4\gamma_2 + 4\gamma_1^2 R)\mathrm{Regret}_T + (4\gamma_2^2 - 4\gamma_1^2 T/\kappa_*) \le 0.$$

Hence,

$$\begin{aligned}
\mathrm{Regret}_T &\le 2\gamma_2 + 2\gamma_1^2 R + 2\gamma_1 \sqrt{\gamma_1^2 R^2 + 2\gamma_2 R + T/\kappa_*} \\
&\le 2\gamma_2 + 4\gamma_1^2 R + 2\gamma_1 \sqrt{2\gamma_2 R} + 2\gamma_1 \sqrt{T/\kappa_*} \\
&\lesssim SR\,(1+\varepsilon_m)^{1/2} \sqrt{(dS^2R + \Phi_m)\,\Phi_m\,\frac{T}{\kappa_*}} + \kappa S^2 R^3\,(1+\varepsilon_m)^{3/2}(dS^2R + \Phi_m)\,\Phi_m,
\end{aligned}$$

where $\Phi_m := m \log T + d\varepsilon_m$ and the second inequality uses $\sqrt{a+b+c} \le \sqrt{a} + \sqrt{b} + \sqrt{c}$. □

C.2 Supporting Lemmas

This section presents several technical lemmas used in the proof of Theorem 2.

Lemma 6 (Approximate Monotonicity). *For all $\mathbf{x} \in \mathbb{R}^d$ and $t \ge 1$,*

$$\|\mathbf{x}\|_{\widetilde{\boldsymbol{H}}_{t+1}^{-1}} \le \sqrt{1+\varepsilon_m}\ \|\mathbf{x}\|_{\widetilde{\boldsymbol{H}}_t^{-1}}.$$

Proof. We combine the norm comparison between $\boldsymbol{H}_t$ and $\widetilde{\boldsymbol{H}}_t$ with the exact monotonicity of $\boldsymbol{H}_t$. First, from

$$\boldsymbol{H}_{t+1} = \widetilde{\boldsymbol{H}}_{t+1} + \bar{\rho}_t \boldsymbol{I}_d, \qquad \bar{\rho}_t \ge 0,$$

we have $\boldsymbol{H}_{t+1} \succeq \widetilde{\boldsymbol{H}}_{t+1}$. Moreover, since $\bar{\rho}_t \le \lambda\varepsilon_m$ by Lemma 1 and $\lambda \boldsymbol{I}_d \preceq \widetilde{\boldsymbol{H}}_{t+1}$, it follows that

$$\boldsymbol{H}_{t+1} = \widetilde{\boldsymbol{H}}_{t+1} + \bar{\rho}_t \boldsymbol{I}_d \preceq \widetilde{\boldsymbol{H}}_{t+1} + \lambda\varepsilon_m \boldsymbol{I}_d \preceq (1+\varepsilon_m)\widetilde{\boldsymbol{H}}_{t+1}.$$

Since Loewner order reverses under matrix inversion, the above display implies

$$\widetilde{\boldsymbol{H}}_{t+1}^{-1} \preceq (1+\varepsilon_m)\boldsymbol{H}_{t+1}^{-1}.$$

Next, by the exact monotonicity of the full Hessian sequence $\boldsymbol{H}_{t+1} \succeq \boldsymbol{H}_t$, and using

$$\boldsymbol{H}_t = \widetilde{\boldsymbol{H}}_t + \bar{\rho}_{t-1}\boldsymbol{I}_d \succeq \widetilde{\boldsymbol{H}}_t,$$

Chaining the three inequalities yields

$$\widetilde{\boldsymbol{H}}_{t+1}^{-1} \preceq (1+\varepsilon_m)\boldsymbol{H}_{t+1}^{-1} \preceq (1+\varepsilon_m)\boldsymbol{H}_t^{-1} \preceq (1+\varepsilon_m)\widetilde{\boldsymbol{H}}_t^{-1}.$$

Taking quadratic forms on both sides gives

$$\mathbf{x}^\top \widetilde{\boldsymbol{H}}_{t+1}^{-1}\mathbf{x} \le (1+\varepsilon_m)\mathbf{x}^\top \widetilde{\boldsymbol{H}}_t^{-1}\mathbf{x},$$

which completes the proof. □

Lemma 7 (FD-sketched Weighted Elliptical Potential)**.** *Following Assumptions 1–3, define*

$$\Lambda_m^T := d\ln(1+\varepsilon_m) + m\ln\left(1+\frac{L_\mu T}{m\lambda g(\tau)}\right). \tag{32}$$

Then

$$\sum_{t=1}^{T}\left\|\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t\right\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \le 2(1+\varepsilon_m)\left(1+\frac{L_\mu}{g(\tau)}\right)\Lambda_m^T.$$

Proof. We begin by relating the sketched and full inverse Hessian norms. By Lemma 6, we have

$$\widetilde{\boldsymbol{H}}_t^{-1} \preceq (1+\varepsilon_m)\boldsymbol{H}_t^{-1}.$$

Taking quadratic forms with $\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t$ on both sides yields

$$\left\|\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t\right\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \le (1+\varepsilon_m)\left\|\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t\right\|^2_{\boldsymbol{H}_t^{-1}}.$$

Thus, to control the sketched potential, it suffices to bound the corresponding full-Hessian potential. We invoke the standard elliptical potential in Lemma 15, Appendix D, which gives

$$\begin{aligned}\sum_{t=1}^{T}\left\|\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t\right\|^2_{\boldsymbol{H}_t^{-1}} &\le 2\left(1+\max_{t\in[T]}\left\|\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t\right\|_2^2\right)\ln\frac{\det(\boldsymbol{H}_{T+1})}{\det(\lambda\boldsymbol{I}_d)}\\ &\le 2\left(1+\frac{L_\mu}{g(\tau)}\right)\ln\frac{\det(\boldsymbol{H}_{T+1})}{\det(\lambda\boldsymbol{I}_d)}.\end{aligned}$$

The second line uses the uniform bound

$$\left\|\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t\right\|_2^2 = w_t\|\boldsymbol{x}_t\|_2^2 \le \frac{L_\mu}{g(\tau)},$$

which follows from Assumptions 1–3. From the FD recursion in Algorithm 1, we have $\boldsymbol{H}_{T+1} = \widetilde{\boldsymbol{H}}_{T+1} + \bar{\rho}_T\boldsymbol{I}_d$. Let $\widetilde{\lambda}_1,\dots,\widetilde{\lambda}_m$ denote the nonzero eigenvalues of $\widetilde{\boldsymbol{H}}_{T+1}-\lambda\boldsymbol{I}_d$, equivalently, the eigenvalues contributed by the sketch term $\boldsymbol{B}_T^\top\boldsymbol{B}_T$. Then

$$\begin{aligned}\ln\frac{\det(\boldsymbol{H}_{T+1})}{\det(\lambda\boldsymbol{I}_d)} &= \sum_{i=1}^{m}\ln\left(1+\frac{\widetilde{\lambda}_i+\bar{\rho}_T}{\lambda}\right)+(d-m)\ln\left(1+\frac{\bar{\rho}_T}{\lambda}\right)\\ &\le m\ln\left(1+\frac{\bar{\rho}_T}{\lambda}+\frac{\mathrm{Tr}(\boldsymbol{B}_T^\top\boldsymbol{B}_T)}{m\lambda}\right)+(d-m)\ln(1+\varepsilon_m)\\ &\le d\ln(1+\varepsilon_m)+m\ln\left(1+\frac{L_\mu T}{m\lambda g(\tau)}\right) =: \Lambda_m^T.\end{aligned}$$

Here, the first inequality uses the AM–GM inequality to upper bound the sum over the m nonzero eigenvalues, and the second uses $\bar{\rho}_T/\lambda \le \varepsilon_m$ from Lemma 1 and $\mathrm{Tr}(\boldsymbol{B}_T^\top\boldsymbol{B}_T) \le L_\mu T/g(\tau)$.

Combining the above inequalities, and then summing over $t=1,\dots,T$, we obtain

$$\sum_{t=1}^{T}\left\|\sqrt{w_s(\boldsymbol{\theta}_{t+1})}\boldsymbol{x}_t\right\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \le (1+\varepsilon_m)\cdot 2\left(1+\frac{L_\mu}{g(\tau)}\right)\Lambda_m^T,$$

which is exactly the claimed bound. □

Lemma 8 (FD-sketched Unweighted Leverage Scores)**.** *Under the same notation as in Lemma 7,*

$$\sum_{t=1}^{T}\|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \le 2\kappa g(\tau)(1+\varepsilon_m)\left(1+\frac{L_\mu}{g(\tau)}\right)\Lambda_m^T,$$

where

$$\kappa = \frac{1}{c_\mu} = \frac{1}{\inf_{\boldsymbol{x}\in\cup_{t=1}^T\mathcal{X}_t,\,\boldsymbol{\theta}\in\Theta}\mu'(\boldsymbol{x}^\top\boldsymbol{\theta})}.$$

Proof. We relate the unweighted leverage score to the weighted one through the lower bound on the curvature weight w_t. By Assumption 2, we have

$$w_t = \frac{\mu'(\boldsymbol{x}_t^\top \boldsymbol{\theta}_{t+1})}{g(\tau)} \geq \frac{c_\mu}{g(\tau)} = \frac{1}{\kappa g(\tau)}.$$

Therefore, for each $t \in [T]$,

$$\|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} = \frac{1}{w_t} \, \|\sqrt{w_t}\, \boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \leq \kappa g(\tau) \, \|\sqrt{w_t}\, \boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}}.$$

Summing the above inequality over $t = 1, \ldots, T$ gives

$$\sum_{t=1}^{T} \|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \leq \kappa g(\tau) \sum_{t=1}^{T} \|\sqrt{w_t}\, \boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}}.$$

We now apply Lemma 7, which yields

$$\sum_{t=1}^{T} \|\sqrt{w_t}\, \boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \leq 2(1+\varepsilon_m) \left(1 + \frac{L_\mu}{g(\tau)}\right) \Lambda_m^T.$$

Substituting this bound into the previous display, we obtain

$$\sum_{t=1}^{T} \|\boldsymbol{x}_t\|^2_{\widetilde{\boldsymbol{H}}_t^{-1}} \leq \kappa g(\tau) \cdot 2(1+\varepsilon_m) \left(1 + \frac{L_\mu}{g(\tau)}\right) \Lambda_m^T,$$

which completes the proof. □

D Omitted Standard Lemmas

In this section, we gather several standard auxiliary results from prior works that appears in our proof. For completeness, we state them here in a unified form, while omitting their proofs.

Lemma 9 (Woodbury and Sherman–Morrison Identities)**.** *Let $\boldsymbol{A} \in \mathbb{R}^{d\times d}$, $\boldsymbol{C} \in \mathbb{R}^{k\times k}$, $\boldsymbol{U} \in \mathbb{R}^{d\times k}$, and $\boldsymbol{V} \in \mathbb{R}^{k\times d}$. Assume that $\boldsymbol{A}$ and $\boldsymbol{C}$ are invertible. Then,*

$$(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{C}\boldsymbol{V})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}\big(\boldsymbol{C}^{-1} + \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}\big)^{-1}\boldsymbol{V}\boldsymbol{A}^{-1}.$$

In particular, for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$, if $1 + \boldsymbol{v}^\top \boldsymbol{A}^{-1}\boldsymbol{u} \neq 0$, then

$$(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u}}.$$

Lemma 10 (Lemma 1 of Zhang et al. (2025))**.** *Let $f : \Theta \to \mathbb{R}$ be convex and differentiable and $\boldsymbol{A} \in \mathbb{R}^{d\times d}$ be a symmetric positive definite matrix. For the OMD-type update $\boldsymbol{\theta}_{t+1} = \arg\min_{\boldsymbol{\theta}\in\Theta} f(\boldsymbol{\theta}) + \frac{1}{2\eta}\|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|^2_{\boldsymbol{A}}$, the following holds for all $\mathbf{u} \in \Theta$:*

$$\|\boldsymbol{\theta}_{t+1} - \mathbf{u}\|^2_{\boldsymbol{A}} \leq 2\eta\langle \nabla f(\boldsymbol{\theta}_{t+1}),\, \mathbf{u} - \boldsymbol{\theta}_{t+1}\rangle + \|\boldsymbol{\theta}_t - \mathbf{u}\|^2_{\boldsymbol{A}} - \|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+1}\|^2_{\boldsymbol{A}}.$$

Lemma 11 (Lemma 9 of Faury et al. (2020))**.** *Let $\mu : \mathbb{R} \to \mathbb{R}$ be a strictly increasing function, and let $\mathcal{Z} \subset \mathbb{R}$ be a bounded interval. Suppose that, for some constant $R > 0$,*

$$|\mu''(z)| \leq R\,\mu'(z), \qquad \forall z \in \mathcal{Z}.$$

Then, for any $z_1, z_2 \in \mathcal{Z}$ and either choice of $z \in \{z_1, z_2\}$, it holds that

$$\int_0^1 \mu'\big(z_1 + \nu(z_2 - z_1)\big)\, d\nu \geq \frac{\mu'(z)}{1 + R|z_1 - z_2|},$$

and, moreover,

$$\frac{\mu'(z)}{2 + R|z_1 - z_2|} \leq \int_0^1 (1-\nu)\,\mu'\big(z_1 + \nu(z_2 - z_1)\big)\, d\nu \leq \exp\big(R^2|z_1 - z_2|^2\big)\,\mu'(z).$$

Lemma 12 (Lemma 5 of Zhang et al. (2025)). *Let $\{\mathcal{F}_t\}_{t\geq 1}$ be the filtration defined by*

$$\mathcal{F}_t := \sigma\big(\{(\boldsymbol{x}_s, r_s)\}_{s=1}^{t-1}\big)\,.$$

Let $\{P_t\}_{t\geq 1}$ be a stochastic process such that, for each $t \geq 1$, P_t is an $\mathcal{F}_t$-measurable random probability distribution over $\mathbb{R}^d$. Assume further that the loss function ℓ_t, defined in (1), is $\mathcal{F}_{t+1}$-measurable. For each $t \geq 1$, define

$$L_t(\theta_\star) := \sum_{s=1}^{t} \ell_s(\theta_\star), \qquad F_t := -\sum_{s=1}^{t} \ln\Big(\mathbb{E}_{\theta\sim P_s}\Big[e^{-\ell_s(\theta)}\Big]\Big)\,.$$

Then, for any $\delta \in (0, 1]$, it holds that

$$\Pr\left(\forall t \geq 1,\ L_t(\theta_\star) \leq F_t + \log\left(\frac{1}{\delta}\right)\right) \geq 1 - \delta.$$

Lemma 13 (Variational Representation of the Mix Loss (Reid et al., 2015)). *Let P be a probability distribution over $\mathbb{R}^d$, and let Δ denote the set of all probability distributions over $\mathbb{R}^d$. For any loss function $\ell : \mathbb{R}^d \to \mathbb{R}$ and any $\alpha > 0$, we have*

$$-\frac{1}{\alpha}\log\Big(\mathbb{E}_{\theta\sim P}\Big[e^{-\alpha\ell(\theta)}\Big]\Big) = \mathbb{E}_{\theta\sim P_*}[\ell(\theta)] + \frac{1}{\alpha}\mathrm{KL}(P_*\|P),$$

where

$$P_* = \arg\min_{P'\in\Delta} \mathbb{E}_{\theta\sim P'}[\ell(\theta)] + \frac{1}{\alpha}\mathrm{KL}(P'\|P)$$

is the optimal solution. Furthermore, Zhang et al. (2025) showed that, for any probability distribution Q over $\mathbb{R}^d$,

$$-\frac{1}{\alpha}\log\Big(\mathbb{E}_{\theta\sim P}\Big[e^{-\alpha\ell(\theta)}\Big]\Big) = \mathbb{E}_{\theta\sim Q}[\ell(\theta)] + \frac{1}{\alpha}\mathrm{KL}(Q\|P) - \frac{1}{\alpha}\mathrm{KL}(Q\|P_*).$$

Lemma 14 (Theorem 1.8.2 of Ihara (1993)). *Let $P = \mathcal{N}(\boldsymbol{u}_p, \Sigma_p)$ and $Q = \mathcal{N}(\boldsymbol{u}_q, \Sigma_q)$ be two d-dimensional Gaussian distributions, where $\Sigma_p, \Sigma_q \in \mathbb{R}^{d\times d}$ are positive definite covariance matrices. Then*

$$\mathrm{KL}(Q\|P) = \frac{1}{2}\left(\log\frac{\det(\Sigma_p)}{\det(\Sigma_q)} + \mathrm{Tr}(\Sigma_p^{-1}\Sigma_q) + \|\boldsymbol{u}_p - \boldsymbol{u}_q\|^2_{\Sigma_p^{-1}} - d\right).$$

Lemma 15 (Lemma 9 of Faury et al. (2022)). *Let $\lambda \geq 1$, and let $\{\boldsymbol{z}_s\}_{s=1}^{\infty} \subset \mathbb{R}^d$ be a sequence satisfying $\|\boldsymbol{z}_s\|_2 \leq Z$ for all $s \in \mathbb{N}$. For each $t \geq 2$, define $\boldsymbol{V}_t := \lambda\boldsymbol{I}_d + \sum_{s=1}^{t-1}\boldsymbol{z}_s\boldsymbol{z}_s^\top$. Then, for any $T \geq 1$,*

$$\sum_{t=1}^{T}\|\boldsymbol{z}_t\|^2_{\boldsymbol{V}_t} \leq 2d(1+Z^2)\log\left(1 + \frac{TZ^2}{d\lambda}\right).$$

E Omitted Algorithms

This section collects algorithmic details omitted from the main text for clarity of presentation. We provide two ingredients needed for an efficient implementation of SGLB: an $O(md)$ procedure for the sketched projection step in Appendix E.1, and a buffered FD implementation that reduces the amortized update cost to $O(md)$ in Appendix E.2. Together, these results make explicit how the sketch-aware structure translates into practical computational gains.

E.1 Efficient Projection by Sketching

We now show that the projection step in SGLB can be solved in sketched form. For reference, the non-sketched GLB-OMD update in (2) can be written as

$$\boldsymbol{\xi}_t = \boldsymbol{\theta}^{\texttt{OMD}}_{t-1} - \eta\overline{\boldsymbol{H}}^{-1}_{t-1}\nabla\ell_{t-1}\big(\boldsymbol{\theta}^{\texttt{OMD}}_{t-1}\big)\,, \qquad \boldsymbol{\theta}^{\texttt{OMD}}_t = \arg\min_{\boldsymbol{\theta}\in\Theta}\|\boldsymbol{\theta} - \boldsymbol{\xi}_t\|^2_{\overline{\boldsymbol{H}}_{t-1}},$$

where $\overline{\boldsymbol{H}}_{t-1} = \boldsymbol{H}_{t-1} + \eta\nabla^2\ell_{t-1}\big(\boldsymbol{\theta}_{t-1}^{\texttt{OMD}}\big)$ is the augmented Hessian. Since $\Theta := \{\boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta}\|_2 \le S\}$ is a Euclidean ball and $\overline{\boldsymbol{H}}_{t-1} \succeq 0$, this projection can be computed via the singular value decomposition of $\overline{\boldsymbol{H}}_{t-1}$, which incurs an $O(d^3)$ time; see Section 4.1 of Mhammedi et al. (2019).

We next show that the corresponding projection in SGLB admits an $O(md)$ implementation due to the sketching low-rank structure. The sketched OMD update in (6) takes the form

$$\boldsymbol{\xi}_{t+1} = \boldsymbol{\theta}_t - \eta\widehat{\boldsymbol{H}}_t^{-1}\nabla\ell_t(\boldsymbol{\theta}_t) \quad \text{and} \quad \boldsymbol{\theta}_{t+1} = \operatorname*{arg\,min}_{\boldsymbol{\theta}\in\Theta} \|\boldsymbol{\theta} - \boldsymbol{\xi}_{t+1}\|^2_{\widehat{\boldsymbol{H}}_t},$$

where

$$\widehat{\boldsymbol{H}}_t := \widetilde{\boldsymbol{H}}_t + \eta\nabla^2\ell_t(\boldsymbol{\theta}_t) = \lambda\boldsymbol{I}_d + \boldsymbol{B}_t^\top\boldsymbol{B}_t + \frac{\eta\,\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_t)}{g(\tau)}\boldsymbol{x}_t\boldsymbol{x}_t^\top.$$

To simplify notation, we suppress the time index in the remainder of this subsection. Our derivation proceeds in two steps. We first identify the $(m+1)$-dimensional subspace on which $\widehat{\boldsymbol{H}}$ differs from the isotropic part $\lambda\boldsymbol{I}_d$. We then decompose the KKT system into this low-dimensional subspace and its orthogonal complement, reducing the projection to a one-dimensional root-searching problem.

By the KKT conditions, there exists a dual variable $\nu \ge 0$ such that

$$(\widehat{\boldsymbol{H}} + \nu\boldsymbol{I}_d)\boldsymbol{\theta} = \widehat{\boldsymbol{H}}\boldsymbol{\xi}, \qquad \nu\big(\|\boldsymbol{\theta}\|_2 - S\big) = 0. \tag{33}$$

If $\|\boldsymbol{\xi}\|_2 \le S$, then $\nu = 0$ and $\boldsymbol{\theta} = \boldsymbol{\xi}$. Otherwise, $\nu > 0$ and $\|\boldsymbol{\theta}\|_2 = S$.

Precomputation. Let $\boldsymbol{V} \in \mathbb{R}^{d\times m}$ denote the right singular matrix in the FD representation

$$\boldsymbol{B} = \sqrt{\boldsymbol{\Sigma}^2 - \rho\boldsymbol{I}_m}\,\boldsymbol{V}^\top.$$

We use $\boldsymbol{V}$ only to simplify the derivation; it need not be stored explicitly. Indeed, for any $\boldsymbol{x} \in \mathbb{R}^d$, $\boldsymbol{V}^\top\boldsymbol{x} = \boldsymbol{\Sigma}^{-1}\boldsymbol{B}\boldsymbol{x}$, so all required products can be computed from $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$ alone.

Define

$$\boldsymbol{a} := \sqrt{\eta\,\mu'(\boldsymbol{x}_t^\top\boldsymbol{\theta}_t)/g(\tau)}\,\boldsymbol{x}_t \in \mathbb{R}^d, \qquad \boldsymbol{p} := \boldsymbol{V}^\top\boldsymbol{a} \in \mathbb{R}^m,$$

and

$$\boldsymbol{q} := \boldsymbol{a} - \boldsymbol{V}\boldsymbol{p} = \boldsymbol{a} - \boldsymbol{B}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{p} \in \mathbb{R}^d.$$

Thus, $\boldsymbol{q}$ is the component of $\boldsymbol{a}$ orthogonal to $\mathrm{col}(\boldsymbol{B}^\top)$. Let $p_\perp := \|\boldsymbol{q}\|_2$, and if $p_\perp > 0$, define $\widehat{\boldsymbol{q}} := \boldsymbol{q}/p_\perp$. Then set

$$\boldsymbol{V}_+ = [\boldsymbol{V},\, \widehat{\boldsymbol{q}}] \in \mathbb{R}^{d\times(m+1)}, \qquad \boldsymbol{V}_+^\top\boldsymbol{V}_+ = \boldsymbol{I}_{m+1}.$$

When $p_\perp = 0$, we simply take $\boldsymbol{V}_+ = \boldsymbol{V}$, and all expressions below remain valid with $m+1$ replaced by m. In the basis induced by $\boldsymbol{V}_+$, the matrix $\widehat{\boldsymbol{H}}$ reduces to

$$\boldsymbol{M} := \boldsymbol{V}_+^\top\widehat{\boldsymbol{H}}\boldsymbol{V}_+ = \begin{pmatrix} \boldsymbol{\Sigma}^2 + (\lambda-\rho)\boldsymbol{I}_m & \\ & \lambda \end{pmatrix} + \begin{pmatrix} \boldsymbol{p} \\ p_\perp \end{pmatrix}\begin{pmatrix} \boldsymbol{p} \\ p_\perp \end{pmatrix}^\top \in \mathbb{R}^{(m+1)\times(m+1)},$$

while $\widehat{\boldsymbol{H}} = \lambda\boldsymbol{I}_d$ on $\mathrm{col}(\boldsymbol{V}_+)^\perp$.

Solving (33). Let $\boldsymbol{\varphi}_\xi := \boldsymbol{V}_+^\top\boldsymbol{\xi}$. Projecting (33) onto $\mathrm{col}(\boldsymbol{V}_+)$ yields

$$\boldsymbol{\alpha}(\nu) := \boldsymbol{V}_+^\top\boldsymbol{\theta}(\nu) = (\boldsymbol{M} + \nu\boldsymbol{I}_{m+1})^{-1}\boldsymbol{M}\,\boldsymbol{\varphi}_\xi \in \mathbb{R}^{m+1}.$$

On the orthogonal complement, we have

$$(\boldsymbol{I}_d - \boldsymbol{V}_+\boldsymbol{V}_+^\top)\boldsymbol{\theta}(\nu) = \frac{\lambda}{\lambda+\nu}(\boldsymbol{I}_d - \boldsymbol{V}_+\boldsymbol{V}_+^\top)\boldsymbol{\xi}.$$

Combining the two parts gives

$$\boldsymbol{\theta}(\nu) = \boldsymbol{V}_+\left(\boldsymbol{\alpha}(\nu) - \frac{\lambda}{\lambda+\nu}\boldsymbol{\varphi}_\xi\right) + \frac{\lambda}{\lambda+\nu}\boldsymbol{\xi}. \tag{34}$$

Hence, $\boldsymbol{\theta}(\nu)$ is obtained from a scalar contraction of intermediate variable $\boldsymbol{\xi}$ together with an $(m+1)$-dimensional correction in $\mathrm{col}(\boldsymbol{V}_+)$.

Algorithm 3 Fast SGLB

Input: Constant $R, L_\mu, S, \lambda > 0$, confidence level $\delta \in (0,1)$, sketch size $m \in \{1, \dots, d-1\}$.

Initialize $\boldsymbol{\theta}_1 \in \Theta := \{\boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta}\|_2 \le S\}$.
Initialize sketch $\boldsymbol{B}_1 \leftarrow \mathbf{0}^{2m \times d}$, $\boldsymbol{G}_1 \leftarrow \frac{1}{\lambda}\boldsymbol{I}_{2m}$.
for $t = 1$ **to** T **do**
  Compute the radius $\widetilde{\beta}_t(\delta)$ by (9).
  Select action $\boldsymbol{x}_t \in \mathcal{X}_t$ according to action selection rule in (10).
  Receive the reward r_t.
  Update sketched OMD estimator $\boldsymbol{\theta}_{t+1}$ by (6).
  Calculate $w_t \leftarrow \mu'(\boldsymbol{x}_t^\top \boldsymbol{\theta}_{t+1})/g(\tau)$.
  if sketch matrix $\boldsymbol{B}_t$ have $2m-1$ non-zero valued rows **then**
    Append $\sqrt{w_t}\boldsymbol{x}_t$ to the last row of $\boldsymbol{B}_t$
    Compute $[\boldsymbol{U}, \boldsymbol{\Sigma}, \boldsymbol{V}] \leftarrow \mathrm{SVD}(\boldsymbol{B}_t)$,where $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \cdots, \sigma_{2m})$
    Set $\rho \leftarrow \sigma_m^2$
    Update $\boldsymbol{B}_{t+1} \leftarrow \sqrt{\max(\boldsymbol{\Sigma}^2 - \rho \boldsymbol{I}_{2m}, 0)} \cdot \boldsymbol{V}^\top$
    Update $\boldsymbol{G}_{t+1} \leftarrow \mathrm{diag}\left\{\frac{1}{\lambda+\sigma_1^2-\rho}, \dots, \frac{1}{\lambda}\right\}$
  else
    Update $\boldsymbol{B}_{t+1} \leftarrow \begin{pmatrix} \boldsymbol{B}_t \\ \sqrt{w_t}\boldsymbol{x}_t \end{pmatrix}$
    Update $\boldsymbol{G}_{t+1}$ by (35).
  end if
end for

Output: Sketch-based estimator $\boldsymbol{\theta}_{T+1}$.

Searching for ν. When $\|\boldsymbol{\xi}\|_2 > S$, we need the unique $\nu > 0$ such that $\|\boldsymbol{\theta}(\nu)\|_2 = S$. Let

$$\boldsymbol{\beta}_\perp := (\boldsymbol{I}_d - \boldsymbol{V}_+ \boldsymbol{V}_+^\top)\boldsymbol{\xi}, \qquad \|\boldsymbol{\beta}_\perp\|_2^2 = \|\boldsymbol{\xi}\|_2^2 - \|\boldsymbol{\varphi}_\xi\|_2^2.$$

Then

$$f(\nu) := \|\boldsymbol{\theta}(\nu)\|_2^2 = \|\boldsymbol{\alpha}(\nu)\|_2^2 + \frac{\lambda^2}{(\lambda+\nu)^2}\|\boldsymbol{\beta}_\perp\|_2^2.$$

Since $f(\cdot)$ is convex and strictly decreasing on $[0, \infty)$, the equation $f(\nu) = S^2$ has a unique solution $\nu_\star$, which can be found by bisection or Newton's method over a scalar variable. Substituting $\nu_\star$ into (34) yields the desired projection.

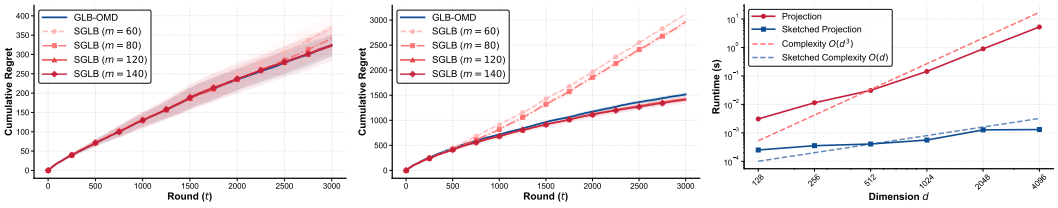
Complexity. The dominant cost comes from two matrix–vector stages: precomputation and reconstruction, both of which require $O(md)$ time. The search for ν only involves matrices of size at most $(m+1) \times (m+1)$ and is therefore costs $O(m)$. Consequently, the projection in (7) can be implemented in $O(md)$ time. In contrast, the corresponding full-matrix projection in (5) requires $O(d^3)$ time. Our solution can thus be viewed as a low-rank counterpart of Mhammedi et al. (2019).

E.2 Fast Algorithm of SGLB

The main computational bottleneck of the vanilla FD update in Algorithm 1 is the singular value decomposition. If an SVD is carried out after every newly arrived row, the update cost is $O(m^2 d)$ per round. As noted in prior work (Liberty, 2013; Ghashami et al., 2016; Luo et al., 2016; Kuzborskij et al., 2019), this cost can be reduced to $O(md)$ amortized time via the standard doubling-space implementation of FD, without violating the spectral approximation guarantee in Lemma 1.

Motivated by this idea, we develop a fast implementation of SGLB. Instead of compressing the sketch after every new observation, we buffer the newly arrived curvature-weighted rows and perform the FD shrinking step only once every m insertions. Equivalently, between two consecutive compression steps, the sketch is allowed to grow temporarily to at most $2m$ rows. This avoids repeated SVD computations while preserving the low-rank structure needed by sketch-aware GLBs.

Within such an epoch, let $\boldsymbol{B}_t$ denote the current buffered sketch matrix and define the gram matrix $\boldsymbol{G}_t := (\boldsymbol{B}_t \boldsymbol{B}_t^\top + \lambda \boldsymbol{I})^{-1}$. When the new row $\sqrt{w_t}\boldsymbol{x}_t$ is appended and no compression is triggered,



(a) Logistic bandits ($d = 500$) (b) Poisson bandits ($d = 500$) (c) Projection time

Figure 4: (a), (b): Cumulative regret of the compared algorithms on synthetic data; (c): Projection time comparison of the compared solvers on synthetic data w.r.t. ambient dimension d.

$\boldsymbol{G}_t$ can be updated recursively by block matrix inversion:

$$\boldsymbol{G}_{t+1} = \begin{pmatrix} \boldsymbol{G}_t + \frac{\boldsymbol{\phi}_t \boldsymbol{\phi}_t^\top}{c} & -\frac{\boldsymbol{\phi}_t}{c} \\ -\frac{\boldsymbol{\phi}_t^\top}{c} & \frac{1}{c} \end{pmatrix}, \tag{35}$$

where

$$\boldsymbol{\phi}_t = \sqrt{w_t}\, \boldsymbol{G}_t \boldsymbol{B}_t \boldsymbol{x}_t, \qquad c = w_t \boldsymbol{x}_t^\top \boldsymbol{x}_t - \sqrt{w_t}\, \boldsymbol{x}_t^\top \boldsymbol{B}_t^\top \boldsymbol{\phi}_t + \lambda,$$

and $w_t := \mu'(\boldsymbol{x}_t^\top \boldsymbol{\theta}_{t+1})/g(\tau)$. Thus, as long as no compression is performed, both the estimator update in (6) and the bonus evaluation in (10) can still be carried out efficiently using the maintained sketched inverse structure and Sherman-Morrison formula.

The overall implementation is summarized in Algorithm 3. Because the buffered sketch has at most $2m$ rows, the matrix $\boldsymbol{G}_t$ has size at most $2m \times 2m$ throughout each epoch. Therefore, each non-compression update only requires matrix–vector multiplications and a rank-one-style inverse update on a matrix of dimension at most $2m$, which costs $O(md)$ amortized time. Since the expensive SVD step is invoked only once every m insertions, its amortized cost is also $O(md)$. Consequently, the fast implementation preserves the $O(md)$ amortized update complexity of SGLB.

F Omitted Details for Section 5

In this section, we provide the omitted details of experiments in section 5. Unless otherwise stated, SGLB employ the efficient implementations described in Appendix E.2. We provide the experimental setups and configurations in Appendix F.1 and additional experiments in Appendix F.2, F.3.

F.1 Experimental Setups

All the experiments in this work were conducted on a machine equipped with a 24-core Intel(R) Xeon(R) Gold 6240R 2.40GHz CPU and 256 GB of memory. All methods were implemented in Python using `NumPy`, with pseudo-inverse and singular value decomposition computed via `np.linalg.pinv` and `np.linalg.svd`, respectively. Unless otherwise stated, each configuration was repeated over 10 independent runs with different random seeds. We report the mean cumulative regret, and the shaded regions in the regret plots indicate ± 1 standard deviation across runs. Wall-clock time was measured using `time.perf_counter()`, and all running-time comparisons are presented on a logarithmic scale.

Configuration. Throughout the experiments, all algorithmic parameters were set according to their theoretical prescriptions, except for the regularization parameter λ. For methods that require explicit regularization, including GLM-UCB (Filippi et al., 2010), GLOC (Jun et al., 2017), ECOLog (Faury et al., 2022), RS-GLinUCB (Sawarni et al., 2024), GLB-OMD (Zhang et al., 2025), and our work SGLB, we adopted a unified tuning protocol and selected λ from the grid $\{10^{-2}, 10^{-1}, 1, 10, 10^2\}$. For methods that do not use a fixed explicit regularization parameter, such as OFUGLB (Lee et al., 2024), we followed the settings prescribed in the previous papers. Following Zhang et al. (2025), we additionally applied a common empirical scaling to the UCB bonus of all methods on real-world datasets to better balance exploration and exploitation.

Table 2: Summary of real-world datasets and experimental configurations. Here, target class denotes the class treated as the positive label, and K denotes the candidate cluster size.

Dataset	OpenML ID	Instances	Features	Classes	Target Class	K
MNIST	554	60,000	784	10	3	100
Covertype	150	581,012	54	7	2	60
MFeat	22	2000	48	10	3	50
W8A	1588	64,700	300	2	1	60

F.2 Additional Results on Synthetic Data

Dataset Construction. We compare methods on a controlled synthetic benchmark designed to isolate the effect of matrix sketching. We consider ambient dimensions $d = \{150, 200, 500\}$. The unknown parameter $\boldsymbol{\theta}_\star$ is sampled uniformly from the d-dimensional sphere of radius $S = 5$. To generate actions with a shared low-rank structure, we first sample a fixed orthonormal basis $\boldsymbol{V} \in \mathbb{R}^{d\times r}$ with latent rank $r = 100$. At each round t, we draw $K = 20$ Gaussian coefficient vectors $\boldsymbol{c}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_r)$, form candidate actions $\boldsymbol{x}_i = \boldsymbol{V}\boldsymbol{c}_i \in \mathbb{R}^d$, and normalize them to the unit ball.

We provide additional experimental results on synthetic data. Figures 4a and 4b report the cumulative regret curves for logistic and Poisson bandits with ambient dimension $d = 500$, which were omitted from Section 5.1 for space reasons. For logistic bandits, all SGLB variants remain close to GLB-OMD throughout the horizon. Even the smaller sketches $m = 60$ and $m = 80$ incur only a moderate regret increase, indicating that SGLB remains effective under aggressive compression in this setting. By contrast, for Poisson bandits, the performance gap becomes much more sensitive to the sketch size. While $m = 120$ and $m = 140$ still track GLB-OMD reasonably well, the smaller sketches $m = 60$ and $m = 80$ accumulate substantially larger regret over time. Overall, these supplementary results reinforce the same trend observed in the main text: the bounded logistic setting is more tolerant to aggressive sketching, whereas the more challenging unbounded Poisson setting requires a larger sketch size to stay close to the non-sketched baseline.

Projection Cost. We benchmark the computational efficiency of the proposed sketched projection routine against a full-matrix baseline for the weighted Euclidean projection subproblem. The baseline explicitly forms the full $d \times d$ matrix $\boldsymbol{H}$ in the projection step of (5), computes its eigendecomposition, and solves the resulting KKT condition by bisection over the dual variable. In contrast, our solver exploits the low-rank structure of $\widetilde{\boldsymbol{H}}$, as detailed in Appendix E.1.

For each ambient dimension d, we generate a random projection instance as follows. We first sample a Gaussian matrix in $\mathbb{R}^{d\times m}$ and orthonormalize it via QR decomposition to obtain $\boldsymbol{U}$. We then draw singular values σ_i independently from the uniform distribution on $[0.5, 2.0]$ and set $\boldsymbol{B} = \mathrm{diag}(\sigma)\boldsymbol{U}^\top$. Next, we sample $\boldsymbol{a}, \boldsymbol{\zeta} \in \mathbb{R}^d$ independently from the standard normal distribution, and rescale $\boldsymbol{\zeta}$ so that $\|\boldsymbol{\zeta}\|_2 \in [2S, 3S]$, ensuring that the projection constraint is active. We uniformly set $\lambda = 1$, $S = 1$, sketch size $m = 16$, and error tolerance $|\|\boldsymbol{\theta}(\nu)\|_2^2 - S^2| < 1 \times 10^{-10}$.

The results are shown in Figure 4c. As d increases, the sketched projection solver is consistently much faster than the original projection solver, and the runtime gap widens rapidly in high dimensions. This empirical behavior is consistent with our theoretical analysis. When the sketch size m is much smaller than d, the sketched projection solver only involves operations on the compressed geometry and therefore admits $O(d)$ complexity, whereas the original projection step requires $O(d^3)$ full-matrix computation. These results confirm that sketching substantially reduces the computational bottleneck of the projection step, especially in high-dimensional regimes.

F.3 Additional Results on Real-World Data

In this section, we report additional experimental results on real-world benchmarks beyond MNIST. Together with the results in Section 5.2, these experiments further demonstrate that the proposed sketch-aware method achieves a favorable balance between statistical performance and computational efficiency on real-world data.

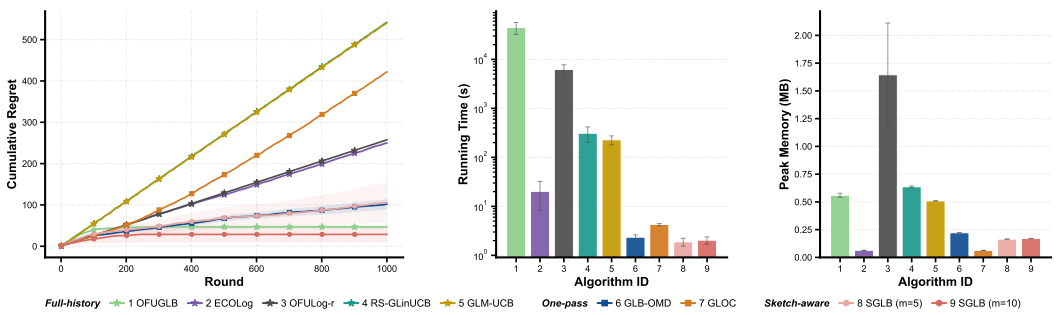


Figure 5: Comparison of full-history, one-pass, and sketch-aware GLB algorithms on `Covertype`. We report cumulative regret (left), total running time (middle), and peak memory usage (right).

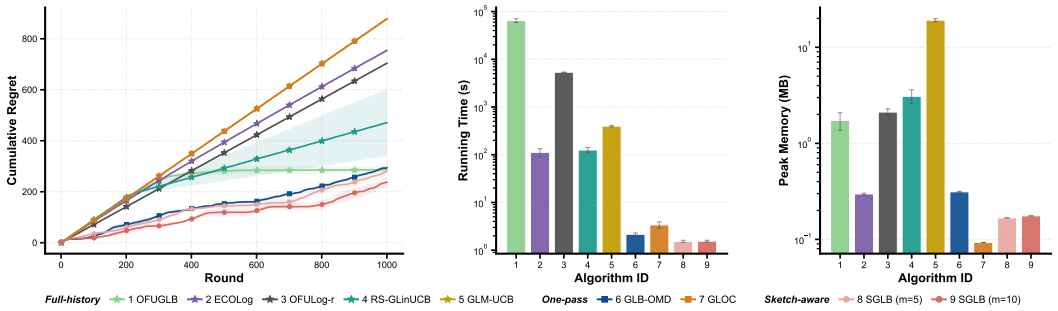


Figure 6: Comparison of full-history, one-pass, and sketch-aware GLB algorithms on `MFeat`. We report cumulative regret (left), total running time (middle), and peak memory usage (right).

Dataset Construction. We conduct experiments on four publicly available multiclass classification datasets from the OpenML repository (Vanschoren et al., 2013), namely `MNIST`, `Covertype`, `MFeat`, and `W8A`; their basic statistics are summarized in Table 2. Following experiments in Filippi et al. (2010), we center and standardize the features and append a constant covariate. We then partition the data into K clusters via unsupervised K-means, and use the resulting cluster centroids as the arm contexts in ambient dimension d. For the logistic reward model, we binarize the labels by assigning reward 1 to samples belonging to the target class and reward 0 otherwise. The reward associated with each arm is then defined as the average binary reward of the samples assigned to its corresponding cluster. We evaluated the compared algorithms with running each method over 5 independent trials and averaging the results to report the regret, the running time and peak memory.

More Baselines. As in Section 5.2, we organize the compared GLB algorithms into three categories. Compared with `MNIST`, the benchmarks considered in this section have smaller ambient dimension d, which allows us to include additional computationally intensive GLB baselines. Specifically, we compare representative methods from three update regimes: *full-history* methods, including GLM-UCB (Filippi et al., 2010), ECOLog (Faury et al., 2020), RS-GLinCB (Sawarni et al., 2024), OFUGLB (Lee et al., 2024), and the logistic-bandit-specific method OFULog-r (Abeille et al., 2021); *one-pass* methods, including GLOC (Jun et al., 2017) and GLB-OMD (Zhang et al., 2025); and the proposed *sketch-aware* method SGLB with varying sketch sizes m.

Results. Figures 5, 6 and 7 summarize the cumulative regret, running time, and peak memory usage of all compared methods on the additional real-world benchmarks. Across both datasets, one-pass and sketch-aware methods achieve the most favorable efficiency: GLB-OMD, GLOC, and the proposed SGLB are consistently faster and more memory-efficient than full-history baselines, while the two SGLB variants remain among the lowest-cost methods overall. In `W8A` dataset, OFULog-r achieves the lowest regret curves, but its running time is four orders of magnitude faster than SGLB, and it requires two orders of magnitude more memory. Compared with the proposed SGLB, GLOC can be slightly more memory-efficient on lower-dimensional instances. However, this advantage is

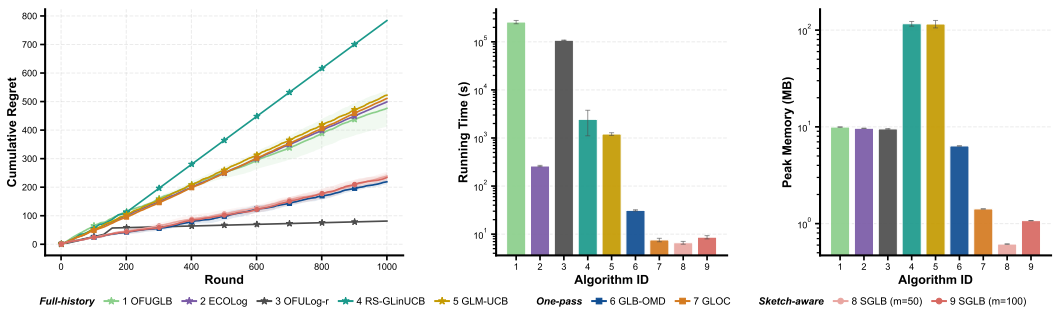


Figure 7: Comparison of full-history, one-pass, and sketch-aware GLB algorithms on `W8A`. We report cumulative regret (left), total running time (middle), and peak memory usage (right).

accompanied by weaker statistical performance: unlike OMD-based methods, GLOC relies on an online-learning-to-confidence-set construction and does not match the improved regret guarantees.