Proceedings of Machine Learning Research vol 134:1–60, 2021 34th Annual Conference on Learning Theory

Efficient Bandit Convex Optimization: Beyond Linear Losses

Arun Sai Suggala ASUGGALA@CS.CMU.EDU
Carnegie Mellon University

Pradeep Ravikumar PRADEEPR@CS.CMU.EDU
Carnegie Mellon University

Praneeth Netrapalli PNETRAPALLI@GOOGLE.COM
Google Research India

Editors: Mikhail Belkin and Samory Kpotufe

Abstract

We study the problem of online learning with bandit feedback, where a learner aims to minimize a sequence of adversarially generated loss functions, while only observing the value of each function at a single point. When the loss functions chosen by the adversary are convex and quadratic, we develop a new algorithm which achieves the *optimal* regret rate of $\tilde{O}\left(T^{1/2}\right)$. Furthermore, our algorithm satisfies three important desiderata: (a) it is practical and can be efficiently implemented for high dimensional problems, (b) the regret bound holds with high probability even against *adaptive adversaries* whose decisions can depend on the learner's previous actions, and (c) it is *robust* to model mis-specification; that is, the regret bound degrades gracefully when the loss functions deviate from convex quadratics. To the best of our knowledge, ours is the first algorithm for bandit convex optimization with quadratic losses which is efficiently implementable and achieves optimal regret guarantees. Existing algorithms for this problem either have sub-optimal regret guarantees (Flaxman et al., 2004; Saha and Tewari, 2011) or are computationally expensive and do not scale well to high-dimensional problems (Bubeck et al., 2017).

Our algorithm is a bandit version of the classical regularized Newton's method. It involves estimation of gradients and Hessians of the loss functions from single point feedback. A key caveat of working with Hessian estimates however is that they typically have large variance. In this work, we show that it is nonetheless possible to finesse this caveat by relying on the idea of "focus regions", where we restrict the algorithm to choose actions from a subset of the action space in which the variance of our estimates can be controlled.

Keywords: Online Learning, Bandit Feedback, Newton's Method, Quadratic Losses

1. Introduction

In this work, we study the problem of online learning with bandit feedback, which can be viewed as a repeated game between a learner and an adversary. In round t of this game, the learner chooses an action $\mathbf{x}_t$ from a known domain $\mathcal{X} \subset \mathbb{R}^d$. The adversary simultaneously selects a loss function $f_t : \mathcal{X} \to \mathbb{R}$ and reveals the loss suffered by the learner $f_t(\mathbf{x}_t)$. The performance of the learner at the end of T rounds is measured using regret

$$\mathrm{Reg}_T = \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x}).$$

For this problem, we would like to design an algorithm for choosing $\mathbf{x}_t$ that satisfies the following key criteria: (a) (**optimal regret**) achieves regret bounds which have optimal dependence on T, and which hold in high-probability against adaptive adversaries, and (b) (**computational efficiency**) the run-time of *each iteration* of algorithm should have a small dependence on dimension d, *e.g.,* polynomial dependence with a small exponent, and *independent* of the number of rounds T (ideally, we would like it to have similar run-time as efficient algorithms for online learning in the full information setting).

The framework of bandit optimization is extremely general and has found numerous practical applications in fields such as computer science, economics, game theory, and medical decision making. Some of these applications include design of clinical trials, market pricing, online ad placement, and recommender systems (Kleinberg, 2005; Bubeck and Cesa-Bianchi, 2012; Hazan, 2016). Owing to its importance, there has been extensive work on designing low-regret algorithms for bandit optimization. Early works on this problem have focused on finite action space $\mathcal{X}$, in which case the problem is called multi-armed bandit (MAB) problem. This problem has been well studied and several efficient algorithms achieving the optimal high-probability regret rate of $O\left(T^{1/2}\right)$ have been developed (Auer et al., 2002; Audibert and Bubeck, 2010; Lee et al., 2020). Later works on bandit optimization have turned to continuous action spaces and convex loss functions. Seminal works along this line have developed online gradient descent style algorithms for regret minimization (Flaxman et al., 2004; Kleinberg, 2005). When the loss functions f_t are convex and bounded, these algorithms achieve $O\left(T^{5/6}\right)$ regret in *expectation.* Improving upon these regret guarantees has remained an open problem until the works of Hazan and Li (2016); Bubeck et al. (2017). The algorithms developed in these latter works achieve the optimal $\tilde{O}\left(T^{1/2}\right)$ regret, albeit they are computationally expensive and are not efficiently implementable in practice. In particular, the run time of the algorithm of Hazan and Li (2016) depends exponentially on the dimension, and the algorithm of Bubeck et al. (2017) involves minimization of an approximately convex function over a nonconvex set, which is non-trivial in practice[1].

Despite years of research, designing efficient algorithms for bandit convex optimization (BCO) has turned out to be challenging. This can be attributed to the extremely limited information available to the learner about the loss functions chosen by the adversary. Consequently, several works have focused on sub-cases of BCO. These works can be classified into two broad categories. One category imposes parametric assumptions such as linearity on the loss functions. The other category imposes structural assumptions such as strong convexity. The most popular parametric assumption that is studied in the literature is the linearity assumption (Abernethy et al., 2009). Recent works have designed efficient algorithms achieving optimal regret guarantees under this assumption (Lee et al., 2020). However, apart from linearity, to the best of our knowledge, no other parametric assumption has been considered in the literature. When it comes to structural assumptions, several works have considered assumptions such as Lipschitzness (Flaxman et al., 2004), smoothness (Saha and Tewari, 2011), and strong convexity, smoothness (Hazan and Levy, 2014; Ito, 2020). Perhaps surprisingly, among all these assumptions, computationally efficient and optimal algorithms are only known for strongly convex, smooth functions (Hazan and Levy, 2014). While these results are interesting, it should be noted that strong convexity is a restrictive assumption which rarely holds in practice. Consequently, it is important

1. Although Bubeck et al. (2017) present a modified algorithm for polytopes which can be implemented in polynomial time, the *per iteration* runtime of this algorithm has a large polynomial dependence on d and a linear dependence on T.

to relax this assumption. However, the oracle lower bounds of Hu et al. (2016) suggest that designing optimal algorithms for non-strongly convex, smooth functions might require new and different algorithmic techniques to those used in existing works. In particular, all existing works which design computationally efficient algorithms first estimate the gradient of the loss function from one-point feedback, and then use Online Mirror Descent (OMD) style updates to choose the next action. The lower bounds of Hu et al. (2016) suggest that such techniques will not be able to achieve the optimal $\tilde{O}\left(T^{1/2}\right)$ regret for non-strongly convex, smooth functions. So, to make progress along this line, we need new algorithmic techniques. Unfortunately, it is unclear how to come up with such techniques for general convex functions.

This Work. In this work, we make progress on this problem by designing an efficient algorithm for convex, quadratic loss functions that achieves optimal high-probability regret guarantees against an adaptive adversary. To be precise, our algorithm achieves a regret of $\tilde{O}\left(d^{16}\sqrt{T}\right)$, which is known to be optimal in T (see Dani et al. (2007) for lower bounds on the regret). In terms of computation, the key computational bottleneck of our algorithm involves generating uniform samples from a convex set. This is a well studied problem and several efficient MCMC algorithms such as Hit-and-run algorithms have been developed for this problem (Lovász and Vempala, 2003; Belloni et al., 2015; Laddha et al., 2020). For action sets which are polytopes with m constraints, the amortized time complexity of each iteration of our algorithm is $\tilde{O}\left(\frac{m^2d^3+md^6}{T}+md^4+m^2d\right)$. In comparison, the only existing computationally efficient and optimal algorithm for this setting has a time complexity of $\tilde{O}\left(\mathrm{poly}(dm)T\right)$ with a much larger exponent on d (Bubeck et al., 2017). Moreover, the runtime of each iteration of this algorithm has a linear dependence on T, thus making it extremely inefficient for large T.

Furthermore, our algorithm is robust to model mis-specification: if each loss function f_t is ϵ-close to a convex, quadratic function in $\|\cdot\|_\infty$ norm, the regret of our algorithm is bounded by $\tilde{O}\left(\epsilon T + d^{16}\sqrt{T}\right)$. We believe robustness is necessary for algorithms which focus on sub-cases of BCO, as the assumptions on loss functions do not typically hold in practice. However, most existing works do not study this property. To the best of our knowledge, ours is the first algorithm for BCO with quadratic functions, which is computationally efficient, robust and achieves optimal regret guarantees.

Techniques. Our algorithm is a regularized Newton's method with self concordant barrier of $\mathcal{X}$ as the regularizer. It involves estimation of gradients and Hessians of the loss functions from single point feedback. This is unlike most existing computationally efficient algorithms which rely *only* on the gradient estimates to choose their actions (Saha and Tewari, 2011; Hazan and Levy, 2014). As previously mentioned, gradient information alone doesn't suffice to design algorithms achieving optimal regret for nonlinear losses (see Section 5.1 for empirical evidence). So, in this work, we estimate both the gradient and Hessian of the unknown loss function and use the estimates in a regularized Newton method. However, estimating the Hessian comes with its own challenges. The variance of the Hessian estimates is typically very large. Consequently, we need new techniques to cancel the effect of variance. In this work, we crucially rely on "focus regions" to handle the variance. This technique is inspired by Bubeck et al. (2017), who use similar ideas to design an optimal, albeit computationally inefficient, algorithm for general convex functions. At a high level, the variance of the Hessian estimates can only be controlled in a small region, which we call focus region. So, we restrict ourselves to this region and always choose actions within this region. However, the resulting algorithm only ensures low regret with respect to (w.r.t) points in the focus region. To ensure low regret even w.r.t points outside the focus region, we perform a test every iteration called "restart condition". Intuitively, this test checks if the minimizer of the cumulative loss over the entire domain falls well within the focus region. If yes, we continue the algorithm, as having a low regret w.r.t points in the focus region ensures the overall regret is low. The test fails when the minimizer gets too close to the boundary of the focus region. In this case, we show that the regret of our actions until now is negative, and restart the algorithm.

While the ideas of focus region and restart condition appeared in Bubeck et al. (2017), we note that new techniques are needed to make this approach computationally efficient. Restricting to quadratics doesn't automatically make Bubeck et al. (2017)'s approach computationally efficient. To make our algorithm efficient, we move away from the exponential weights update scheme used by Bubeck et al. (2017) and instead rely on Newton method and OMD framework. Moreover, we design a new test for the restart condition that is much more computationaly efficient than the test of Bubeck et al. (2017) (see Section 5 for more details).

Before we proceed, we note that our algorithm requires access to a self-concordant barrier (SCB) of $\mathcal{X}$ which satisfies certain assumption on the behavior of its Hessian (see Assumption 3). If $\mathcal{X} \subset \mathbb{R}$, then any SCB satisfies this property (see Proposition 4). Moreover, we show that the log-barrier of any polyhedral set satisfies this property. We conjecture that any SCB of any convex action set $\mathcal{X} \subset \mathbb{R}^d$ satisfies this property.

Paper Organization. Section 2 presents necessary background. Section 3 presents our main results. Section 4 discusses some of the related works. Section 5 presents our algorithm and Section 6 discusses the key ideas used in the algorithm. In Section 7 we discuss the computational aspects of our algorithm. We conclude with Section 8. Due to the lack of space, most proofs are presented in the appendix.

2. Problem Setting and Background

Notation. Throughout the paper, we denote vectors by bold-faced letters ($\mathbf{x}$), and matrices by capital letters (A). $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$ and $\|\cdot\|_A$ is the weighted Euclidean norm, *i.e.,* $\|\mathbf{x}\|_A = \langle A\mathbf{x}, \mathbf{x}\rangle^{1/2}$, where A is a positive definite matrix. We let $B_r(\mathbf{x})$ denote an ℓ_2 ball of radius r centered at $\mathbf{x}$, *i.e.,* $B_r(\mathbf{x}) = \{\mathbf{y} : \|\mathbf{y}-\mathbf{x}\| \leq r\}$. We let $B_{r,A}(\mathbf{x}) = \{\mathbf{y} : \|\mathbf{x}-\mathbf{y}\|_A \leq r\}$. For any strictly convex twice differentiable function f, we define the local norm at $\mathbf{x}$ as $\|\mathbf{v}\|_{\mathbf{x},f} = \left\langle \mathbf{v}, \nabla^2 f(\mathbf{x})\mathbf{v}\right\rangle^{1/2}$. $\partial\mathcal{X}$ denotes the boundary of a set $\mathcal{X}$. $b = \tilde{O}\left(a\right)$ implies $b \leq Ca\log a$ for a large enough constant C independent of a.

A function $f : \mathcal{X} \to \mathbb{R}$ is ϵ-close to a function $g : \mathcal{X} \to \mathbb{R}$ if $\sup_{\mathbf{x}\in\mathcal{X}} |f(\mathbf{x}) - g(\mathbf{x})| \leq \epsilon$. A function f is a quadratic function if it can be parameterized as $f(\mathbf{x}; A, \mathbf{b}, c) = \frac{1}{2}\langle \mathbf{x}, A\mathbf{x}\rangle + \langle \mathbf{b}, \mathbf{x}\rangle + c$, for some $A \in \mathbb{R}^{d\times d}, \mathbf{b} \in \mathbb{R}^d, c \in \mathbb{R}$. In addition, if $(A + A^T)$ is positive semi-definite, then f is called a *convex* quadratic function. Note that the set of *linear* functions is a subset of the set of convex quadratic functions. We let $\mathbb{E}_t$ denote the conditional expectation conditioned on all randomness in the first $t-1$ rounds. We use $\mathbb{B}^d, \mathbb{S}^{d-1}$ to denote the d-dimensional unit ball and unit sphere w.r.t Euclidean norm. We let $\mathbf{u} \sim \mathbb{B}^d, \mathbf{v} \sim \mathbb{S}^{d-1}$ denote the random variables chosen uniformly from these sets.

Problem Setting. In this work, we assume that the action space $\mathcal{X}$ is convex, compact and has non-empty interior. Without loss of generality, we assume $\mathcal{X}$ contains an Euclidean ball of radius 1, and has an ℓ_2 diameter of D, *i.e.,* $\sup_{\mathbf{x},\mathbf{y}\in\mathcal{X}} \|\mathbf{x}-\mathbf{y}\| \leq D$. We assume that each loss function f_t is ϵ-close to a convex quadratic function q_t which is bounded and Lipschitz, *i.e.,* $\sup_{\mathbf{x}\in\mathcal{X}} |q_t(\mathbf{x})| \leq B$, and for all $\mathbf{x},\mathbf{y}\in\mathcal{X}$, $|q_t(\mathbf{x}) - q_t(\mathbf{y})| \leq L\|\mathbf{x}-\mathbf{y}\|$. Finally, we assume the adversary is adaptive, *i.e.,* the decisions of the adversary can depend on the learner's previous actions.

2.1. One-point Gradient and Hessian Estimates

A major component of our algorithm involves estimating the gradient and Hessian of the unknown loss function from one-point feedback provided by the adversary. These estimates are then used in OMD to pick the next move of the learner. In this work, we rely on the following randomized sampling scheme to compute these estimates.

Proposition 1 *Let $f:\mathbb{R}^d \to \mathbb{R}$ be a quadratic function. Let $C \in \mathbb{R}^{d\times d}$ be any symmetric positive definite matrix. Then*

$$\begin{aligned}
\nabla f(\mathbf{x}) &= d\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1 f(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right], \\
\nabla^2 f(\mathbf{x}) &= \frac{d^2}{2}\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T + \mathbf{v}_2\mathbf{v}_1^T)C^{-1} f(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right].
\end{aligned}$$

To generate unbiased estimates of the gradient and Hessian of f at $\mathbf{x}$, we first randomly sample $\mathbf{v}_1$ and $\mathbf{v}_2$ from uniform distribution on $\mathbb{S}^{d-1}$, and get the one-point feedback from the adversary about $f(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)$, and then rely on the above proposition. We note that one can also rely on Gaussian smoothing to estimate this information (see Proposition 27 in Appendix). For the simplicity and clarity of analysis, in this work, we use the above sampling scheme instead of Gaussian smoothing. However, our algorithm and its analysis can be modified in a straightforward way to rely on Gaussian smoothing.

2.2. Self Concordant Barriers

Self Concordant Barriers (SCBs) play a crucial role in our algorithm and its analysis. So, in this section, we define SCB and present some of its useful properties.

Definition 2 *Let $\mathcal{X} \subseteq \mathbb{R}^d$ be a closed convex set with non-empty interior. A function $R : int(\mathcal{X}) \to \mathbb{R}$ is called a ν-self-concordant barrier of $\mathcal{X}$, if*

1. *(Barrier Property) R is three times continuously differentiable with $R(\mathbf{x}_k) \to \infty$ along every sequence $\{\mathbf{x}_k \in int(\mathcal{X})\}$ converging to a boundary point of $\mathcal{X}$, as $k \to \infty$*

2. *R satisfies the following for all $\mathbf{x} \in int(\mathcal{X}), h \in \mathbb{R}^d$,*

$$|\nabla^3 R(\mathbf{x})[h,h,h]| \leq 2|\nabla^2 R(\mathbf{x})[h,h]|^{3/2}, \quad |\langle \nabla R(\mathbf{x}), h\rangle| \leq \sqrt{\nu}|\nabla^2 R(\mathbf{x})[h,h]|^{1/2}.$$

Without loss of generality, we assume $\min_{\mathbf{x}\in\mathcal{X}} R(\mathbf{x}) = 0$. It is well known that R satisfies the following properties (see Appendix G for a more comprehensive review)

- **(P1)** *Dikin Ellipsoid*: For any $\mathbf{x} \in \text{int}(\mathcal{X})$, the Dikin ellipsoid centered at $\mathbf{x}$, $B_{1,\nabla^2 R(\mathbf{x})}(\mathbf{x})$, is entirely contained in $\mathcal{X}$.

- **(P2)** For any $\mathbf{x} \in \text{int}(\mathcal{X})$, and $\mathbf{y} \in B_{1,\nabla^2 R(\mathbf{x})}(\mathbf{x})$, we have

$$(1 - \|\mathbf{x}-\mathbf{y}\|_{\nabla^2 R(\mathbf{x})})^2 \nabla^2 R(\mathbf{x}) \preceq \nabla^2 R(\mathbf{y}) \preceq \frac{1}{(1 - \|\mathbf{x}-\mathbf{y}\|_{\nabla^2 R(\mathbf{x})})^2} \nabla^2 R(\mathbf{x}). \tag{1}$$

In this work, we assume that $\mathcal{X}$ has an SCB which satisfies the following additional property.

Assumption 3 *For any $\mathbf{x}, \mathbf{y} \in$ int$(\mathcal{X})$ such that $\|\mathbf{y}-\mathbf{x}\|_{\nabla^2 R(\mathbf{x})} \leq \lambda$*

$$\nabla^2 R(\mathbf{y}) \succeq \frac{1}{(1+\lambda)^2} \nabla^2 R(\mathbf{x}). \tag{2}$$

The following propositions show that a wide range of action spaces have SCBs which satisfy this property. We conjecture that any SCB satisfies this property.

Proposition 4 *Suppose $\mathcal{X} \subseteq \mathbb{R}$. Then any SCB of $\mathcal{X}$ satisfies Assumption 3.*

Proposition 5 *Suppose $\mathcal{X} \subseteq \mathbb{R}^d$ is polyhedral,* i.e., *it is the intersection of a finite number of closed half spaces. Then the logarithmic barrier of $\mathcal{X}$ is an SCB which satisfies Assumption 3.*

3. Main Results

Theorem 6 (Approximately quadratic losses) *Suppose f_t is ϵ-close to a convex, quadratic function $q_t(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A_t \mathbf{x} + \langle \mathbf{b}_t, \mathbf{x}\rangle + c_t$, for $\epsilon = O\left(dBT^{-1/2}\right)$. Let R be a ν-self-concordant barrier of $\mathcal{X}$ that satisfies Assumption 3. Suppose Algorithm 1 is run for T iterations with appropriate choice of hyper-parameters. Suppose the diameter of $\mathcal{X}$ is bounded by T, and the Lipschitz constants of $\{q_t\}_{t=1}^T$ are bounded by T. Then with probability at least $1-\delta$, the regret of the algorithm is upper bounded by $\tilde{O}\left(d^{11}(d+\nu)^5\sqrt{T}\right)$.*

Remarks. We now briefly discuss the above result. See Table 1 for a detailed comparison of our algorithm with other related algorithms.

- Our algorithm achieves the optimal regret guarantees in high probability, against adaptive adversaries. In comparison with Bubeck et al. (2017), our regret bound has similar dependence on T and slightly worse dependence on dimension d. We believe the dimension dependence of our regret can be improved to d^8 using a tighter analysis. Also note that the OMD based algorithm of Saha and Tewari (2011), which only relies on gradient estimates of loss functions, achieves a sub-optimal regret of $\tilde{O}\left(T^{2/3}\right)$.
- There are two key computational bottlenecks in our approach: (a) (**uniform sampling**) on an average, each iteration of our algorithm involves generating $\tilde{O}\left(\frac{d}{T}\right)$ samples from uniform distribution over a convex set. This is a well studied problem and several efficient

Paper	Regret	Adversary	amortized time complexity of each iteration (dependence on d, T)
Hazan and Li (2016)	$\tilde{O}\left(2^{d^4}(\log T)^{2d}T^{1/2}\right)$ (h.p)	adaptive	$O\left((\log T)^{\text{poly}(d)}\right)$
Bubeck et al. (2017)	$\tilde{O}\left(d^{9.5}T^{1/2}\right)$ (h.p)	adaptive	$O\left(2^d\right)$
Bubeck et al. (2017) (computationally efficient variant)	$\tilde{O}\left(d^{10.5}T^{1/2}\right)$ (h.p)	adaptive	$\tilde{O}\left(\text{poly}(dm)T\right)$
Saha and Tewari (2011)	$\tilde{O}\left(d^{2/3}T^{2/3}\right)$ (exp)	oblivious	involves minimization of a self concordant function
Flaxman et al. (2004)	$\tilde{O}\left(d^{1/2}T^{3/4}\right)$ (exp)	oblivious	involves projecting a point onto a convex set
This paper (instantiation for polytopes)	$\tilde{O}\left(d^{16}T^{1/2}\right)$ (h.p)	adaptive	$\tilde{O}\left(\frac{m^2d^3+(m+d)d^5}{T}\right)$ $+\tilde{O}\left((m+d)d^3+m^2d\right)$

Table 1: Comparison of various approaches for BCO with quadratic losses. "h.p", "exp" in the second column denote high probability and expected regret bounds respectively. m in the last column denotes the number of constraints in the polytope.

algorithms are known for uniform sampling from various classes of convex sets. To derive concrete runtime bounds, we consider the special case of the action set $\mathcal{X}$ being a polytope with m constraints. By relying on the algorithm of Laddha et al. (2020), we can generate a single sample in $\tilde{O}\left(m^2d^2+(m+d)d^4\right)$ time. (b) (**Newton update**) The Newton update in our algorithm involves minimization of a convex objective. This objective can be minimized using plethora of convex optimization techniques that have been developed. For the special case of action set being a polytope with m constraints, this objective can be minimized in $\tilde{O}\left(m^2d+(m+d)d^3\right)$ time using interior point methods (IPM).

- Our algorithm is robust to model mis-specification. In particular, even if each loss function f_t is $O\left(T^{-1/2}\right)$ away from a convex, quadratic function, our algorithm achieves the optimal regret. This result can be improved in a straightforward fashion: suppose each f_t is ϵ_t close to a convex, quadratic function. Then our algorithm achieves the optimal regret as long as $\sum_{t=1}^T \epsilon_t = O\left(T^{1/2}\right)$.

4. Related Work

In this section, we present a review of bandit optimization that is necessarily incomplete but is relevant to the current work. Multi-armed bandits is perhaps the simplest and most well studied sub-case of bandit optimization. Several efficient and optimal algorithms have been proposed for this problem (Audibert et al., 2009; Audibert and Bubeck, 2010; Abernethy et al., 2015; Lee et al., 2020). These algorithms first estimate the unknown loss function from one-point feedback, and then rely on Follow-the-Regularized-Leader (FTRL) framework with appropriate regularizer to choose the next action.

Moving beyond MAB, several recent works on bandit optimization have focused on BCO. For bounded, convex functions, Flaxman et al. (2004); Kleinberg (2005) developed online gradient descent style algorithms which achieve $\tilde{O}\left(T^{5/6}\right)$ regret. Recent works of Bubeck et al. (2017); Hazan and Li (2016) improved upon this result and developed algorithms which achieve the optimal $\tilde{O}\left(T^{1/2}\right)$ regret (also see Lattimore (2020) for information-theoretic upper bounds). However, these algorithms are computationally expensive. Moreover, the regret bounds of Hazan and Li (2016) have exponential dependence on dimension. As previously mentioned, several works have studied sub-cases of BCO. The most popular among these sub-cases is bandit linear optimization. For this problem, Abernethy et al. (2009) provided the first efficient algorithm with optimal $O\left(T^{1/2}\right)$ regret in expectation (see Dani et al. (2007) for lower bounds on regret for linear losses). This algorithm uses one-point estimate of the gradient and relies on OMD with SCB of $\mathcal{X}$ as the regularizer to choose the next action. Subsequent works have attempted to develop algorithms which achieve optimal regret in high-probability (Bartlett et al., 2008; Abernethy and Rakhlin, 2009). However, this turned out to be a difficult problem. It is only recently that an efficient and optimal algorithm for this problem was designed (Lee et al., 2020). A related line of work studied generalizations of linear bandits in euclidean space to the framework of Reproducing Kernel Hilbert Spaces (RKHS) (Chatterji et al., 2019; Takemori and Sato, 2020). As an application of this general framework, Chatterji et al. (2019) study convex quadratic losses. However, their algorithm, which is based on exponential weights update scheme, is computationally inefficient as it involves sampling from non log-concave distributions, which is NP-hard in general. Moving beyond linear losses, Flaxman et al. (2004) provided an algorithm with $O\left(T^{3/4}\right)$ regret for convex, Lipschitz loss functions. Saha and Tewari (2011) provided an algorithm for convex, smooth loss functions with $\tilde{O}\left(T^{2/3}\right)$ regret. For strongly convex, smooth functions, Hazan and Levy (2014); Ito (2020) provide algorithms which achieve the optimal $\tilde{O}\left(T^{1/2}\right)$ regret (see Shamir (2013) for lower bounds on regret for strongly convex losses).

Another active line of research on bandit optimization has focused on handling weaker adversary models. One such popular model is the stochastic adversary model, where it is assumed that the loss functions seen by the learner are independent samples generated from an unknown but fixed distribution (Lai and Robbins, 1985; Agrawal and Goyal, 2012; Filippi et al., 2010; Kveton et al., 2020; Agarwal et al., 2011; Srinivas et al., 2009). Recently, there has been a flurry of research on designing computationally efficient and optimal regret algorithms for this setting. However, these algorithms usually have poor performance in the stronger adversary model considered in this work. Yet another line of research on bandit optimization has focused on multi-point feedback models where the player can query each loss function at multiple points. Several recent works have designed efficient algorithms for this setting (Agarwal et al., 2010; Duchi et al., 2015; Shamir, 2017). These works show that it is possible to achieve similar regret guarantees in this setting as in the full-information setting.

5. Regularized Bandit Newton Algorithm

In this section we describe our algorithm for BCO (see Algorithm 1). At a high level, our algorithm tries to estimate the missing information (*i.e.,* gradient and Hessian) about the unknown loss function and pass it to the OMD framework, which chooses the next action.

Gradient and Hessian estimation. To estimate the gradient and Hessian of f_t at $\mathbf{x}_t$, we rely on the following randomized sampling scheme. We first randomly sample a point from the uniform distirbution on a ellipsoid with mean $\mathbf{x}_t$ and whose covariance matrix depends on the Hessian estimates of the past loss functions $\{f_s\}_{s=1}^{t-1}$. Next, we get one-point feedback from the adversary about the loss value at the sampled point, and use it to estimate the gradient and Hessian (see lines 6-13 of Algorithm 1). Our choice of the covariance matrix ensures that the sampling scheme adapts to the geometry of the cumulative loss $\sum_{s=1}^{t-1} f_s(\mathbf{x})$. In particular, it reduces exploration along directions which are strongly convex, and increases exploration along directions which are linear. This choice of exploration helps us achieve the right balance between exploration and exploitation, and plays a crucial role in achieving optimal regret guarantees.

Focus Region. Once we have an estimate of the gradient and Hessian, we construct a quadratic approximation of f_t around $\mathbf{x}_t$ (see line 14 of Algorithm 1). One caveat with this approximation, however, is that it is *not* guaranteed to have a low variance. To see this, first note that the variance of our estimate $\hat{f}_t(\mathbf{x})$ scales with $\|\mathbf{x} - \mathbf{x}_t\|_{M_t}$ (look at line 6 for definition of M_t). If $\mathbf{x}_t$ gets too close to the boundary of $\mathcal{X}$, then $\|\nabla^2 R(\mathbf{x}_t)\|_2$ and $\|M_t\|_2$ become very large. This in turn increases the variance of $\hat{f}_t(\mathbf{x})$, for $\mathbf{x}$ far away from $\mathbf{x}_t$. Consequently, we can not directly plug in the estimate $\hat{f}_t(\mathbf{x})$ into the OMD framework to choose the next action. To handle this issue, we rely on focus regions. In each iteration of the algorithm, we maintain a focus region F_t which satisfies the following key property: *the variance of the quadratic approximation within F_t is small and bounded.* To this end, we choose an F_t such that $\|\mathbf{x} - \mathbf{x}_t\|_{M_t}$ is bounded for any $\mathbf{x} \in F_t$. When picking the next action $\mathbf{x}_{t+1}$ using OMD, we restrict ourselves to the focus region F_t.

At the beginning of the algorithm, we set F_1 to $\mathcal{X}_\xi$, a scaled version of $\mathcal{X}$, which is defined as $\mathcal{X}_\xi = \xi \mathbf{x}_1 + (1-\xi)\mathcal{X}$, where $\xi = T^{-4}$ and $\mathbf{x}_1$ is the minimizer of $R(\mathbf{x})$ over $\mathcal{X}$. We use $\mathcal{X}_\xi$ instead of $\mathcal{X}$ purely for theoretical reasons, as it simplifies our proofs. In practice, one can set F_1 to $\mathcal{X}$. To ensure F_t satisfies the above mentioned property on low variance, we perform a check in each iteration of the algorithm (see lines 20-25). Intuitively, this checks if the current focus region has a large overlap with $B_{\alpha, M_t}(\mathbf{x}_t)$, the region of low variance of the quadratic approximation. If yes, we do not change the focus region. If not, we shrink the focus region so that it overlaps with the low variance region. Moreover, we simultaneously increase the learning rate (η_t) of OMD. This learning rate change ensures that the algorithm can quickly adapt to any changes of the adversary. If the adversary attempts to move the minimizer of $\min_{\mathbf{x} \in \mathcal{X}_\xi} \sum_{s=0}^{t} f_s(\mathbf{x})$ outside of the focus region, then increasing the learning rate helps us quickly detect this change. This plays a crucial role in the restart condition, which we explain next. Several recent works have used the idea of increasing learning schedule for various purposes (Agarwal et al., 2017; Bubeck et al., 2017; Lee et al., 2020).

Restart Condition. By relying on focus regions, we can only guarantee low regret w.r.t points within the focus region. To ensure low regret even w.r.t points outside the focus region, we perform another test every iteration, which we call "restart condition" (see lines 15-18). Intuitively, this test checks if the minimizer of $\min_{\mathbf{x} \in \mathcal{X}} \sum_{s=1}^{t} f_s(\mathbf{x})$ is well within the focus region. If yes, we continue the algorithm, as having a low regret w.r.t points in the focus region ensures the overall regret is low. If instead the test fails, then it usually implies that the minimizer is too close to the boundary of the focus region $\partial F_t \cap \mathrm{int}(\mathcal{X})$. In this case we show that the regret of our actions until now is negative. So, we can safely restart the algorithm. That is, we act as if time step $t+1$ is time step 1 and run the algorithm for $T - t$ steps.

Algorithm 1 Regularized Bandit Newton Algorithm

Input: ν-self-concordant barrier R, initial learning rate η_1, number of iterations T, radius of initial focus region α, learning rate increment γ, exploration parameter λ, β.
Denote $\hat{g}_0 = 0, \hat{H}_0 = 0, \eta_0 = 0, \xi = T^{-4}$
Let $\mathbf{x}_1 = \mathrm{argmin}_{\mathbf{x}\in\mathcal{X}} R(\mathbf{x})$
Focus Region $F_1 = \mathcal{X}_\xi$, where $\mathcal{X}_\xi = \xi\mathbf{x}_1 + (1-\xi)\mathcal{X}$
for $t = 1 \dots T$ **do**
    Let $M_t = \left(\nabla^2 R(\mathbf{x}_t) + \sum_{s=0}^{t-1}\eta_s\hat{H}_s\right)$.
    Sample $\mathbf{v}_{1,t}, \mathbf{v}_{2,t} \sim \mathbb{S}^{d-1}$, and compute $\mathbf{y}_t = \mathbf{x}_t + \lambda M_t^{-1/2}(\mathbf{v}_{1,t} + \mathbf{v}_{2,t})$
    if $\mathbf{y}_t \in \mathcal{X}$ **then**
        Play $\mathbf{y}_t$ and observe $f_t(\mathbf{y}_t)$
        Estimate gradient and Hessian of f_t at $\mathbf{x}_t$ as

$$\hat{g}_t = \lambda^{-1} df_t(\mathbf{y}_t) M_t^{1/2}\mathbf{v}_{1,t}, \quad \hat{H}_t = \frac{\lambda^{-2}}{2} d^2 f_t(\mathbf{y}_t) M_t^{1/2}\left(\mathbf{v}_{1,t}\mathbf{v}_{2,t}^T + \mathbf{v}_{2,t}\mathbf{v}_{1,t}^T\right) M_t^{1/2}$$

    else
        Play $\mathbf{x}_t$ and set $\hat{g}_t = 0, \hat{H}_t = 0$.
    end if
    Let $\hat{f}_t(\mathbf{x}) = \left\langle \hat{g}_t - \hat{H}_t\mathbf{x}_t, \mathbf{x}\right\rangle + \frac{1}{2}\mathbf{x}^T\hat{H}_t\mathbf{x}$ be the quadratic approximation of f_t at $\mathbf{x}_t$
    `//restart condition`
    if $\sum_{s=0}^{t}\hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_t}\sum_{s=0}^{t}\hat{f}_s(\mathbf{y}) \leq -\frac{\beta}{\eta_1}$ **then**
        Restart
    end if
    Compute $\mathbf{x}_{t+1}$ using OMD

$$\mathbf{x}_{t+1} = \underset{\mathbf{x}\in F_t}{\mathrm{argmin}}\, \eta_t \left\langle \hat{g}_t, \mathbf{x}\right\rangle + \Phi_{R_{t+1}}(\mathbf{x}, \mathbf{x}_t).$$

    Here $\Phi_{R_{t+1}}$ is Bregman divergence w.r.t $R_{t+1}(\mathbf{x}) \overset{\text{def}}{=} R(\mathbf{x}) + \sum_{s=0}^{t}\frac{\eta_s}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s)$
    `//Update focus region`
    if $\mathrm{Vol}(F_t \cap B_{\alpha, M_{t+1}}(\mathbf{x}_{t+1})) \leq \frac{1}{2}\mathrm{Vol}(F_t)$ **then**
        $F_{t+1} = F_t \cap B_{\alpha, M_{t+1}}(\mathbf{x}_{t+1})$ and $\eta_{t+1} = (1+\gamma)\eta_t$
    else
        $F_{t+1} = F_t$ and $\eta_{t+1} = \eta_t$
    end if
end for

We note that the ideas of focus region and restart condition appeared in the work of Bubeck et al. (2017). However, their approach is computationally expensive, even after restricting the loss functions to convex quadratics. There are two main reasons for this:

1. the algorithm of Bubeck et al. (2017) relies on exponential weights update scheme. Each iteration of this algorithm involves generating $\tilde{O}(d)$ samples from an approximately log-concave distribution, which can be computationally expensive in high dimensions. In

contrast, we rely on OMD framework in our work, which doesn't require access to an approximately log-concave sampler.

2. the restart condition of Bubeck et al. (2017) involves optimization of an approximately convex objective over a *non-convex* set. To be precise, the authors use the following restart condition

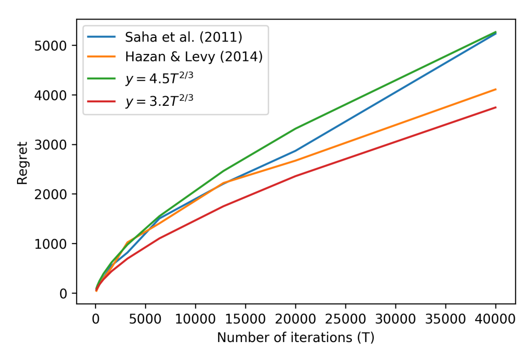
$$\min_{\mathbf{y}\in\partial F_t\cap\mathrm{int}(\mathcal{X})}\sum_{s=0}^{t}\hat{f}_s(\mathbf{y})-\min_{\mathbf{y}\in F_t}\sum_{s=0}^{t}\hat{f}_s(\mathbf{y})\leq\frac{1}{\eta_1}.$$

Implementing this can be NP-hard in general because the domain of the first optimization is a nonconvex set. While the authors present a modified algorithm to handle this issue, it is still computationally expensive (the runtime of each iteration is $\tilde{O}\left(d^aT\right)$ for some large a). Moreover, the modified algorithm only works for constraint sets which are polytopes and whose coefficients in the constraints are rational numbers with absolute values of numerators and denominators bounded by $\mathrm{poly}(T)$. In contrast, the restart condition we use only involves minimization of $\min_{\mathbf{x}\in F_t}\sum_{s=0}^{t}\hat{f}_s(\mathbf{x})$, which we show is approximately convex and can be optimized efficiently (see Section 7).

5.1. Importance of Hessian Estimates

In this section we empirically demonstrate that existing OMD algorithms that only rely on gradient information don't achieve optimal regret bounds for quadratic losses (Abernethy et al., 2009; Saha and Tewari, 2011; Hazan and Levy, 2014).

Lets consider a simple example where the adversary always selects the following loss function in each iteration: $f_t(\mathbf{x})=\sum_{i=1}^{d/2}x_i^2+\sum_{i=1}^{d}x_i$. Here, we choose the domain $\mathcal{X}$ to be $\mathbb{B}^d$. In this case, all the three algorithms mentioned above get sub-optimal regret of $\Omega(T^{2/3})$ (see image for empirical evidence). This is because M_t (defined in line 6 of Algorithm 1), which controls the exploration, is not chosen appropriately by these algorithms. Ideally, we should explore the first $d/2$ directions less and the last $d/2$ directions more. This is because the expected regret of these algorithms depends on the following term: $\mathbb{E}[f_t(\mathbf{y}_t)-f_t(\mathbf{x}_t)]$ $=\mathbb{E}[f_t(\mathbf{x}_t+M_t^{-1/2}\mathbf{v}_t)-f_t(\mathbf{x}_t)]=\sum_{i=1}^{d/2}\mathbb{E}[(M_t^{-1/2}\mathbf{v}_t)_i^2]$. So a good choice of M_t should ensure $\mathbb{E}[(M_t^{-1/2}\mathbf{v}_t)_i^2]$ is low for the first $d/2$ coordinates. We achieve this in our algorithm by relying on Hessian estimates, which tell us how much exploration to do in each direction. For the example considered here, M_t in our algorithm is approximately equal to $\nabla^2R(\mathbf{x}_t)+\sum_{s=0}^{t-1}2\eta_s\nabla^2f_s(\mathbf{x})$. For this choice of M_t, $\mathbb{E}[(M_t^{-1/2}\mathbf{v}_t)_i^2]$ goes down with t along the first $d/2$ directions. As a result, our algorithm performs less exploration along directions with large curvature, and more exploration along directions with small curvature, and achieves the optimal trade-off between exploration and exploitation. If we do uniform exploration in all directions (similar to existing algorithms), then we don't achieve the optimal regret.



6. Analysis

In this section we provide an outline of the proof of our main result stated in Theorem 6. We prove the following Theorem from which Theorem 6 follows readily.

Theorem 7 (Regret) *Consider the setting of Theorem 6. Suppose Algorithm 1 is run for T iterations with the following hyper-parameters*

$$\lambda = \frac{1}{4},\ \alpha = c_1(\nu + d)d\log^2 dT,\ \beta = d\log dT,\ \gamma = \frac{c_2}{d\log T},\ \eta_1 = \frac{c_3}{d^7(B+\epsilon)\alpha^4\sqrt{T}\log T},$$

for some universal constants $c_1, c_2, c_3 > 0$. Let $\mathcal{T}$ be the minimum between T and the first time at which the algorithm restarts. Then with probability at least $1-\delta$

$$\sum_{t=1}^{\mathcal{T}} f_t(\mathbf{y}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{\mathcal{T}} f_t(\mathbf{x}) \leq \begin{cases} \tilde{O}\left(d^{11}(d+\nu)^5\sqrt{T}\right) & \text{if } \mathcal{T} = T \\ 0 & \text{otherwise} \end{cases}.$$

Proof (Sketch) We first consider the case where the restart condition triggered for the first time at iteration $\mathcal{T} < T$. Then we show that the regret of the learner until $\mathcal{T}$ is negative. There are several key steps involved in showing this result:

1. We first show that the minimizer of the cumulative loss $\sum_{s=0}^{\mathcal{T}} f_s(\mathbf{x})$ over the entire domain $\mathcal{X}$ lies in $F_{\mathcal{T}}$; that is, $\min_{\mathbf{x}\in\mathcal{X}} \sum_{s=0}^{\mathcal{T}} f_s(\mathbf{x}) = \min_{\mathbf{x}\in F_{\mathcal{T}}} \sum_{s=0}^{\mathcal{T}} f_s(\mathbf{x})$. This immediately entails that the regret after $\mathcal{T}$ iterations satisfies.

$$\mathrm{Reg}_{\mathcal{T}} = \sum_{s=0}^{\mathcal{T}} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in F_{\mathcal{T}}} \sum_{s=0}^{\mathcal{T}} f_s(\mathbf{x}).$$

2. Next, consider the following for any $\mathbf{x} \in F_{\mathcal{T}}$

$$\sum_{s=0}^{\mathcal{T}} f_s(\mathbf{y}_s) - \sum_{s=0}^{\mathcal{T}} f_s(\mathbf{x}) = \sum_{s=0}^{\mathcal{T}} [f_s(\mathbf{y}_s) - f_s(\mathbf{x}_s)] + \sum_{s=0}^{\mathcal{T}} \left[f_s(\mathbf{x}_s) - f_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) + \hat{f}_s(\mathbf{x})\right]$$
$$+ \sum_{s=0}^{\mathcal{T}} \left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right].$$

Recall $\mathbf{y}_t - \mathbf{x}_t = \lambda M_t^{-1/2}(\mathbf{v}_{1,t} + \mathbf{v}_{2,t})$. Relying on standard martingale concentration inequalities, the first term in the RHS above can be bounded as $\tilde{O}\left(d\eta_1^{-1}\right)$. To bound the second term, we rely on a key property of our loss estimates $\{\hat{f}_t\}_{t=1}^T$: the cumulative loss estimate concentrates well around the true cumulative loss (see Proposition 8). Using this property, the second term can be bounded as $O\left(\eta_1^{-1}\right)$. To bound the last term, we rely on the definition of restart condition which says that $\sum_{s=0}^{\mathcal{T}} \hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x}) \leq -\beta\eta_1^{-1}$. Combining these bounds shows that the regret after $\mathcal{T}$ iterations is negative.

Next, consider the case where the restart condition never triggered. Here, we can again show that the minimizer of the cumulative loss over the entire domain lies in the focus region F_T. So it suffices to bound $\sum_{s=0}^{T} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in F_T} \sum_{s=0}^{T} f_s(\mathbf{x})$. Consider the same decomposition of regret as above. We use the same arguments as above to bound the first two terms in the decomposition. To bound the thrid term, we consider the following

$$\sum_{s=0}^{T}\left[\hat{f}_s(\mathbf{x}_s)-\hat{f}_s(\mathbf{x})\right]=\sum_{s=0}^{T}\left[\hat{f}_s(\mathbf{x}_s)-\hat{f}_s(\mathbf{x}_{s+1})\right]+\sum_{s=0}^{T}\left[\hat{f}_s(\mathbf{x}_{s+1})-\hat{f}_s(\mathbf{x})\right].$$

The first term in the RHS can be upper bounded using stability of the iterates $\|\mathbf{x}_t-\mathbf{x}_{t+1}\|_{M_t}$ (in our proof we show that $\|\mathbf{x}_t-\mathbf{x}_{t+1}\|_{M_t}$ is upper bounded by $\tilde{O}\left(\eta_t\right)$). The second term is the regret of Be-The-Regularized-Leader and can be upper bounded as $\tilde{O}\left(\eta_1^{-1}\right)$. Combining these two bounds, we show that the regret is $\tilde{O}\left(T^{1/2}\right)$. ■

The proof of Theorem 7 relies on several crucial properties of the iterates produced by our algorithm. First, we need to ensure that the matrix M_t is positive definite and the iterates $\mathbf{y}_t$ produced by our algorithm lie within the domain $\mathcal{X}$. Second, we need to show that the algorithm is stable, *i.e.,* $\|\mathbf{x}_t-\mathbf{x}_{t+1}\|_{M_t}$ is small. The following proposition plays a crucial role in showing these properties. It is concerned about concentration of the Hessian estimates $\{\hat{H}_t\}_{t=1}^T$, and the loss estimates $\{\hat{f}_t\}_{t=1}^T$ computed by the Algorithm.

Proposition 8 *Consider the setting of Theorem 7. Let $\mathcal{T}$ be the minimum between T and the first time at which the algorithm restarts. Then for any $t\leq\mathcal{T}$, the following properties hold with probability at least $1-T^{-2}$*

- *Let $H_t=\frac{1}{2}(A_t+A_t^T)$ be the Hessian of $q_t(\mathbf{x})$, and let $\tilde{M}_t=\nabla^2R(\mathbf{x}_t)+\sum_{s=0}^{t-1}\eta_sH_s$. Then M_t defined in line 6 of Algorithm 1 satisfies*

$$\|\tilde{M}_t^{-1/2}(\tilde{M}_t-M_t)\tilde{M}_t^{-1/2}\|_2=O\left(\alpha^2\eta_1\lambda^{-2}d^5B\sqrt{T\log(dT)}\right).$$

- *The cumulative loss estimate $\sum_{s=1}^{t}\hat{f}_s(\mathbf{x})$ satisfies*

$$\sup_{\mathbf{x}\in F_t}\left|\sum_{s=1}^{t}\eta_1(\hat{f}_s(\mathbf{x})-\hat{f}_s(\mathbf{x}_s)-q_s(\mathbf{x})+q_s(\mathbf{x}_s))\right|\leq O\left(\alpha^2\eta_1\lambda^{-2}Bd^{4.5}\sqrt{T\log dT}\right).$$

7. Implementation

In this section, we discuss the implementation aspects of our algorithm.

Focus region update. To estimate the ratio $\frac{\mathrm{Vol}(F_t\cap B_{\alpha,M_{t+1}}(\mathbf{x}_{t+1}))}{\mathrm{Vol}(F_t)}$, we generate sufficiently many independent uniformly distributed samples in F_t and count what fraction of them fall in $F_t\cap B_{\alpha,M_{t+1}}(\mathbf{x}_{t+1})$. By sampling just $O(\log T)$ points, we can show that with probability at least $1-\frac{1}{T^4}$, the focus region gets updated whenever the true ratio is less than $\frac{1}{4}$ and doesn't get updated whenever the true ratio is greater than $\frac{3}{4}$. The intermediate values don't effect our argument. Next, note that we need not generate the samples every iteration. It suffices to generate them only when the focus region gets updated. We can reuse the old samples in rest of the iterations. In Appendix E we show that the focus region doesn't get updated more than $O\left(d\log T\right)$ times (see Lemma 22). So, over T iterations of the Algorithm, we only need to generate $O\left(d\log^2 T\right)$ samples.

As previously mentioned, uniform sampling from a convex set is a well studied problem. For the special case of the action set being a polytope with m constraints, we rely on the recent work of Laddha et al. (2020) which uses Dikin walk for sampling. The authors show that the Dikin walk mixes in $O(d\bar{\nu})$ steps, where $\bar{\nu}$ is the strong self concordant parameter of the set. For our problem, $\bar{\nu}$ is $O(m + O(d \log T))$ (this follows from the fact that each of our focus regions is an intersection of $O(d \log T)$ elliposoids and a polytope). Moreover, each iteration of Dikin walk takes $O\left(d^3 \log T + md + d^2\right)$ time. So generating a single sample from uniform distribution in F_t takes $\tilde{O}\left(m^2 d^2 + (m+d)d^4\right)$ time.

OMD Update. Our results in Appendix E entail that the objective in line 19 is strictly convex (see Lemma 21). So we can use IPM to solve the objective. As a concrete example, lets again consider the case of action set being a polytope with m constraints. Since there are $O(d \log T)$ elliposoidal constrains and m linear constraints, the self concordant parameter of the entire objective is $m + O(d \log T)$. So, the number of Newton updates we perform is $\tilde{O}(m+d)$. Moreoever, performing each newton update takes $O\left(d^3 \log T + md\right)$ time. So, the overall compute time of IPM is $\tilde{O}\left(m^2 d + (m+d)d^3\right)$.

Restart Condition. Checking the restart condition involves minimizing $\sum_{s=0}^{t} \hat{f}_s(\mathbf{y})$ over the focus region F_t. We note that this need not be a convex function. However, it is pointwise close to the following convex function: $\sum_{s=0}^{t} \hat{f}_s(\mathbf{y}) + (d^2\alpha^2\eta_1)^{-1}(\mathbf{y}-\mathbf{x}_t)^T \nabla^2 R(\mathbf{x}_t)(\mathbf{y}-\mathbf{x}_t)$ (see Remark 24 in Appendix for a discussion on the convexity of this objective). To see why this objective is pointwise close to $\sum_{s=0}^{t} \hat{f}_s(\mathbf{y})$, first note that our choice of F_t always ensures $\|\mathbf{y}-\mathbf{x}_t\|_{\nabla^2 R(\mathbf{x}_t)} \leq O(d\alpha)$ for any $\mathbf{y} \in F_t$ (see Lemma 21). So the modified objective is $O\left(\eta_1^{-1}\right)$ close to the original objective. So we can rely on IPM to solve the modified objective and obtain $O\left(\eta_1^{-1}\right)$-approximate solution to the original objective (note that an approximate solution suffices for our argument). The computational complexity of IPM in this case is same as the complexity of OMD update described above.

8. Conclusion

In this paper, we presented a new algorithm for bandit optimization with convex (approximately) quadratic functions. Our algorithm achieves the optimal regret rate of $\tilde{O}(\sqrt{T})$ and is computationally much more efficient than any other known algorithms for this problem. To obtain these results, we (i) estimate the Hessian of the loss functions and use it in a controlled fashion to minimize the effect of variance in this estimation and (ii) develop new algorithmic ideas to implement this efficiently.

Future work. While our work focuses on the convex quadratic setting, we believe our ideas can be extended to other convex, parameteric loss functions such as generalized linear models. However, extending the idea of using Hessian (or more generally k^{th} order derivatives for $k > 1$) estimates to obtain efficient algorithms with optimal regret rates seems challenging, even for highly smooth functions as the estimates of higher order derivatives come with high variance and new ideas seem necessary to make effective use of them. This is an interesting future direction to explore. Finally, we believe the dimension dependence in our regret bound can improved to d^8 by tightening the Hessian concentration result in Proposition 8. We base this claim on the results in Appendix D, where we show that Algorithm 1 achieves $\tilde{O}\left(d^{5.5}\sqrt{T}\right)$ regret when the Hessian of f_t is known to the learner ahead of round t.

Acknowledgment

A.S., and P.R. acknowledge the support of NSF via IIS-1909816.

References

Jacob Abernethy and Alexander Rakhlin. Beating the adaptive bandit with high probability. In *2009 Information Theory and Applications Workshop*, pages 280–289. IEEE, 2009.

Jacob D Abernethy, Elad Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. 2009.

Jacob D Abernethy, Chansoo Lee, and Ambuj Tewari. Fighting bandits with a new kind of smoothness. *Advances in Neural Information Processing Systems*, 28:2197–2205, 2015.

Alekh Agarwal, Ofer Dekel, and Lin Xiao. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *COLT*, pages 28–40. Citeseer, 2010.

Alekh Agarwal, Dean P Foster, Daniel J Hsu, Sham M Kakade, and Alexander Rakhlin. Stochastic convex optimization with bandit feedback. In *Advances in Neural Information Processing Systems*, pages 1035–1043, 2011.

Alekh Agarwal, Haipeng Luo, Behnam Neyshabur, and Robert E Schapire. Corralling a band of bandit algorithms. In *Conference on Learning Theory*, pages 12–38. PMLR, 2017.

Shipra Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *Conference on learning theory*, pages 39–1, 2012.

Jean-Yves Audibert and Sébastien Bubeck. Regret bounds and minimax policies under partial monitoring. *The Journal of Machine Learning Research*, 11:2785–2836, 2010.

Jean-Yves Audibert, Sébastien Bubeck, et al. Minimax policies for adversarial and stochastic bandits. In *COLT*, volume 7, pages 1–122, 2009.

Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

Peter Bartlett, Varsha Dani, Thomas Hayes, Sham Kakade, Alexander Rakhlin, and Ambuj Tewari. High-probability regret bounds for bandit online linear optimization. In *Proceedings of the 21st Annual Conference on Learning Theory-COLT 2008*, pages 335–342. Omnipress, 2008.

Alexandre Belloni, Tengyuan Liang, Hariharan Narayanan, and Alexander Rakhlin. Escaping the local minima via simulated annealing: Optimization of approximately convex functions. In *Conference on Learning Theory*, pages 240–265, 2015.

Sébastien Bubeck and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *arXiv preprint arXiv:1204.5721*, 2012.

Sébastien Bubeck, Yin Tat Lee, and Ronen Eldan. Kernel-based methods for bandit convex optimization. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 72–85, 2017.

Niladri Chatterji, Aldo Pacchiano, and Peter Bartlett. Online learning with kernel losses. In *International Conference on Machine Learning*, pages 971–980. PMLR, 2019.

Varsha Dani, Thomas P Hayes, and Sham M Kakade. The price of bandit information for online optimization. In *Proceedings of the 20th International Conference on Neural Information Processing Systems*, pages 345–352, 2007.

John C Duchi, Michael I Jordan, Martin J Wainwright, and Andre Wibisono. Optimal rates for zero-order convex optimization: The power of two function evaluations. *IEEE Transactions on Information Theory*, 61(5):2788–2806, 2015.

Sarah Filippi, Olivier Cappe, Aurélien Garivier, and Csaba Szepesvári. Parametric bandits: The generalized linear case. In *Advances in Neural Information Processing Systems*, pages 586–594, 2010.

Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*, 2004.

Elad Hazan. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

Elad Hazan and Kfir Levy. Bandit convex optimization: Towards tight bounds. *Advances in Neural Information Processing Systems*, 27:784–792, 2014.

Elad Hazan and Yuanzhi Li. An optimal algorithm for bandit convex optimization. *arXiv preprint arXiv:1603.04350*, 2016.

Xiaowei Hu, LA Prashanth, András György, and Csaba Szepesvari. (bandit) convex optimization with biased noisy gradient oracles. In *Artificial Intelligence and Statistics*, pages 819–828. PMLR, 2016.

Shinji Ito. An optimal algorithm for bandit convex optimization with strongly-convex and smooth loss. In *International Conference on Artificial Intelligence and Statistics*, pages 2229–2239. PMLR, 2020.

Robert D Kleinberg. Nearly tight bounds for the continuum-armed bandit problem. In *Advances in Neural Information Processing Systems*, pages 697–704, 2005.

Branislav Kveton, Manzil Zaheer, Csaba Szepesvari, Lihong Li, Mohammad Ghavamzadeh, and Craig Boutilier. Randomized exploration in generalized linear bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 2066–2076, 2020.

Aditi Laddha, Yin Tat Lee, and Santosh Vempala. Strong self-concordance and sampling. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1212–1222, 2020.

Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

Tor Lattimore. Improved regret for zeroth-order adversarial bandit convex optimisation. *arXiv preprint arXiv:2006.00475*, 2020.

Chung-Wei Lee, Haipeng Luo, Chen-Yu Wei, and Mengxiao Zhang. Bias no more: high-probability data-dependent regret bounds for adversarial bandits and mdps. *arXiv preprint arXiv:2006.08040*, 2020.

László Lovász and Santosh Vempala. Where to start a geometric random walk, 2003.

Arkadi Nemirovski. Interior point polynomial time methods in convex programming. *Lecture notes*, 2004.

Yurii Nesterov. *Lectures on convex optimization*, volume 137. Springer, 2018.

Ankan Saha and Ambuj Tewari. Improved regret guarantees for online smooth convex optimization with bandit feedback. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 636–642, 2011.

Ohad Shamir. On the complexity of bandit and derivative-free stochastic convex optimization. In *Conference on Learning Theory*, pages 3–24. PMLR, 2013.

Ohad Shamir. An optimal algorithm for bandit and zero-order convex optimization with two-point feedback. *The Journal of Machine Learning Research*, 18(1):1703–1713, 2017.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.

Sho Takemori and Masahiro Sato. Approximation methods for kernelized bandits. *arXiv preprint arXiv:2010.12167*, 2020.

Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*, 12(4):389–434, 2012.

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.

Contents

A Proof of Proposition 1 **19**

B Proof of Proposition 4 **20**

C Proof of Proposition 5 **21**

D Warm up: Hypothetical case of known Hessians **21**
D.1 Intermediate Results 22
D.2 Proof of Theorem 9 23

E Proof of Theorem 7 **38**
E.1 Proof of Proposition 8 48
E.2 Main argument for Theorem 7 54

F Additional Results **58**

G Review of Self Concordant Barriers **59**

Appendix A. Proof of Proposition 1

Let $f(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, A\mathbf{x}\rangle + \langle \mathbf{b}, \mathbf{x}\rangle + c$, for some $A \in \mathbb{R}^{d\times d}, \mathbf{b} \in \mathbb{R}^d, c \in \mathbb{R}$. The gradient and Hessian of f at $\mathbf{x}$ are given by

$$\nabla f(\mathbf{x}) = \frac{1}{2}(A + A^T)\mathbf{x} + \mathbf{b}, \quad \nabla^2 f(\mathbf{x}) = \frac{1}{2}(A + A^T).$$

Gradient. From the definition of f, we have

$$\begin{aligned}\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1 f(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right] &= \frac{1}{2}\underbrace{\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)^T A(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]}_{T_1}\\ &+ \underbrace{\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1\langle \mathbf{b}, \mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2\rangle\right]}_{T_2}.\end{aligned}$$

First consider T_1

$$\begin{aligned}&\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)^T A(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]\\ &= \mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1\right]\mathbf{x}^T A\mathbf{x} + \mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1(\mathbf{v}_1+\mathbf{v}_2)^T CAC(\mathbf{v}_1+\mathbf{v}_2)\right]\\ &\quad+ \mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1\mathbf{x}^T AC(\mathbf{v}_1+\mathbf{v}_2) + C^{-1}\mathbf{v}_1(\mathbf{v}_1+\mathbf{v}_2)^T CA\mathbf{x}\right].\end{aligned}$$

Since $\mathbf{v}_1, \mathbf{v}_2$ are independent random variables whose distributions are symmetric around origin, it is easy to see that the first two terms in the RHS are 0. So we get

$$\begin{aligned}&\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)^T A(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]\\ &\quad= \mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1\mathbf{x}^T AC(\mathbf{v}_1+\mathbf{v}_2) + C^{-1}\mathbf{v}_1(\mathbf{v}_1+\mathbf{v}_2)^T CA\mathbf{x}\right]\\ &\quad= \mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1\mathbf{x}^T AC\mathbf{v}_1 + C^{-1}\mathbf{v}_1\mathbf{v}_1^T CA\mathbf{x}\right]\\ &\quad= C^{-1}\mathbb{E}_{\mathbf{v}_1\sim\mathbb{S}^{d-1}}\left[\mathbf{v}_1\mathbf{v}_1^T\right]C(A\mathbf{x}+A^T\mathbf{x}) = \frac{1}{d}(A+A^T)\mathbf{x},\end{aligned}$$

where we used the fact that $\mathbb{E}_{\mathbf{v}_1\sim\mathbb{S}^{d-1}}\left[\mathbf{v}_1\mathbf{v}_1^T\right] = \frac{1}{d}I_{d\times d}$. Now consider T_2

$$\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1\langle \mathbf{b}, \mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2\rangle\right] = \mathbb{E}_{\mathbf{v}_1}\left[C^{-1}\mathbf{v}_1\langle \mathbf{b}, C\mathbf{v}_1\rangle\right] = \frac{1}{d}\mathbf{b}.$$

Substituting the above expressions for T_1, T_2 in the first display gives us

$$\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}\mathbf{v}_1 f(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right] = \frac{1}{d}\nabla f(\mathbf{x}).$$

Hessian. From the definition of f, we have

$$\begin{aligned}&\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T+\mathbf{v}_2\mathbf{v}_1^T)C^{-1}f(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]\\ &\quad= \frac{1}{2}\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T+\mathbf{v}_2\mathbf{v}_1^T)C^{-1}(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)^T A(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]\\ &\quad\quad+ \mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T+\mathbf{v}_2\mathbf{v}_1^T)C^{-1}\langle \mathbf{b}, \mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2\rangle\right]\end{aligned}$$

Since $\mathbf{v}_1, \mathbf{v}_2$ are independent random variables whose distributions are symmetric around origin, it is easy to see that the second term in the RHS above is 0. So, consider the first term

$$\begin{aligned}
&\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T+\mathbf{v}_2\mathbf{v}_1^T)C^{-1}(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)^TA(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]\\
&\quad=\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T+\mathbf{v}_2\mathbf{v}_1^T)C^{-1}(\mathbf{v}_1+\mathbf{v}_2)^TCAC(\mathbf{v}_1+\mathbf{v}_2)\right]\\
&\quad=2\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_1^T)CAC(\mathbf{v}_2\mathbf{v}_2^T)C^{-1}+C^{-1}(\mathbf{v}_1\mathbf{v}_2^T)CAC(\mathbf{v}_1\mathbf{v}_2^T)C^{-1}\right],
\end{aligned}$$

where we relied on the fact that odd moments of $\mathbf{v}_1, \mathbf{v}_2$ are zero. Continuing, we get

$$\begin{aligned}
&\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T+\mathbf{v}_2\mathbf{v}_1^T)C^{-1}(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)^TA(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]\\
&\quad=\frac{2}{d^2}(A+A^T),
\end{aligned}$$

where we used the fact that $\mathbb{E}_{\mathbf{v}_1\sim\mathbb{S}^{d-1}}\left[\mathbf{v}_1\mathbf{v}_1^T\right]=\frac{1}{d}I$ and $\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[(\mathbf{v}_1\mathbf{v}_2^T)W(\mathbf{v}_1\mathbf{v}_2^T)\right]=\frac{1}{d^2}W^T$. Substituting this in the first display gives us

$$\mathbb{E}_{\mathbf{v}_1,\mathbf{v}_2\sim\mathbb{S}^{d-1}}\left[C^{-1}(\mathbf{v}_1\mathbf{v}_2^T+\mathbf{v}_2\mathbf{v}_1^T)C^{-1}f(\mathbf{x}+C\mathbf{v}_1+C\mathbf{v}_2)\right]=\frac{2}{d^2}\nabla^2 f(\mathbf{x}).$$

Appendix B. Proof of Proposition 4

This proposition was proved in Nemirovski (2004). For the sake of completeness, we reproduce the proof here. Let $\mathbf{h}=\mathbf{y}-\mathbf{x}$ and $r=\|\mathbf{h}\|_{\nabla^2R(\mathbf{x})}$. Let $\phi(t)=\nabla^2R(\mathbf{x}+t\mathbf{h})[\mathbf{h},\mathbf{h}]$. The function ϕ satisfies the following properties

$$0\le\phi(t),\ r^2=\phi(0),\ |\phi'(t)|=|\nabla^3R(\mathbf{x}+t\mathbf{h})[\mathbf{h},\mathbf{h},\mathbf{h}]|\le 2\phi^{3/2}(t).$$

So, for all positive ϵ, we have

$$0<\phi_\epsilon(t)=\epsilon+\phi(t),\ |\phi_\epsilon'(t)|\le 2\phi_\epsilon^{3/2}(t).$$

Continuing,

$$\left|\frac{d}{dt}\phi_\epsilon^{-1/2}(t)\right|\le 1.$$

It follows that

$$\phi_\epsilon^{-1/2}(t)\le\phi_\epsilon^{-1/2}(0)+t.$$

This gives us

$$\frac{\phi_\epsilon(0)}{(1+t\phi_\epsilon^{1/2}(0))^2}\le\phi_\epsilon(t).$$

The above inequality holds for any $t\in[0,1]$ and any $\epsilon>0$. Passing to limit as $\epsilon\to 0+$, we get

$$\frac{r^2}{(1+rt)^2}\le\phi(t)=\nabla^2R(\mathbf{x}+t\mathbf{h})[\mathbf{h},\mathbf{h}]$$

Setting $t=1$, we get $\nabla^2R(\mathbf{y})\ge\frac{1}{(1+r)^2}\nabla^2R(\mathbf{x})$. Using the fact that $r\le\lambda$ gives us the required result.

Appendix C. Proof of Proposition 5

Let $\mathcal{X} = \{\mathbf{x} : \langle \mathbf{a}_i, \mathbf{x} \rangle \geq b_i, \text{ for } i = 1, \dots m\}$. Consider the logarithmic barrier for $\mathcal{X}$

$$R(\mathbf{x}) = -\sum_i \log(\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i).$$

It is well know that $R(\mathbf{x})$ is a m-self concordant barrier for $\mathcal{X}$ (Nemirovski, 2004). The Hessian of R is given by

$$\nabla^2 R(\mathbf{x}) = \sum_i \frac{\mathbf{a}_i \mathbf{a}_i^T}{(\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i)^2}.$$

Since $\|\mathbf{y} - \mathbf{x}\|_{\nabla^2 R(\mathbf{x})} \leq \lambda$, we have

$$\begin{aligned}
&\sum_i \frac{\langle \mathbf{a}_i, \mathbf{y} - \mathbf{x} \rangle^2}{(\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i)^2} \leq \lambda^2 \\
&\implies \forall i, \frac{\langle \mathbf{a}_i, \mathbf{y} - \mathbf{x} \rangle^2}{(\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i)^2} \leq \lambda^2 \\
&\implies \forall i, \langle \mathbf{a}_i, \mathbf{y} - \mathbf{x} \rangle \leq \lambda(\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i),
\end{aligned}$$

where we used the fact that $\mathbf{x} \in \mathcal{X}$ and hence $\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i \geq 0$ in the last step. This then implies that

$$\langle \mathbf{a}_i, \mathbf{y} \rangle - b_i \leq (1 + \lambda)(\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i).$$

Since $\mathbf{y} \in \mathcal{X}$ and hence $\langle \mathbf{a}_i, \mathbf{y} \rangle - b_i \geq 0$, we have $(\langle \mathbf{a}_i, \mathbf{y} \rangle - b_i)^2 \leq (1 + \lambda)^2 (\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i)^2$. So, we have

$$\nabla^2 R(\mathbf{y}) = \sum_i \frac{\mathbf{a}_i \mathbf{a}_i^T}{(\langle \mathbf{a}_i, \mathbf{y} \rangle - b_i)^2} \succeq \frac{1}{(1+\lambda)^2} \sum_i \frac{\mathbf{a}_i \mathbf{a}_i^T}{(\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i)^2} = \frac{1}{(1+\lambda)^2} \nabla^2 R(\mathbf{x}).$$

This finishes the proof of the Proposition.

Appendix D. Warm up: Hypothetical case of known Hessians

In this section, we consider a hypothetical scenario where we are given access to the Hessian H_t of loss function f_t at the beginning of iteration t. In such a scenario, instead of estimating the Hessian from single point feedback (as done in Algorithm 1), one can rely on H_t. In this section, we study such an algorithm; that is, we study a variant of Algorithm 1 where we replace the Hessian estimate $\hat{H}_t$ with H_t.

Studying this hypothetical scenario helps the readers understand the intuition behind Algorithm 1. Moreover, it greatly simplifies our proofs and makes it easier to understand the key ideas in the proof of Theorem 6. Finally, this hypothetical scenario encompasses the important special case of linear loss functions (*i.e.,* $H_t = 0$) that is often studied in the literature of bandit optimization (Abernethy et al., 2009).

The following Theorem bounds the regret of this hypothetical algorithm. To further simplify the analysis, we assume the loss functions are exactly quadratic (*i.e.,* $\epsilon = 0$).

Theorem 9 (Approximately quadratic losses) *Suppose f_t is a convex, quadratic function $f_t(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A_t \mathbf{x} + \langle \mathbf{b}_t, \mathbf{x}\rangle + c_t$. Let R be a ν-self-concordant barrier of $\mathcal{X}$ that satisfies Assumption 3. Suppose the diameter of $\mathcal{X}$ is bounded by T, and the Lipschitz constants of $\{f_t\}_{t=1}^T$ are bounded by T. Suppose Algorithm 1 is run for T iterations with $\hat{H}_t = \frac{1}{2}(A_t + A_t^T)$ and the following hyper-parameters*

$$\lambda = \frac{1}{4},\ \alpha = c_1(\nu + d)d\log^2 dT,\ \beta = 4d\log dT,\ \gamma = \frac{c_2}{d\log T},\ \eta_1 = \frac{c_3}{d^{2.5}B\alpha\sqrt{T}\log T},$$

for some universal constants $c_1, c_2, c_3 > 0$. Let $\mathcal{T}$ be the minimum between T and the first time at which the algorithm restarts. Then with probability at least $1-\delta$

$$\sum_{t=1}^{\mathcal{T}} f_t(\mathbf{y}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{\mathcal{T}} f_t(\mathbf{x}) \leq \begin{cases} \tilde{O}\left(d^{3.5}(d+\nu)^2\sqrt{T}\right) & \text{if } \mathcal{T} = T \\ 0 \quad \text{otherwise} \end{cases}.$$

Remark 10 (Linear losses) *The above regret bound can be improved to $\tilde{O}\left(d^{3.5}\nu^2\sqrt{T}\right)$ for linear loss functions. This is because for linear losses, we can obtain a tighter bound for $\sum_{s=1}^{\mathcal{T}} f_s(\mathbf{y}_s) - f_s(\mathbf{x}_s)$ than the one we obtained for general quadratic functions in the proof of Theorem 9 (see Equation 6 below).*

Remark 11 (Convex losses) *The above Theorem can be generalized in a straightforward way to general convex loss functions. Suppose f_t's are general convex loss functions and suppose we have access to a lower bound for of $\nabla^2 f_t$'s. In particular, suppose at the beginning iteration t, we have access to H_t which satisfies: $\forall \mathbf{x} \in \mathcal{X}, H_t \preceq \nabla^2 f_t(\mathbf{x})$. Suppose we run Algorithm 1 with $\hat{H}_t = H_t$. Then we can use similar proof techniques as in Theorem 9 to obtain regret bounds. There are two special cases of particular interest here.*

1. *(**Strongly convex and smooth**) Suppose f_t's are strongly convex and smooth and we have access to the strong convex parameter of f_t (say κ_t) at each iteration t. Suppose Algorithm 1 is run with $\hat{H}_t = \kappa_t I$. Then its regret is $\tilde{O}\left(d^{3.5}\nu^2\sqrt{T}\right)$.*
2. *(**Smooth**) Suppose f_t's are smooth and we run Algorithm 1 with $\hat{H}_t = 0$. Then its regret can be bounded by $\tilde{O}\left(d^{7/3}T^{2/3}\right)$.*

Before we present a proof of this Theorem, we present some useful intermediate results.

D.1. Intermediate Results

Lemma 12 (Initial focus region) *For any $\alpha \geq \nu + 2\sqrt{\nu}$,*

$$F_1 \subseteq \mathcal{X} \subseteq B_{\alpha,\nabla^2 R(\mathbf{x}_1)}(\mathbf{x}_1).$$

Proof Consider property (P4) of self-concordant barriers stated in Equation (21) of Appendix G. It says that for any $\mathbf{x} \in \mathrm{int}(\mathcal{X})$

$$\mathcal{X} \cap \{\mathbf{y} : \langle \nabla R(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle \geq 0\} \subseteq B_{\nu+2\sqrt{\nu},\nabla^2 R(\mathbf{x})}(\mathbf{x}).$$

Since $\mathbf{x}_1$ is the minimizer of $R(\mathbf{x})$ over $\mathcal{X}$, and since it is in the interior of $\mathcal{X}$, we have $\nabla R(\mathbf{x}_1) = 0$. So, from property (P4) we have $\mathcal{X} \subseteq B_{\nu+2\sqrt{\nu},\nabla^2 R(\mathbf{x}_1)}(\mathbf{x}_1)$. The lemma then immediately follows from the definition of F_1 (recall $F_1 = \mathcal{X}_\xi \subseteq \mathcal{X}$). ∎

Lemma 13 (Lemma 5 of Bubeck et al. (2017)) *Let $\mathcal{K}$ be a convex body and $\mathcal{E}$ be an ellipsoid centered at the origin. Suppose that $Vol(\mathcal{K} \cap \mathcal{E}) \geq \frac{1}{2} Vol(\mathcal{K})$. Then $\mathcal{K} \subset 4d\mathcal{E}$.*

Lemma 14 (Lemma 4.6 of Hazan (2016)) *Let B_0 be a symmetric positive definite matrix and let $\{B_t\}_{t=1}^T$ be symmetric positive semi-definite matrices. Let $A_t = \sum_{s=0}^t B_s$. Then*

$$\sum_{t=1}^T tr\left(A_t(A_t - A_{t-1})\right) \leq \log_2 \frac{\det A_T}{\det A_0}.$$

Lemma 15 (Wainwright (2019)) *Let $X_1, \ldots X_K \in \mathbb{R}$ be a martingale difference sequence, where $\mathbb{E}\left[X_i | \mathcal{F}_{i-1}\right] = 0$. Assume that X_i satisfy the following tail condition, for some scalar $B_i > 0$*

$$\mathbb{P}\left(\left|\frac{X_i}{B_i}\right| \geq z \middle| \mathcal{F}_{i-1}\right) \leq 2\exp(-z^2).$$

Then

$$\mathbb{P}\left(\left|\sum_{i=1}^K X_i\right| \geq z\right) \leq 2\exp\left(-c\frac{z^2}{\sum_{i=1}^K B_i^2}\right),$$

where $c > 0$ is a universal constant.

Lemma 16 (Matrix Azuma; Tropp (2012)) *Consider a finite adapted sequence $\{X_i\}$ of symmetric matrices in dimension d, and fixed sequence $\{A_i\}$ of symmetric matrices that satisfy*

$$\mathbb{E}_i\left[X_i\right] = 0 \text{ and } X_i^2 \preceq A_i^2 \text{ almost surely.}$$

Compute the variance parameter $\sigma^2 \coloneqq \|\sum_i A_i^2\|_2$. Then, for all $t \geq 0$,

$$\mathbb{P}\left(\lambda_{max}\left(\sum_i X_i\right) \geq t\right) \leq d e^{-t^2/8\sigma^2}.$$

D.2. Proof of Theorem 9

To prove Theorem 9, we work with a slightly modified algorithm and show that with high probability, the iterates of the modified algorithm are exactly same as the actual algorithm. Consequently, proving the proposition for the modified algorithm entails that the Theorem also holds for the actual algorithm. In the modified algorithm, we slightly change $\hat{g}_t, \hat{H}_t$ and work with the following sequence of random variables

$$\hat{g}_t = \lambda^{-1} d\iota_t f_t(\mathbf{y}_t) M_t^{1/2} \mathbf{v}_{1,t}, \quad \hat{H}_t = \frac{\iota_t}{2}\left(A_t + A_t^T\right).$$

where ι_t is an indicator random variable which is equal to 1 if and only if the following event happen

$$\sup_{\mathbf{x} \in F_t} \left| \sum_{s=1}^{t-1} (\hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) - \iota_s f_s(\mathbf{x}) + \iota_s f_s(\mathbf{x}_s)) \right| \leq \frac{1}{\eta_1}.$$

This event happens when the cumulative loss estimate $\sum_{s=1}^{t-1}\hat{f}_s(\mathbf{x})$ is close to the true cumulative loss $\sum_{s=1}^{t-1} f_s(\mathbf{x})$ over the focus region F_t. We assume the algorithm is run with these modified estimates of gradients and Hessians. The main benefit of working with the modified gradient and Hessian estimates is that they are more amenable to analysis. Our proof shows that with high probability, the modified random variables $\hat{g}_t, \hat{H}_t$ are exactly equal to the original definitions of $\hat{g}_t, \hat{H}_t$. In particular, we show that in every iteration before the algorithm restarts, $\iota_t = 1$ with high probability. This entails that the actions output by the modified algorithm are exactly same as the actual algorithm, with high probability. As a result, it suffices to prove Theorem 9 for the modified algorithm.

We now derive some useful properties of the iterates produced by the modified algorithm. Some of these properties are very basic and pertain to the well-behavedness of the iterates of the algorithm. For example, the first property ensures that $\mathbf{y}_t$ always lies in $\mathcal{X}$.

Lemma 17 (Properties of iterates) *Consider the setting of Theorem 9. Let $\mathcal{T}$ be the minimum between T and the first iteration at which the modified algorithm restarts. For any $t < \mathcal{T}$ such that $\eta_t \leq 10\eta_1$, the iterates of the algorithm satisfy the following stability properties*

1. *M_t is positive definite and $\mathbf{y}_t \in \mathcal{X}$.*

2. *$R_t(\mathbf{x})$ is a strictly convex function over F_t.*

3. *For all $\mathbf{x} \in F_t$, $\|\mathbf{x} - \mathbf{x}_t\|_{M_t} \leq 4d\alpha$ and $\nabla^2 R(\mathbf{x}) \succeq \frac{1}{(1+4d\alpha)^2}\nabla^2 R(\mathbf{x}_t)$.*

4. *$\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t} \leq 2\lambda^{-1}dB\eta_t$ and $\|I - M_t^{-1/2}M_{t+1}M_t^{-1/2}\|_2 \leq 12\lambda^{-2}d^2B\eta_t$.*

5. *if $\iota_t = 0$, then $\iota_t = \iota_{t+1} = \cdots = \iota_{\mathcal{T}}$, $\mathbf{x}_t = \mathbf{x}_{t+1} \cdots = \mathbf{x}_{\mathcal{T}}$ and $F_t = F_{t+1} \cdots = F_{\mathcal{T}}$.*

Proof We use induction to prove the lemma.

Base Case (t=1).

1. First note that $M_1 = \nabla^2 R(\mathbf{x}_1)$. From property P3 of SCB stated in Appendix G, we know that $R(\mathbf{x})$ is strictly convex over $\mathrm{int}(\mathcal{X})$. So M_1 is positive definite and invertible. Moreover, from the Dikin ellipsoid property (P1) of SCB stated in Section 2, and from our choice of λ, it is easy to see that $\mathbf{y}_1 \in \mathcal{X}$.

2. The strict convexity property of $R(\mathbf{x})$ over F_1 follows from property P3 of SCB stated in Appendix G.

3. To show that for all $\mathbf{x} \in F_1$, $\nabla^2 R(\mathbf{x}) \succeq \frac{1}{(1+4d\alpha)^2}\nabla^2 R(\mathbf{x}_1)$, we rely on Assumption 3 and Lemma 12. In particular, from Assumption 3 we know that if $\|\mathbf{x} - \mathbf{x}_1\|_{\nabla^2 R(\mathbf{x}_1)} \leq \lambda$, then $\nabla^2 R(\mathbf{x}) \succeq \frac{1}{(1+\lambda)^2}\nabla^2 R(\mathbf{x}_1)$. Moreover, from Lemma 12 we know that any $\mathbf{x} \in \mathcal{X}$ satisfies
$$\|\mathbf{x} - \mathbf{x}_1\|_{\nabla^2 R(\mathbf{x}_1)} \leq \nu + 2\sqrt{\nu} \leq \alpha.$$
Combining these two facts gives us the required result.

4. We now show that $\mathbf{x}_2$ and $\mathbf{x}_1$ are close to each other. Note that $\mathbf{x}_2$ is the minimizer of the following objective

$$\mathbf{x}_2 \in \operatorname*{argmin}_{\mathbf{x} \in F_1} \eta_1 \langle \hat{g}_1, \mathbf{x} \rangle + \Phi_{R_2}(\mathbf{x}, \mathbf{x}_1). \tag{3}$$

From first order optimality conditions we have

$$\forall \mathbf{x} \in F_1, \quad \langle \nabla R_2(\mathbf{x}_2) - \nabla R_2(\mathbf{x}_1) + \eta_1 \hat{g}_1, \mathbf{x} - \mathbf{x}_2 \rangle \geq 0.$$

Substituting $\mathbf{x}_1$ in the above equation gives us

$$\langle \nabla R_2(\mathbf{x}_2) - \nabla R_2(\mathbf{x}_1) + \eta_1 \hat{g}_1, \mathbf{x}_1 - \mathbf{x}_2 \rangle \geq 0.$$

This can equivalently be written as

$$\left\langle \nabla R(\mathbf{x}_2) - \nabla R(\mathbf{x}_1) + \eta_1 \hat{g}_1 + \eta_1 \hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \leq 0. \tag{4}$$

Now suppose $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} > 2\lambda^{-1} dB\eta_1$. Then we have

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}_2) - \nabla R(\mathbf{x}_1) + \eta_1 \hat{g}_1 + \eta_1 \hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \\
&\overset{(a)}{\geq} \frac{\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}^2}{1 + \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}} + \left\langle \eta_1 \hat{g}_1 + \eta_1 \hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \\
&\overset{(b)}{\geq} \frac{\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}^2}{1 + \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}} - \eta_1 \|\hat{g}_1\|_{M_1}^* \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \\
&= \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \left(\frac{\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}}{1 + \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}} - \eta_1 \|\hat{g}_1\|_{M_1}^* \right),
\end{aligned}$$

where (a) follows from property P7 of SCBs stated in Appendix G and (b) follows from the fact that $\hat{H}_1$ is a positive semi-definite matrix. Next, consider the following

$$(\|\hat{g}_1\|_{M_1}^*)^2 = \hat{g}_1^T M_1^{-1} \hat{g}_1 = \lambda^{-2} d^2 f_1^2(\mathbf{y}_1) \mathbf{v}_{1,1}^T \mathbf{v}_{1,1} \leq \lambda^{-2} d^2 B^2.$$

Substituting this in the previous inequality and using the fact that $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} > 2\lambda^{-1} dB\eta_1$ gives us

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}_2) - \nabla R(\mathbf{x}_1) + \eta_1 \hat{g}_1 + \eta_1 \hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \\
&\quad \geq \lambda^{-1} dB\eta_1 \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \left(\frac{2}{1 + 2\lambda^{-1} dB\eta_1} - 1 \right) \\
&\quad \overset{(a)}{>} 0,
\end{aligned}$$

where (a) follows from the fact that $\lambda^{-1} dB\eta_1 < 1/2$. This contradicts the first order optimality condition in Equation (4). This shows that $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \leq 2\lambda^{-1} dB\eta_1$.

Next, we show that $M_1^{-1/2} M_2 M_1^{-1/2}$ is close to identity. From the definitions of M_1, M_2, we have

$$M_1^{-1/2} M_2 M_1^{-1/2} - I = M_1^{-1/2} (\nabla^2 R(\mathbf{x}_2) - \nabla^2 R(\mathbf{x}_1)) M_1^{-1/2} + \eta_1 M_1^{-1/2} \hat{H}_1 M_1^{-1/2}.$$

Since $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \leq 2\lambda^{-1}dB\eta_1$, we can rely on property P2 of SCB stated in Section 2 to infer that

$$\nabla^2 R(\mathbf{x}_2) \preceq \frac{1}{(1-2\lambda^{-1}dB\eta_1)^2}\nabla^2 R(\mathbf{x}_1) \preceq (1+6\lambda^{-1}dB\eta_1)\nabla^2 R(\mathbf{x}_1),$$

where the last inequality follows since $\lambda^{-1}dB\eta_1 < 1/10$. Next, note that $\hat{H}_1$ can be written as

$$\hat{H}_1 = \mathbb{E}\left[\frac{\lambda^{-2}}{2}d^2 f_1(\mathbf{y}_1) M_1^{1/2}\left(\mathbf{v}_{1,1}\mathbf{v}_{2,1}^T + \mathbf{v}_{2,1}\mathbf{v}_{1,1}^T\right) M_1^{1/2}\right].$$

So we have $M_1^{-1/2}\hat{H}_1 M_1^{-1/2} = \mathbb{E}\left[\frac{\lambda^{-2}}{2}d^2 f_1(\mathbf{y}_1)\left(\mathbf{v}_{1,1}\mathbf{v}_{2,1}^T + \mathbf{v}_{2,1}\mathbf{v}_{1,1}^T\right)\right]$ which is a bounded quantity. Substituting the previous two bounds in our expression for $M_1^{-1/2}M_2M_1^{-1/2} - I$ we get

$$\|M_1^{-1/2}M_2M_1^{-1/2} - I\|_2 \leq 6\lambda^{-1}dB\eta_1 + \lambda^{-2}d^2B\eta_1$$

5. Note that ι_1 is always equal to 1. So the last property trivially holds. This finishes the proof of the base case.

Induction Step. Suppose the proposition holds for the first $t-1$ iterations. We now show that it also holds for the t^{th} iteration.

1. The first part on positive definiteness of M_t and $\mathbf{y}_t \in \mathcal{X}$ follows from the same arguments as in the base case.

2. Note that $R_t(\mathbf{x}) = R(\mathbf{x}) + \sum_{s=0}^{t-1}\frac{\eta_s}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s)$. Since $\hat{H}_s$ is positive semi-definite, we have $\nabla^2 R_t(\mathbf{x}) \succeq \nabla^2 R(\mathbf{x})$. The strict convexity of $R_t(\mathbf{x})$ then follows from the fact that $R(\mathbf{x})$ is strictly convex over $\mathrm{int}(\mathcal{X})$.

3. The focus region update condition of our algorithm (lines 21-25 of Algorithm 1) always ensures that
$$\mathrm{Vol}(F_t \cap B_{\alpha,M_t}(\mathbf{x}_t)) \geq \frac{1}{2}\mathrm{Vol}(F_t).$$
So, from Lemma 13 we know that for any $\mathbf{x} \in F_t$, $\|\mathbf{x}-\mathbf{x}_t\|_{M_t} \leq 4d\alpha$. By relying on Assumption 3 on SCB, we then get
$$\forall \mathbf{x} \in F_t,\ \nabla^2 R(\mathbf{x}) \succeq \frac{1}{(1+4d\alpha)^2}\nabla^2 R(\mathbf{x}_t).$$

4. We now prove stability of the iterates. In particular, we show that $\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t} \leq 2\lambda^{-1}dB\eta_t$. If $\iota_{t-1} = 0$, then this trivially holds (because $\mathbf{x}_{t+1} = \mathbf{x}_t$). So lets consider the case where $\iota_{t-1} = 1$. From the first order optimality conditions, we have
$$\forall \mathbf{x} \in F_t, \quad \langle \nabla R_{t+1}(\mathbf{x}_{t+1}) - \nabla R_{t+1}(\mathbf{x}_t) + \eta_t\hat{g}_t, \mathbf{x} - \mathbf{x}_{t+1}\rangle \geq 0. \tag{5}$$
Note that from our definition of F_t, F_{t-1} we always have $F_t \subseteq F_{t-1}$ and $\mathbf{x}_t \in F_t$. So substituting $\mathbf{x}_t$ in the above equation and rearranging terms gives us
$$\left\langle \nabla R(\mathbf{x}_{t+1}) - \nabla R(\mathbf{x}_t) + \eta_t\hat{g}_t + \sum_{s=1}^{t}\eta_s\hat{H}_s(\mathbf{x}_{t+1}-\mathbf{x}_t), \mathbf{x}_{t+1} - \mathbf{x}_t\right\rangle \leq 0.$$

Now suppose $\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t} > 2\lambda^{-1} dB\eta_t$. Then we have

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}_{t+1}) - \nabla R(\mathbf{x}_t) + \eta_t \hat{g}_t + \sum_{s=1}^{t} \eta_s \hat{H}_s(\mathbf{x}_{t+1} - \mathbf{x}_t), \mathbf{x}_{t+1} - \mathbf{x}_t \right\rangle \\
&\quad \overset{(a)}{\geq} \frac{\|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2_{\nabla^2 R(\mathbf{x}_t)}}{1 + \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{\nabla^2 R(\mathbf{x}_t)}} + \|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2_{\eta_{1:t}\hat{H}_{1:t}} + \langle \eta_t \hat{g}_t, \mathbf{x}_{t+1} - \mathbf{x}_t \rangle \\
&\quad \overset{(b)}{\geq} \frac{\|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2_{\nabla^2 R(\mathbf{x}_t)}}{1 + \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{\nabla^2 R(\mathbf{x}_t)}} + \|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2_{\eta_{1:t-1}\hat{H}_{1:t-1}} - \eta_t \|\hat{g}_t\|^*_{M_t} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t},
\end{aligned}$$

where (a) follows from property P7 of SCBs stated in Appendix G and (b) follows from the fact that $\hat{H}_t$ is a positive semi-definite matrix. Here $\eta_{1:t}\hat{H}_{1:t} = \sum_{s=1}^{t} \eta_s \hat{H}_s$. Continuing

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}_{t+1}) - \nabla R(\mathbf{x}_t) + \eta_t \hat{g}_t + \sum_{s=1}^{t} \eta_s \hat{H}_s(\mathbf{x}_{t+1} - \mathbf{x}_t), \mathbf{x}_{t+1} - \mathbf{x}_t \right\rangle \\
&\quad \overset{(b)}{\geq} \frac{\|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2_{M_t}}{1 + \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t}} - \eta_t \|\hat{g}_t\|^*_{M_t} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t},
\end{aligned}$$

Next, consider the following

$$(\|\hat{g}_t\|^*_{M_t})^2 = \hat{g}_t^T M_t^{-1} \hat{g}_t = \lambda^{-2} d^2 f_t^2(\mathbf{y}_t) \mathbf{v}_{1,t}^T \mathbf{v}_{1,t} \leq \lambda^{-2} d^2 B^2.$$

Substituting this in the previous inequality and using the fact that $\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t} > 2\lambda^{-1} dB\eta_t$ gives us

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}_{t+1}) - \nabla R(\mathbf{x}_t) + \eta_t \hat{g}_t + \sum_{s=1}^{t} \eta_s \hat{H}_s(\mathbf{x}_{t+1} - \mathbf{x}_t), \mathbf{x}_{t+1} - \mathbf{x}_t \right\rangle \\
&\quad \geq \lambda^{-1} dB\eta_t \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t} \left(\frac{2}{1 + 2\lambda^{-1} dB\eta_t} - 1 \right) \\
&\quad \overset{(a)}{>} 0,
\end{aligned}$$

where (a) follows from the fact that $\lambda^{-1} dB\eta_t < 1/2$. This contradicts the first order optimality condition in Equation (5). This shows that $\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{M_t} \leq 2\lambda^{-1} dB\eta_t$.

Next, we show that $M_t^{-1/2} M_{t+1} M_t^{-1/2}$ is close to identity. From the definitions of M_t, M_{t+1}, we have

$$M_t^{-1/2} M_{t+1} M_t^{-1/2} - I = M_t^{-1/2} (\nabla^2 R(\mathbf{x}_{t+1}) - \nabla^2 R(\mathbf{x}_t)) M_t^{-1/2} + \eta_t M_t^{-1/2} \hat{H}_t M_t^{-1/2}.$$

Using similar arguments as in the base case, we get

$$\nabla^2 R(\mathbf{x}_{t+1}) \preceq (1 + 6\lambda^{-1} dB\eta_t) \nabla^2 R(\mathbf{x}_t), \ M_t^{-1/2} \hat{H}_t M_t^{-1/2} = \mathbb{E}_t \left[\frac{\lambda^{-2}}{2} d^2 f_t(\mathbf{y}_t) \left(\mathbf{v}_{1,t} \mathbf{v}_{2,t}^T + \mathbf{v}_{2,t} \mathbf{v}_{1,t}^T \right) \right].$$

Substituting these quantities in our expression for $M_t^{-1/2} M_{t+1} M_t^{-1/2} - I$ we get

$$\|M_t^{-1/2} M_{t+1} M_t^{-1/2}\|_2 \leq 12\lambda^{-2} d^2 B\eta_t.$$

5. The last property that remains to be shown is that if $\iota_t = 0$, then $\iota_t = \iota_{t+1} = \cdots = \iota_{\mathcal{T}}$, $\mathbf{x}_t = \mathbf{x}_{t+1} \cdots = \mathbf{x}_{\mathcal{T}}$ and $F_t = F_{t+1} \cdots = F_{\mathcal{T}}$. We assume $\iota_{t-1} = 1$, since otherwise the property is trivially true. Also note that $R_t(\mathbf{x})$ is strictly convex over F_t and so the OMD update in line 19 of Algorithm 1 has a unique minimizer.

   When $\iota_t = 0$, we have $\hat{g}_t = 0, \hat{H}_t = 0$. So the OMD update in line 19 of Algorithm 1 is given by $\mathbf{x}_{t+1} = \text{argmin}_{\mathbf{x}\in F_t} \Phi_{R_{t+1}}(\mathbf{x}, \mathbf{x}_t)$. Since $R_{t+1}(\mathbf{x}) = R_t(\mathbf{x})$ and $\mathbf{x}_t \in F_t$, it is easy to see that $\mathbf{x}_{t+1} = \mathbf{x}_t$. So the algorithm wouldn't make any progress in further rounds.

This finishes the proof of the lemma. ■

We now show that the focus region doesn't get updated more than $12d \log T$ times. This helps us show that the learning η_t doesn't gets too large.

Lemma 18 (Focus region updates) *Consider the setting of Theorem 9. Let $\mathcal{T}$ be the minimum between T and the first time at which the modified algorithm restarts. Then the focus region gets updated no more than $12d \log T$ times before $\mathcal{T}$. Moreover, $\eta_s \leq 10\eta_1$ for any $s \leq \mathcal{T}$.*

Proof We prove the proposition using contradiction. Assume that the focus region gets updated more than $12d \log T$ times before the algorithm restarts. Let $\tau < \mathcal{T}$ be the iteration where the focus region update happens for $12d \log T^{th}$ time. We now show that the restart condition should have triggered in iteration τ.

We have the following upper bound on the volume of $F_{\tau+1}$:

$$\text{Vol}(F_{\tau+1}) \leq \text{Vol}(F_\tau) \leq \frac{1}{T^{6d}} \text{Vol}(\mathcal{X}_\xi).$$

This follows from the fact that the volume of the focus region reduces by a factor of $1/2$ whenever the focus region update condition triggers. In the rest of the proof, we show that if the volume of focus region is less than $\frac{1}{T^{6d}}\text{Vol}(\mathcal{X}_\xi)$, then the restart condition should have triggered.

Step 1. First of all, for our choice of γ, we have $(1+\gamma)^{12d\log T} \leq 10$. Consequently, $\eta_\tau \leq 10\eta_1$. So the properties of the iterates we proved in Lemma 17 apply to our setting here. From this Lemma, we can infer that $\iota_\tau = 1$. Otherwise, we know that the focus region shouldn't have changed in the τ^{th} iteration (recall, in Lemma 17 we showed that if $\iota_\tau = 0$, then $F_\tau = F_{\tau+1}$). Moreoever, from this Lemma we can infer that $\forall t \leq \tau, \iota_t = 1$. So the cumulative loss estimate is close to the true cumulative loss and satisfies

$$\sup_{\mathbf{x}\in F_\tau} \left| \sum_{s=1}^{\tau-1} (\hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) - f_s(\mathbf{x}) + f_s(\mathbf{x}_s)) \right| \leq \frac{1}{\eta_1}.$$

Step 2. Let $\mathbf{u}_{\tau+1}$ be the minimizer of $\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x})$ over F_τ. Suppose $B\left(\mathbf{u}_{\tau+1}, \frac{1}{T^2}\right) \cap \mathcal{X}_\xi \subset F_\tau$. Then

$$\text{Vol}(F_\tau) \geq \text{Vol}\left(B\left(\mathbf{u}_{\tau+1}, \frac{1}{T^2}\right) \cap \mathcal{X}_\xi\right).$$

Next, from our assumption that $\mathcal{X}$ contains a euclidean ball of radius 1, we can infer that $\mathcal{X}_\xi = \xi\mathbf{x}_1 + (1-\xi)\mathcal{X}$ contains a ball of radius $(1-\xi)$ in it. Let $\tilde{B}$ be the ball of radius $(1-\xi)$

that lies in $\mathcal{X}_\xi$. By convexity of $\mathcal{X}$ and the fact that the diameter of $\mathcal{X}$ is less than or equal to T, we have

$$\left(1-\frac{1}{T^3}\right)\mathbf{u}_{\tau+1}+\frac{1}{T^3}\tilde{B}\subseteq B\left(\mathbf{u}_{\tau+1},\frac{1}{T^2}\right)\cap\mathcal{X}_\xi.$$

This shows that $\text{Vol}(F_\tau)\geq T^{-4d}\omega_d$, where ω_d is the volume of unit sphere in $\mathbb{R}^d$. Combining this with the previous upper bound on $\text{Vol}(F_\tau)$, we get

$$T^{-4d}\omega_d,\leq \text{Vol}(F_\tau)\leq T^{-6d}\text{Vol}(\mathcal{X})\overset{(a)}{\leq} T^{-5d}\omega_d,$$

where (a) follows from the fact that the diameter of $\mathcal{X}$ is upper bounded by T. We arrived at a contradiction. This shows that $B\left(\mathbf{u}_{\tau+1},\frac{1}{T^2}\right)\cap\mathcal{X}_\xi\not\subset F_\tau$.

Step 3. Since $B\left(\mathbf{u}_{\tau+1},\frac{1}{T^2}\right)\cap\mathcal{X}_\xi\not\subset F_\tau$, the following holds: $\exists\mathbf{x}\in\partial F_\tau\cap\text{int}(\mathcal{X}_\xi)$ such that $\|\mathbf{x}-\mathbf{u}_{\tau+1}\|_2\leq\frac{1}{T^2}$. Now, consider the following for such an $\mathbf{x}$

$$\begin{aligned}\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x})-\hat{f}_s(\mathbf{u}_{\tau+1}) &= \sum_{s=1}^{\tau}f_s(\mathbf{x})-f_s(\mathbf{u}_{\tau+1})\\ &\quad+\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x})-\hat{f}_s(\mathbf{u}_{\tau+1})-f_s(\mathbf{x})+f_s(\mathbf{u}_{\tau+1}).\end{aligned}$$

Since each f_s is T-Lipschitz, the first term in the RHS above is upper bounded by 1. Since the cumulative loss estimate is close to the true cumulative loss, the second term can be bounded as

$$\begin{aligned}\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x})-\hat{f}_s(\mathbf{u}_{\tau+1})-f_s(\mathbf{x})+f_s(\mathbf{u}_{\tau+1}) &\leq \frac{1}{\eta_1}+\hat{f}_\tau(\mathbf{x})-\hat{f}_\tau(\mathbf{u}_{\tau+1})-f_\tau(\mathbf{x})+f_\tau(\mathbf{u}_{\tau+1})\\ &\overset{(a)}{=}\frac{1}{\eta_1}+\left\langle\hat{g}_\tau-\mathbb{E}_\tau\left[\hat{g}_\tau\right],\mathbf{x}-\mathbf{u}_{\tau+1}\right\rangle,\end{aligned}$$

where (a) follows from the definitions of $f_\tau,\hat{f}_\tau$. Next, from Lemma 17 we know that for any $\mathbf{x}\in F_\tau$, $\|\mathbf{x}-\mathbf{x}_\tau\|_{M_\tau}\leq 4d\alpha$. Since $\mathbf{x},\mathbf{u}_{\tau+1}$ are points in F_τ, we have $\|\mathbf{x}-\mathbf{u}_{\tau+1}\|_{M_\tau}\leq 8d\alpha$. Using, this we get

$$\begin{aligned}\left\langle\hat{g}_\tau-\mathbb{E}_\tau\left[\hat{g}_\tau\right],\mathbf{x}-\mathbf{u}_{\tau+1}\right\rangle &\leq \|\hat{g}_\tau-\mathbb{E}_\tau\left[\hat{g}_\tau\right]\|_{M_\tau}^*\|\mathbf{x}-\mathbf{u}_{\tau+1}\|_{M_\tau}\\ &\leq 16\lambda^{-1}d^2\alpha B,\end{aligned}$$

where the last inequality follows from the fact that $\|\hat{g}_\tau\|_{M_\tau}^*$ is a bounded random variable which satisfies $\|\hat{g}_\tau\|_{M_\tau}^*\leq\lambda^{-1}dB$. Since $16\lambda^{-1}d^2B\eta_1\leq 1$, we have

$$\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x})-\hat{f}_s(\mathbf{u}_{\tau+1})-f_s(\mathbf{x})+f_s(\mathbf{u}_{\tau+1})\leq\frac{2}{\eta_1}.$$

This shows that $\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{u}_{\tau+1}) \leq \frac{4}{\eta_1}$. We now show that this implies the restart condition should have triggered. Consider the following

$$\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) &= \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{u}_{\tau+1}) \\
&\leq \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x}) \\
&= \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}\rangle - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) \\
&= \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}_{s+1}\rangle + \langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \\
&\quad - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) \\
&\overset{(a)}{\leq} \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \\
&\quad - \sum_{s=1}^{\tau} \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s),
\end{aligned}$$

where (a) follows from the stability of the iterates we proved in Lemma 17. Since $\mathbf{x}_{s+1}$ is the minimizer of $\min_{\mathbf{y}\in F_s} \eta_s \langle \hat{g}_s, \mathbf{y}\rangle + \Phi_{R_{s+1}}(\mathbf{y}, \mathbf{x}_s)$, we have the following from the first order optimality conditions

$$\langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \leq \frac{\Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_{s+1}) - \Phi_{R_{s+1}}(\mathbf{x}_{s+1}, \mathbf{x}_s)}{\eta_s}.$$

Using this in the previous display, we get

$$\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \sum_{s=1}^{\tau} \frac{\Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_{s+1})}{\eta_s} \\
- \sum_{s=1}^{\tau} \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s).
\end{aligned}$$

Rearranging the terms in the RHS above, we get

$$\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \frac{\Phi_R(\mathbf{x}, \mathbf{x}_1)}{\eta_1} - \frac{\Phi_{R_{\tau+1}}(\mathbf{x}, \mathbf{x}_{\tau+1})}{\eta_\tau} \\
+ \sum_{s=2}^{\tau} \left(\frac{1}{\eta_s} - \frac{1}{\eta_{s-1}}\right) \Phi_{R_s}(\mathbf{x}, \mathbf{x}_s).
\end{aligned}$$

Recall, $\mathbf{x} \in \partial F_\tau \cap \operatorname{int}(\mathcal{X}_\xi)$. Let τ' be such that $\mathbf{x} \in \partial B_{\alpha, M_{\tau'}}(\mathbf{x}_{\tau'})$. Then

$$\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \frac{\Phi_R(\mathbf{x}, \mathbf{x}_1)}{\eta_1} - \gamma \frac{\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'})}{\eta_{\tau'}}.$$

Since $\|\mathbf{x} - \mathbf{x}_{\mathcal{T}'}\|_{M_{\mathcal{T}'}} = \alpha$, we have the following lower bound on $\Phi_{R_{\mathcal{T}'}}(\mathbf{x}, \mathbf{x}_{\mathcal{T}'})$ which follows from property (P6) of SCB stated in Appendix G

$$\Phi_{R_{\mathcal{T}'}}(\mathbf{x}, \mathbf{x}_{\mathcal{T}'}) \geq \alpha - \log(1+\alpha).$$

For our choice of α, $\Phi_{R_{\mathcal{T}'}}(\mathbf{x}, \mathbf{x}_{\mathcal{T}'})$ can be lower bounded by $\alpha/2$. We now upper bound $\Phi_R(\mathbf{x}, \mathbf{x}_1)$. Since $\mathbf{x} \in \mathcal{X}_\xi$, using property P8 of SCB stated in Appendix G, we can upper bound $\Phi_R(\mathbf{x}, \mathbf{x}_1)$ as

$$\Phi_R(\mathbf{x}, \mathbf{x}_1) = R(\mathbf{x}) \leq 4\nu \log T.$$

Substituting the above two bounds in the previous display and using the fact that $\eta_\tau \leq 10\eta_1$, we get

$$\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y} \in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 20\lambda^{-2} d^2 B^2 \eta_1 T + \frac{4\nu \log T}{\eta_1} - \frac{\alpha\gamma}{20\eta_1} \leq -\frac{\beta}{\eta_1}.$$

This implies, the restart condition should have triggered. This shows that the focus region doesn't get updated more than $12d\log T$ times. ∎

Lemma 19 *Consider the setting of Theorem 9. Let $\mathcal{T}$ be the minimum between T and the first time at which the modified algorithm restarts. Then for any $t \leq \mathcal{T}$,*

$$M_t \preceq T^8(\nu + 2\sqrt{\nu})^2(\nabla^2 R(\mathbf{x}_1) + I).$$

Proof First note that the iterates generated by the algorithm lie in $\mathcal{X}_\xi$, where $\xi = T^{-4}$. So using property P8 of SCB stated in Appendix G, we have

$$\forall t \leq \mathcal{T}, \quad \nabla^2 R(\mathbf{x}_t) \preceq \left(\frac{\nu + 2\sqrt{\nu}}{\xi}\right)^2 \nabla^2 R(\mathbf{x}_1) = T^8(\nu + 2\sqrt{\nu})^2 \nabla^2 R(\mathbf{x}_1).$$

Next, since f_t is T Lipschitz and since $\mathcal{X}$ contains a euclidean ball of radius 1 in it, we have $\nabla^2 f_t(\mathbf{x}) \preceq TI$. We now use the above two inequalities to bound M_t

$$\begin{aligned} M_t = \nabla^2 R(\mathbf{x}_t) + \sum_{s=1}^{t-1} \eta_s \hat{H}_s &\preceq T^8(\nu + 2\sqrt{\nu})^2 \nabla^2 R(\mathbf{x}_1) + \sum_{s=1}^{t-1} \eta_s T I \\ &\overset{(a)}{\preceq} T^8(\nu + 2\sqrt{\nu})^2(\nabla^2 R(\mathbf{x}_1) + I), \end{aligned}$$

where (a) relied on the fact that $\eta_s \leq 10\eta_1$ for any $s \leq \mathcal{T}$ which we proved in Lemma 18. ∎

The following Lemma is concerned about concentration of loss estimates $\{\hat{f}_t\}_{t=1}^T$ computed by the modified algorithm. This Lemma helps us show that with high probability, the iterates of the modified and the original algorithms are exactly the same. Before we proceed, note that the focus region gets updated at most $12d\log T$ times before the algorithm restarts. So, for our choice of γ, we have $(1+\gamma)^{12d\log T} \leq 10$. Consequently, for all $t \leq \mathcal{T}$, $\eta_t \leq 10\eta_1$. So the results of Lemma 17 apply to all the iterates in the first $\mathcal{T}$ iterations of the modified algorithm.

Lemma 20 (Concentration of loss estimates) *Let $\mathcal{T}$ be the minimum between T and the first time at which the modified algorithm restarts. Then for any $t \le \mathcal{T}$, the following statement holds with probability at least $1 - T^{-2}$*

$$\sup_{\mathbf{x}\in F_t} \left| \sum_{s=1}^{t-1} \eta_1(\hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) - \iota_s f_s(\mathbf{x}) + \iota_s f_s(\mathbf{x}_s)) \right| \le \tilde{O}\left(\lambda^{-1} d^{5/2} \alpha B \eta_1 \sqrt{T}\right).$$

Proof First, note that

$$\begin{aligned} \hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) &= \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) + \langle \hat{g}_s, \mathbf{x}-\mathbf{x}_s\rangle \\ \iota_s f_s(\mathbf{x}) - \iota_s f_s(\mathbf{x}_s) &= \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) + \iota_s \langle \nabla f_s(\mathbf{x}_s), \mathbf{x}-\mathbf{x}_s\rangle. \end{aligned}$$

So $\hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) - \iota_s f_s(\mathbf{x}) + \iota_s f_s(\mathbf{x}_s) = \langle \hat{g}_s - \iota_s \nabla f_s(\mathbf{x}_s), \mathbf{x}-\mathbf{x}_s\rangle$. For any $\mathbf{x} \in F_t$, define random variables $Z_{\mathbf{x},s}$ as

$$Z_{\mathbf{x},s} = \begin{cases} \eta_1 \langle \hat{g}_s - \iota_s \nabla f_s(\mathbf{x}_s), \mathbf{x}-\mathbf{x}_s\rangle & \text{if } s \le \mathcal{T} \\ 0 \quad \text{otherwise} \end{cases}.$$

Since $\mathbb{E}_s[\hat{g}_s] = \iota_s \nabla f_s(\mathbf{x}_s)$, it is easy to see that $\{Z_{\mathbf{x},s}\}_{s=1}^T$ is a martingale difference sequence. Moreover, $Z_{\mathbf{x},s}$ is a bounded random variable. This follows from the fact that $\|\hat{g}_s\|_{M_s}^*$ is bounded and satisfies $\|\hat{g}_s\|_{M_s}^* \le \lambda^{-1} dB$. Moreover, for any $\mathbf{x} \in F_s$, $\|\mathbf{x}-\mathbf{x}_s\|_{M_s} \le 4d\alpha$ (see Lemma 17). So we have

$$|Z_{\mathbf{x},s}| \le \eta_1 | \langle \hat{g}_s - \iota_s \nabla f_s(\mathbf{x}_s), \mathbf{x}-\mathbf{x}_s\rangle | \le 8\lambda^{-1} d^2 \alpha B \eta_1.$$

By relying on standard concentration bounds for martingale difference sequences (see Lemma 15), we get that with probability at least $1-\delta$,

$$\sup_{t\le T} |\sum_{s=1}^{t-1} Z_{\mathbf{x},s}| = O\left(\lambda^{-1} d^2 \alpha B \eta_1 \sqrt{T \log T/\delta}\right).$$

Next, we bound $\sup_{\mathbf{x}\in F_t} \sup_{t\le T} |\sum_{s=1}^{t-1} Z_{\mathbf{x},s}|$ using ϵ-net arguments. Let $\mathcal{N}_\epsilon$ be an ϵ-net over F_t which satisfies the following: for every $\mathbf{x}$, there exists a $\mathbf{x}_\epsilon \in \mathcal{N}_\epsilon$ such that $\|\mathbf{x}-\mathbf{x}_\epsilon\|_{M_t} \le \epsilon$. Then

$$\sup_{\mathbf{x}\in F_t} \sup_{t\le T} |\sum_{s=1}^{t-1} Z_{\mathbf{x},s}| \le \underbrace{\sup_{\mathbf{x}\in F_t} \sup_{t\le T} |\sum_{s=0}^{t-1} Z_{\mathbf{x}_\epsilon,s}|}_{T_1} + \underbrace{\sup_{\mathbf{x}\in F_t} \sup_{t\le T} |\sum_{s=0}^{t-1} Z_{\mathbf{x}_\epsilon,s} - Z_{\mathbf{x},s}|}_{T_2}.$$

Using a simple union bound, T_1 can be bounded as

$$T_1 \le O\left(\lambda^{-1} d^2 \alpha B \eta_1 \sqrt{T \log T|\mathcal{N}_\epsilon|/\delta}\right) \overset{(a)}{\le} O\left(\lambda^{-1} d^{5/2} \alpha B \eta_1 \sqrt{T \log \frac{\alpha d T}{\epsilon\delta}}\right),$$

where the bound holds with probability at least $1-\delta$ and (a) holds since $\forall \mathbf{x} \in F_t, \|\mathbf{x}-\mathbf{x}_t\| \leq 4d\alpha$ and as a result $|\mathcal{N}_\epsilon| \leq \left(\frac{4d\alpha}{\epsilon}\right)^d$. T_2 can be bounded as follows

$$\begin{aligned}
\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=0}^{t-1} Z_{\mathbf{x}_\epsilon,s} - Z_{\mathbf{x},s}| &= \sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=0}^{t-1}\eta_1\left\langle \hat{g}_s - \iota_s\nabla f_s(\mathbf{x}_s), \mathbf{x}-\mathbf{x}_\epsilon\right\rangle| \\
&\overset{(a)}{\leq} 2\eta_1\lambda^{-1}dB\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}\left(\sum_{s=0}^{t-1}\|\mathbf{x}-\mathbf{x}_\epsilon\|_{M_s}\right) \\
&\overset{(b)}{\leq} 2(1+4d\alpha)^2\eta_1\lambda^{-1}dB\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}\left(\sum_{s=0}^{t-1}\|\mathbf{x}-\mathbf{x}_\epsilon\|_{M_t}\right) = O\left(\lambda^{-1}d^3\alpha^2 B\eta_1\epsilon T\right),
\end{aligned}$$

where (a) follows from the fact that $\|\hat{g}_s\|^*_{M_s} \leq \lambda^{-1}dB$ and (b) follows from Lemma 17 where we showed that $M_s \preceq (1+4d\alpha)^2 M_t$. Choosing $\epsilon = \frac{1}{\alpha\sqrt{dT}}$, and plugging the above bounds for T_1, T_2 in the upper bound for $\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=1}^{t-1} Z_{\mathbf{x},s}|$ gives us the required result. ■

Proof of Theorem 9. From Lemma 20, we know that with high probability, the iterates of the modified algorithm which relies on indicator variables ι_t are exactly same as the original algorithm. So it suffices to prove the regret bound for the modified algorithm. In the sequel, we work with the modified algorithm. Throughout the proof, we let $\mathcal{T}$ be the minimum between T and the first time step at which the algorithm restarts. Let τ be the minimum between $\mathcal{T}$ and the last time step where $\iota_\tau = 1$. Our goal is to bound the following quantity

$$\sum_{s=1}^{\mathcal{T}}\iota_s f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}}\sum_{s=1}^{\mathcal{T}}\iota_s f_s(\mathbf{x}) = \sum_{s=1}^{\tau} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}}\sum_{s=1}^{\tau} f_s(\mathbf{x}).$$

Case 1 ($\mathcal{T} = T$). We first consider the case where the restart condition didn't trigger in the first T iterations (i.e., $\mathcal{T} = T$). In this case, we show that the regret is $\tilde{O}\left(T^{1/2}\right)$. Since the restart condition hasn't triggered, we know that

$$\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau}\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{y}) \geq -\frac{\beta}{\eta_1}.$$

From the proof of Lemma 18, this implies $\forall \mathbf{x} \in \partial F_\tau \cap \mathrm{int}(\mathcal{X}_\xi)$

$$\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x}) - \min_{\mathbf{y}\in F_\tau}\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{y}) \geq \frac{4}{\eta_1}.$$

For the sake of clarity, we reproduce the argument we used in Lemma 18. To show this, we prove the contrapositive statement. Suppose $\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x}) - \min_{\mathbf{y}\in F_\tau}\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1}$ for some $\mathbf{x} \in \partial F_\tau \cap \mathrm{int}(\mathcal{X}_\xi)$. We now show that this implies the restart condition should have triggered. Consider the following

$$
\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) &\leq \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x}) \\
&= \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}\rangle - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) \\
&= \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}_{s+1}\rangle + \langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \\
&\quad - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) \\
&\overset{(a)}{\leq} \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \\
&\quad - \sum_{s=1}^{\tau} \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s),
\end{aligned}
$$

where (a) follows from the stability of the iterates we proved in Lemma 17. Since $\mathbf{x}_{s+1}$ is the minimizer of $\min_{\mathbf{y}\in F_s} \eta_s \langle \hat{g}_s, \mathbf{y}\rangle + \Phi_{R_{s+1}}(\mathbf{y}, \mathbf{x}_s)$, we have the following from the first order optimality conditions

$$
\langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \leq \frac{\Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_{s+1}) - \Phi_{R_{s+1}}(\mathbf{x}_{s+1}, \mathbf{x}_s)}{\eta_s}.
$$

Using this in the previous display, we get

$$
\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \sum_{s=1}^{\tau} \frac{\Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_{s+1})}{\eta_s} \\
- \sum_{s=1}^{\tau} \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s).
\end{aligned}
$$

Rearranging the terms in the RHS above, we get

$$
\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \frac{\Phi_R(\mathbf{x}, \mathbf{x}_1)}{\eta_1} - \frac{\Phi_{R_{\tau+1}}(\mathbf{x}, \mathbf{x}_{\tau+1})}{\eta_\tau} \\
+ \sum_{s=2}^{\tau} \left(\frac{1}{\eta_s} - \frac{1}{\eta_{s-1}}\right) \Phi_{R_s}(\mathbf{x}, \mathbf{x}_s).
\end{aligned}
$$

Recall, $\mathbf{x} \in \partial F_\tau \cap \mathrm{int}(\mathcal{X})$. Let τ' be such that $\mathbf{x} \in \partial B_{\alpha, M_{\tau'}}(\mathbf{x}_{\tau'})$. Then

$$
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 2\lambda^{-2}d^2B^2 \sum_{s=1}^{\tau} \eta_s + \frac{\Phi_R(\mathbf{x}, \mathbf{x}_1)}{\eta_1} - \gamma\frac{\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'})}{\eta_{\tau'}}.
$$

Since $\|\mathbf{x} - \mathbf{x}_{\tau'}\|_{M_{\tau'}} = \alpha$, we have the following lower bound on $\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'})$ which follows from property (P6) of SCB stated in Appendix G

$$
\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'}) \geq \alpha - \log(1+\alpha).
$$

For our choice of α, $\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'})$ can be lower bounded by $\alpha/2$. We now upper bound $\Phi_R(\mathbf{x}, \mathbf{x}_1)$. Since $\mathbf{x} \in \mathcal{X}_\xi$, using property P8 of SCB stated in Appendix G, we can upper bound $\Phi_R(\mathbf{x}, \mathbf{x}_1)$ as

$$\Phi_R(\mathbf{x}, \mathbf{x}_1) = R(\mathbf{x}) \leq 4\nu \log T.$$

Substituting the above two bounds in the previous display and using the fact that $\eta_\tau \leq 10\eta_1$, we get

$$\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 20\lambda^{-2}d^2B^2\eta_1 T + \frac{4\nu \log T}{\eta_1} - \frac{\alpha\gamma}{20\eta_1} \leq -\frac{\beta}{\eta_1}.$$

This implies, the restart condition should have triggered. But since the restart condition hasn't triggered, this result shows that $\forall \mathbf{x} \in \partial F_\tau \cap \mathrm{int}(\mathcal{X}_\xi), \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \geq \frac{4}{\eta_1}$. Next, since our cumulative loss estimate concentrates well around the true cumulative loss (i.e., $\iota_\tau = 1$), this implies

$$\forall \mathbf{x} \in \partial F_\tau \cap \mathrm{int}(\mathcal{X}_\xi), \quad \sum_{s=1}^{\tau} f_s(\mathbf{x}) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} f_s(\mathbf{y}) \geq \frac{2}{\eta_1}.$$

Since f_s's are convex, this implies the minimizer of $\min_{\mathbf{x}\in\mathcal{X}_\xi} \sum_{s=0}^{T} f_s(\mathbf{x})$ is in F_τ. So, the regret of the algorithm can be bounded as follows

$$\begin{aligned}
\mathrm{Reg}_T = \sum_{s=1}^{\tau} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{s=1}^{\tau} f_s(\mathbf{x}) &\overset{(a)}{\leq} 1 + \sum_{s=1}^{\tau} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}_\xi} \sum_{s=1}^{\tau} f_s(\mathbf{x}) \\
&= 1 + \sum_{s=1}^{\tau} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in F_\tau} \sum_{s=1}^{\tau} f_s(\mathbf{x}),
\end{aligned}$$

where (a) follows from the definition of $\mathcal{X}_\xi = (1-\xi)\mathcal{X} + \xi\mathbf{x}_1$ and the fact that the loss functions are Lipschitz and the diameter of $\mathcal{X}$ is bounded. Next, consider the following for any $\mathbf{x} \in F_\tau$

$$\begin{aligned}
\sum_{s=1}^{\tau} f_s(\mathbf{y}_s) - \sum_{s=1}^{\tau} f_s(\mathbf{x}) = &\underbrace{\sum_{s=1}^{\tau} [f_s(\mathbf{y}_s) - f_s(\mathbf{x}_s)]}_{T_1} + \underbrace{\sum_{s=1}^{\tau} \left[f_s(\mathbf{x}_s) - f_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) + \hat{f}_s(\mathbf{x})\right]}_{T_2} \\
&+ \underbrace{\sum_{s=1}^{\tau} \left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right]}_{T_3}.
\end{aligned}$$

Bounding T_1. We first bound T_1. Since f_s is a quadratic function with Hessian $\hat{H}_s$, we have

$$\sum_{s=1}^{\tau} f_s(\mathbf{y}_s) - f_s(\mathbf{x}_s) = \sum_{s=1}^{\tau} \lambda \left\langle \nabla f_s(\mathbf{x}_s), M_s^{-1/2}(\mathbf{v}_{1,s} + \mathbf{v}_{2,s})\right\rangle + \frac{\lambda^2}{2}(\mathbf{v}_{1,s} + \mathbf{v}_{2,s})^T M_s^{-1/2} \hat{H}_s M_s^{-1/2}(\mathbf{v}_{1,s} + \mathbf{v}_{2,s})$$

Let $Z_s = \lambda \left\langle \nabla f_s(\mathbf{x}_s), M_s^{-1/2}(\mathbf{v}_{1,s}+\mathbf{v}_{2,s})\right\rangle$ if $s \leq \tau$ and 0 if $s > \tau$. Note that $\{Z_s\}_{s=1}^T$ is a martingale difference sequence with each Z_s being bounded: $|Z_s| \leq 2dB$. This follows from the observation that $\nabla f_s(\mathbf{x}_s) = \mathbb{E}_s\left[\hat{g}_s\right]$ and the fact that $M_s^{-1/2}\hat{g}_s$ is a bounded random variable. By relying on standard concentration bounds for martingale difference sequences (see Lemma 15), we get that with probability at least $1-\delta$, $\sum_{s=1}^T Z_s = O\left(dB\sqrt{T\log 1/\delta}\right)$. We now bound the last term in the RHS above. Consider the following

$$\begin{aligned}(\mathbf{v}_{1,s}+\mathbf{v}_{2,s})^T M_s^{-1/2}\hat{H}_s M_s^{-1/2}(\mathbf{v}_{1,s}+\mathbf{v}_{2,s}) &\leq 4\|M_s^{-1/2}\hat{H}_s M_s^{-1/2}\|_2\\ &\leq 4\|M_{s+1}^{-1/2}\hat{H}_s M_{s+1}^{-1/2}\|_2\|M_s^{-1/2}M_{s+1}M_s^{-1/2}\|_2\end{aligned}$$

From Lemma 17 we know that $\|M_s^{-1/2}M_{s+1}M_s^{-1/2}\|_2 \leq 1+12\lambda^{-2}d^2B\eta_t \leq 2$. So we have

$$(\mathbf{v}_{1,s}+\mathbf{v}_{2,s})^T M_s^{-1/2}\hat{H}_s M_s^{-1/2}(\mathbf{v}_{1,s}+\mathbf{v}_{2,s}) \leq 8\|M_{s+1}^{-1/2}\hat{H}_s M_{s+1}^{-1/2}\|_2 = 8\|\hat{H}_s^{1/2}M_{s+1}^{-1}\hat{H}_s^{1/2}\|_2.$$

Define $N_t = (1+4d\alpha)^{-2}\nabla^2 R(\mathbf{x}_1) + \sum_{s=1}^{t-1}\eta_s\hat{H}_s$. From Lemma 17 we know that $\nabla^2 R(\mathbf{x}_t) \succeq (1+4d\alpha)^{-2}\nabla^2 R(\mathbf{x}_1)$. So $N_t \preceq M_t$ for all t. Using this in the previous inequality we get

$$\begin{aligned}(\mathbf{v}_{1,s}+\mathbf{v}_{2,s})^T M_s^{-1/2}\hat{H}_s M_s^{-1/2}(\mathbf{v}_{1,s}+\mathbf{v}_{2,s}) &\leq 8\|\hat{H}_s^{1/2}N_{s+1}^{-1}\hat{H}_s^{1/2}\|_2\\ &\leq 8\mathrm{tr}\left(N_{s+1}^{-1}\hat{H}_s\right)\\ &= \frac{8}{\eta_s}\mathrm{tr}\left(N_{s+1}^{-1}(N_{s+1}-N_s)\right)\end{aligned}$$

By relying on Lemma 14 we can upper bound $\sum_{s=1}^{\tau}\frac{8}{\eta_s}\mathrm{tr}\left(N_{s+1}^{-1}(N_{s+1}-N_s)\right)$ as

$$\sum_{s=1}^{\tau}\frac{8}{\eta_s}\mathrm{tr}\left(N_{s+1}^{-1}(N_{s+1}-N_s)\right) \leq \frac{8}{\eta_1}\sum_{s=1}^{\tau}\mathrm{tr}\left(N_{s+1}^{-1}(N_{s+1}-N_s)\right) \leq \frac{8}{\eta_1}\log\frac{\det N_T}{\det N_1} \tag{6}$$

From Lemma 19 we know that $N_T \preceq \mathrm{poly}(dT)$. Assuming $\nabla^2 R(\mathbf{x}_1) \succeq \frac{1}{\mathrm{poly}(dT)}I$, the RHS above can be upper bounded as $O\left(\frac{d\log dT}{\eta_1}\right)$. To summarize, we have the following upper bound T_1: $O\left(dB\sqrt{T\log 1/\delta} + \frac{d\log dT}{\eta_1}\right)$

Bounding T_2. Since $\iota_\tau = 1$, T_2 can be upper bounded as

$$\begin{aligned}T_2 &\leq \frac{1}{\eta_1} + \left[f_\tau(\mathbf{x}_\tau) - f_\tau(\mathbf{x}) - \hat{f}_\tau(\mathbf{x}_\tau) + \hat{f}_\tau(\mathbf{x})\right]\\ &= \frac{1}{\eta_1} + \left\langle \hat{g}_\tau - \mathbb{E}_\tau\left[\hat{g}_\tau\right], \mathbf{x}-\mathbf{x}_\tau\right\rangle \leq \frac{2}{\eta_1},\end{aligned}$$

where the last inequality follows from the fact that $\|\mathbf{x}-\mathbf{x}_\tau\|_{M_\tau} \leq 4d\alpha$ and $\|\hat{g}_\tau\|_{M_\tau}^* \leq \lambda^{-1}dB$.

Bounding T_3. To bound T_3, we consider the following

$$\begin{aligned}\sum_{s=1}^{\tau}\left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right] &= \sum_{s=1}^{\tau}\left\langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}\right\rangle - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s)\\ &= \sum_{s=1}^{\tau}\left\langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}_{s+1}\right\rangle + \left\langle \hat{g}_s, \mathbf{x}_{s+1}-\mathbf{x}\right\rangle - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s)\end{aligned}$$

Using similar arguments as at the beginning of Case 1, this can be bounded as

$$\sum_{s=1}^{\tau}\left[\hat{f}_s(\mathbf{x}_s)-\hat{f}_s(\mathbf{x})\right]\leq 2\lambda^{-2}d^2B^2\sum_{s=1}^{\tau}\eta_s+\frac{\Phi_R(\mathbf{x},\mathbf{x}_1)}{\eta_1},$$

Since $\mathbf{x}\in\mathcal{X}_\xi$, using property P8 of SCB stated in Appendix G, we can upper bound $\Phi_R(\mathbf{x},\mathbf{x}_1)$ as

$$\Phi_R(\mathbf{x},\mathbf{x}_1)=R(\mathbf{x})\leq 4\nu\log T.$$

Combining the bounds for T_1,T_2,T_3 shows that with probability at least $1-T^{-2}$ the regret is upper bounded by

$$\tilde{O}\left(dB\sqrt{T}+\frac{(\nu+d)}{\eta_1}+\lambda^{-2}d^2B^2\eta_1T\right)=\tilde{O}\left(d^{3.5}(d+\nu)^2\sqrt{T}\right).$$

Case 2 ($\mathcal{T}<T$). We now consider the case where the restart condition triggered at some iteration $\mathcal{T}<T$. Using the fact that the restart condition hasn't triggered in iteration $\mathcal{T}-1$ and using similar arguments as in the beginning of Case 1, we can again show that the minimizer of the cumulative loss over the entire domain lies in the focus region $F_{\mathcal{T}}$, and $\iota_{\mathcal{T}}=1$. So regret until $\mathcal{T}$ is given by

$$\begin{aligned}\mathrm{Reg}_{\mathcal{T}}=\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{y}_s)-\min_{\mathbf{x}\in\mathcal{X}}\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{x})&\overset{(a)}{\leq}1+\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{y}_s)-\min_{\mathbf{x}\in\mathcal{X}_\xi}\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{x})\\&=1+\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{y}_s)-\min_{\mathbf{x}\in F_{\mathcal{T}}}\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{x}),\end{aligned}$$

where (a) follows from the definition of $\mathcal{X}_\xi$. Using the same regret decomposition as in Case 1, for any $\mathbf{x}\in F_{\mathcal{T}}$

$$\begin{aligned}\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{y}_s)-\sum_{s=1}^{\mathcal{T}}f_s(\mathbf{x})=&\underbrace{\sum_{s=1}^{\mathcal{T}}[f_s(\mathbf{y}_s)-f_s(\mathbf{x}_s)]}_{T_1}+\underbrace{\sum_{s=1}^{\mathcal{T}}\left[f_s(\mathbf{x}_s)-f_s(\mathbf{x})-\hat{f}_s(\mathbf{x}_s)+\hat{f}_s(\mathbf{x})\right]}_{T_2}\\&+\underbrace{\sum_{s=1}^{\mathcal{T}}\left[\hat{f}_s(\mathbf{x}_s)-\hat{f}_s(\mathbf{x})\right]}_{T_3}.\end{aligned}$$

We use the same arguments as in Case 1 to bound T_1,T_2 as

$$T_1=O\left(dB\sqrt{T\log 1/\delta}+\frac{d\log dT}{\eta_1}\right),\quad T_2=\frac{2}{\eta_1}.$$

Since the restart condition triggered in round $\mathcal{T}$, T_3 is bounded by $-\frac{\beta}{\eta_1}$. Combining all these bounds, we get the following bound on regret

$$\mathrm{Reg}_{\mathcal{T}}\leq O\left(dB\sqrt{T\log 1/\delta}+\frac{d\log dT}{\eta_1}\right)+\frac{2}{\eta_1}-\frac{\beta}{\eta_1}.$$

For our choice of hyper-parameters, the above bound is less than 0.

Appendix E. Proof of Theorem 7

The proof of this Theorem uses similar arguments as the proof of "known Hessian" case in Appendix D. The additional complexity in proving Theorem 7 comes from dealing with Hessian estimates instead of exact Hessians used in Appendix D. In particular, in Theorem 7, we need to prove one additional result regarding the concentration of cumulative Hessian estimates.

We first introduce some notation we use in the proof. We let $r_t(\mathbf{x}) = f_t(\mathbf{x}) - q_t(\mathbf{x})$, where $q_t(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A_t \mathbf{x} + \langle \mathbf{b}_t, \mathbf{x}\rangle + c_t$. Recall, $\sup_{\mathbf{x}\in\mathcal{X}} |r_t(\mathbf{x})| \leq \epsilon$. We let $H_t = \frac{1}{2}(A_t + A_t^T)$ denote the Hessian of $q_t(\mathbf{x})$. Define random variable Z_t as

$$Z_t = 2^{-1}\lambda^{-2}d^2 f_t(\mathbf{y}_t)\left(\mathbf{v}_{1,t}\mathbf{v}_{2,t}^T + \mathbf{v}_{2,t}\mathbf{v}_{1,t}^T\right).$$

Since f_t is bounded, it is easy to see that Z_t is a bounded random variable (assuming M_t is positive definite and $\mathbf{y}_t \in \mathcal{X}$). In particular, Z_t can be bounded as

$$\|Z_t\|_2 \leq \lambda^{-2}d^2(B+\epsilon). \tag{7}$$

Another important thing to note here is that

$$\mathbb{E}_t\left[2^{-1}\lambda^{-2}d^2 q_t(\mathbf{y}_t)\left(\mathbf{v}_{1,t}\mathbf{v}_{2,t}^T + \mathbf{v}_{2,t}\mathbf{v}_{1,t}^T\right)\right] = M_t^{-1/2}H_t M_t^{-1/2}.$$

Consequently, $\|Z_t - M_t^{-1/2}H_tM_t^{-1/2}\|_2 \leq 2\lambda^{-2}d^2(B+\epsilon)$.

Similar to Appendix D, to prove Theorem 7, we work with a slightly modified algorithm and show that with high probability, the iterates of the modified algorithm are exactly same as the original algorithm. Consequently, proving the Theorem for the modified algorithm entails that the Theorem also holds for the actual algorithm. In the modified algorithm, we slightly change the random variables $\hat{g}_t, \hat{H}_t$ and work with the following sequence of random variables

$$\hat{g}_t = \lambda^{-1}d\iota_t f_t(\mathbf{y}_t)M_t^{1/2}\mathbf{v}_{1,t}, \quad \hat{H}_t = \frac{\lambda^{-2}}{2}d^2\iota_t f_t(\mathbf{y}_t)M_t^{1/2}\left(\mathbf{v}_{1,t}\mathbf{v}_{2,t}^T + \mathbf{v}_{2,t}\mathbf{v}_{1,t}^T\right)M_t^{1/2}.$$

where ι_t is an indicator random variable which is equal to 1 if and only if the following two events happen

$$\|I - \tilde{M}_t^{-1/2}M_t\tilde{M}_t^{-1/2}\|_2 \leq \frac{1}{10(1+8d\alpha)^2},$$

$$\sup_{\mathbf{x}\in F_t}\left|\sum_{s=1}^{t-1}(\hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) - \iota_s q_s(\mathbf{x}) + \iota_s q_s(\mathbf{x}_s))\right| \leq \frac{1}{\eta_1}.$$

Here, we define $\tilde{M}_t$ as $\tilde{M}_t = \nabla^2 R(\mathbf{x}_t) + \sum_{s=0}^{t-1}\eta_s\iota_s H_s$. Intuitively, the first event happens when M_t is spectrally close to $\tilde{M}_t$, and the second event happens when the cumulative loss estimate $\sum_{s=1}^{t-1}\hat{f}_s(\mathbf{x})$ is close to the true cumulative loss $\sum_{s=1}^{t-1}q_s(\mathbf{x})$. We assume the algorithm is run with these modified estimates of gradients and Hessians[2]. The main benefit

2. It should be noted that this is a hypothetical algorithm. We can not actually run this algorithm in practice as we can not compute ι_t

of working with the modified gradient and Hessian estimates is that they are bounded and are more amenable to analysis. Our proof shows that with high probability, the modified random variables $\hat{g}_t, \hat{H}_t$ are exactly equal to the original random variables. As a result, it suffices to prove Theorem 7 for the hypothetical algorithm.

We now derive some useful properties of the iterates produced by the modified algorithm.

Lemma 21 (Properties of iterates) *Consider the setting of Theorem 7. Let $\mathcal{T}$ be the minimum between T and the first iteration at which the modified algorithm restarts. For any $t < \mathcal{T}$ such that $\eta_t \leq 10\eta_1$, the iterates of the algorithm satisfy the following stability properties*

1. *M_t is positive definite and $\mathbf{y}_t \in \mathcal{X}$.*
2. *$R_t(\mathbf{x})$ is a strictly convex function over F_t.*
3. *For all $\mathbf{x} \in F_t$, $\|\mathbf{x} - \mathbf{x}_t\|_{M_t} \leq 4d\alpha$ and $\nabla^2 R(\mathbf{x}) \succeq \frac{1}{(1+8d\alpha)^2}\nabla^2 R(\mathbf{x}_t)$.*
4. *$\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{\tilde{M}_t} \leq c\eta_t$ and $\|I - \tilde{M}_t^{-1/2}\tilde{M}_{t+1}\tilde{M}_t^{-1/2}\|_2 \leq 4c\eta_t$. Here $c = 10(B+\epsilon)(\lambda^{-1}d + \lambda^{-2}d^3\alpha)$.*
5. *if $\iota_t = 0$, then $\iota_t = \iota_{t+1} = \cdots = \iota_{\mathcal{T}}$, $\mathbf{x}_t = \mathbf{x}_{t+1} \cdots = \mathbf{x}_{\mathcal{T}}$ and $F_t = F_{t+1} \cdots = F_{\mathcal{T}}$.*

Proof The proof uses similar arguments as in the proof of Lemma 17. So to avoid redundancy, we often directly rely on some of the results proved in Lemma 17. We use induction to prove the lemma.

Base Case (t=1).

1. First note that $\tilde{M}_1 = M_1 = \nabla^2 R(\mathbf{x}_1)$. So the proof follows from the proof of corresponding part in Lemma 17.
2. The proof of this part follows from the proof of corresponding part in Lemma 17.
3. The proof of this part follows from the proof of corresponding part in Lemma 17.
4. We now show that $\mathbf{x}_2$ and $\mathbf{x}_1$ are close to each other. Note that $\mathbf{x}_2$ is the minimizer of the following objective

$$\mathbf{x}_2 \in \operatorname*{argmin}_{\mathbf{x} \in F_1} \eta_1 \langle \hat{g}_1, \mathbf{x} \rangle + \Phi_{R_2}(\mathbf{x}, \mathbf{x}_1). \tag{8}$$

From first order optimality conditions we have

$$\forall \mathbf{x} \in F_1, \quad \langle \nabla R_2(\mathbf{x}_2) - \nabla R_2(\mathbf{x}_1) + \eta_1 \hat{g}_1, \mathbf{x} - \mathbf{x}_2 \rangle \geq 0.$$

Substituting $\mathbf{x}_1$ in the above equation gives us

$$\langle \nabla R_2(\mathbf{x}_2) - \nabla R_2(\mathbf{x}_1) + \eta_1 \hat{g}_1, \mathbf{x}_1 - \mathbf{x}_2 \rangle \geq 0.$$

This can equivalently be written as

$$\left\langle \nabla R(\mathbf{x}_2) - \nabla R(\mathbf{x}_1) + \eta_1 \hat{g}_1 + \eta_1 \hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \leq 0. \tag{9}$$

Now suppose $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} > c\eta_1$, where $c = 10(B+\epsilon)(\lambda^{-1}d + \lambda^{-2}d^3\alpha)$. Then we have

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}_2) - \nabla R(\mathbf{x}_1) + \eta_1\hat{g}_1 + \eta_1\hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \\
&\overset{(a)}{\geq} \frac{\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}^2}{1 + \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}} + \left\langle \eta_1\hat{g}_1 + \eta_1\hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \\
&\geq \frac{\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}^2}{1 + \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}} - \eta_1\left(\|\hat{g}_1\|_{M_1}^* + \|\hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1)\|_{M_1}^*\right)\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \\
&= \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}\left(\frac{\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}}{1 + \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}} - \eta_1\|\hat{g}_1\|_{M_1}^* - \eta_1\|\hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1)\|_{M_1}^*\right),
\end{aligned}$$

where (a) follows from property P7 of SCBs stated in Appendix G. Next, consider the following

$$(\|\hat{g}_1\|_{M_1}^*)^2 = \hat{g}_1^T M_1^{-1}\hat{g}_1 = \lambda^{-2}d^2 f_1^2(\mathbf{y}_1)\mathbf{v}_{1,1}^T\mathbf{v}_{1,1} \leq \lambda^{-2}d^2(B+\epsilon)^2.$$

$$\begin{aligned}
(\|\hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1)\|_{M_1}^*)^2 &= (\mathbf{x}_2 - \mathbf{x}_1)^T\hat{H}_1 M_1^{-1}\hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1)^T \\
&\leq \left(\frac{d^2 f_1(\mathbf{y}_1)}{2\lambda^2}\right)^2 \|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}^2 \|\mathbf{v}_{1,1}\mathbf{v}_{2,1}^T + \mathbf{v}_{2,1}\mathbf{v}_{1,1}^T\|_2^2 \\
&\overset{(a)}{\leq} 16\lambda^{-4}d^6(B+\epsilon)^2\alpha^2,
\end{aligned}$$

where (a) follows from the fact that $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \leq 4d\alpha$ proved in point (3). Substituting this in the previous inequality and using the fact that $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} > c\eta_1$ gives us

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}_2) - \nabla R(\mathbf{x}_1) + \eta_1\hat{g}_1 + \eta_1\hat{H}_1(\mathbf{x}_2 - \mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \right\rangle \\
&\geq \frac{c}{2}\eta_1\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1}\left(\frac{2}{1 + c\eta_1} - 1\right) \overset{(a)}{>} 0,
\end{aligned}$$

where (a) follows from the fact that $c\eta_1 < 1/2$. This contradicts the first order optimality condition in Equation (9). This shows that $\|\mathbf{x}_2 - \mathbf{x}_1\|_{M_1} \leq c\eta_1$.

Next, we show that $\tilde{M}_1^{-1/2}\tilde{M}_2\tilde{M}_1^{-1/2}$ is close to identity. From the definitions of $\tilde{M}_1, \tilde{M}_2$, we have

$$\tilde{M}_1^{-1/2}\tilde{M}_2\tilde{M}_1^{-1/2} - I = \tilde{M}_1^{-1/2}(\nabla^2 R(\mathbf{x}_2) - \nabla^2 R(\mathbf{x}_1))\tilde{M}_1^{-1/2} + \eta_1\tilde{M}_1^{-1/2}H_1\tilde{M}_1^{-1/2}.$$

Since $\|\mathbf{x}_2 - \mathbf{x}_1\|_{\tilde{M}_1} \leq c\eta_1 < 1$, we can rely on property P2 of SCB stated in Section 2 to infer that

$$\nabla^2 R(\mathbf{x}_2) \preceq \frac{1}{(1 - c\eta_1)^2}\nabla^2 R(\mathbf{x}_1) \preceq (1 + 3c\eta_1)\nabla^2 R(\mathbf{x}_1),$$

where the last inequality follows since $c\eta_1 < 1/10$. Next, note that H_1 can be written as

$$H_1 = \mathbb{E}\left[\frac{\lambda^{-2}}{2}d^2 f_1(\mathbf{y}_1)\tilde{M}_1^{1/2}\left(\mathbf{v}_{1,1}\mathbf{v}_{2,1}^T + \mathbf{v}_{2,1}\mathbf{v}_{1,1}^T\right)\tilde{M}_1^{1/2}\right].$$

So we have $\tilde{M}_1^{-1/2} H_1 \tilde{M}_1^{-1/2} = \mathbb{E}\left[\frac{\lambda^{-2}}{2} d^2 f_1(\mathbf{y}_1)\left(\mathbf{v}_{1,1}\mathbf{v}_{2,1}^T + \mathbf{v}_{2,1}\mathbf{v}_{1,1}^T\right)\right]$ which is a bounded quantity. Substituting the previous two bounds in our expression for $\tilde{M}_1^{-1/2}\tilde{M}_2\tilde{M}_1^{-1/2} - I$, we get

$$\|\tilde{M}_1^{-1/2}\tilde{M}_2\tilde{M}_1^{-1/2} - I\|_2 \leq 4c\eta_1.$$

5. Since $M_1 = \tilde{M}_1$, ι_1 is always equal to 1. So the last property trivially holds. This finishes the proof of the base case.

Induction Step. Suppose the Lemma holds for the first $t-1$ iterations. We now show that it also holds for the t^{th} iteration.

1. **Invertibility.** We first show that M_t is positive definite. If $\iota_{t-1} = 0$, then it is easy to see that M_t is equal to M_{t-1}, which we know is positive definite. So lets consider the where $\iota_{t-1} = 1$. We know that $\iota_1 = \iota_2 = \cdots = \iota_{t-1}$ and $\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{\tilde{M}_{t-1}} \leq c\eta_{t-1}$. Now, consider the following

$$\begin{aligned} M_t &= \nabla^2 R(\mathbf{x}_t) + \sum_{s=1}^{t-1} \eta_s \hat{H}_s \\ &= M_{t-1} + \eta_{t-1}\hat{H}_{t-1} + \nabla^2 R(\mathbf{x}_t) - \nabla^2 R(\mathbf{x}_{t-1}) \\ &\overset{(a)}{\succeq} M_{t-1} + \eta_{t-1}\hat{H}_{t-1} - 2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{\nabla^2 R(\mathbf{x}_{t-1})} \nabla^2 R(\mathbf{x}_{t-1}), \end{aligned}$$

where (a) follows from the property of self-concordant functions stated in Equation (1). From stability, we have

$$\begin{aligned} M_t &\succeq M_{t-1} + \eta_{t-1}\hat{H}_{t-1} - 2c\eta_{t-1}\nabla^2 R(\mathbf{x}_{t-1}) \\ &= M_{t-1}^{1/2}\left[I + \eta_{t-1}Z_{t-1} - 2c\eta_{t-1}M_{t-1}^{-1/2}\nabla^2 R(\mathbf{x}_{t-1})M_{t-1}^{-1/2}\right] M_{t-1}^{1/2}. \end{aligned}$$

We now show that $\left[I + \eta_{t-1}Z_{t-1} - 2c\eta_{t-1}M_{t-1}^{-1/2}\nabla^2 R(\mathbf{x}_{t-1})M_{t-1}^{-1/2}\right] \succ 0$. To show this, we rely on the following argument

$$\begin{aligned} &\|Z_{t-1} - 2cM_{t-1}^{-1/2}\nabla^2 R(\mathbf{x}_{t-1})M_{t-1}^{-1/2}\|_2 \\ &\quad \leq \|Z_{t-1}\|_2 + 2c\|M_{t-1}^{-1/2}\tilde{M}_{t-1}M_{t-1}^{-1/2}\|_2\|\tilde{M}_{t-1}^{-1/2}\nabla^2 R(\mathbf{x}_{t-1})\tilde{M}_{t-1}^{-1/2}\|_2 \\ &\quad \leq \lambda^{-2}d^2(B+\epsilon) + 3c \leq 4c, \end{aligned}$$

where the last inequality follows from the fact that Z_t is a bounded random variable, and $\|I - \tilde{M}_{t-1}^{-1/2}M_{t-1}\tilde{M}_{t-1}^{-1/2}\|_2 \leq \frac{1}{10}$ (consequently, $\|M_{t-1}^{-1/2}\tilde{M}_{t-1}M_{t-1}^{-1/2}\|_2 \leq \frac{3}{2}$). This shows that for our choice of hyper-parameters, M_t is invertible.

Valid Iterates. Next, we show that $\mathbf{y}_t \in \mathcal{X}$. If $\iota_{t-1} = 0$, then it is easy to see that this is the case (because $M_t = M_{t-1}$). So we assume $\iota_{t-1} = 1$. In this case, we first bound $\|I - \tilde{M}_t^{-1/2}M_t\tilde{M}_t^{-1/2}\|_2$ (i.e., we show that M_t and $\tilde{M}_t$ are spectrally close). Consider the following

$$\|I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\|_2 = \|\tilde{M}_t^{-1/2}(\tilde{M}_{t-1} - M_{t-1} + \eta_{t-1}(\hat{H}_{t-1} - H_{t-1}))\tilde{M}_t^{-1/2}\|_2 \quad (10)$$
$$\leq \|\tilde{M}_t^{-1/2}(\tilde{M}_{t-1} - M_{t-1})\tilde{M}_t^{-1/2}\|_2 \quad (11)$$
$$+ \eta_{t-1}\|\tilde{M}_t^{-1/2}(\hat{H}_{t-1} - H_{t-1})\tilde{M}_t^{-1/2}\|_2 \quad (12)$$

Consider the first term in the RHS above

$$\|\tilde{M}_t^{-1/2}(\tilde{M}_{t-1} - M_{t-1})\tilde{M}_t^{-1/2}\|_2 \leq \|\tilde{M}_{t-1}^{-1/2}(\tilde{M}_{t-1} - M_{t-1})\tilde{M}_{t-1}^{-1/2}\|\|\tilde{M}_t^{-1/2}\tilde{M}_{t-1}\tilde{M}_t^{-1/2}\|$$
$$\leq \frac{1}{5(1+8d\alpha)^2},$$

where the last inequality follows from the fact that $\|I - \tilde{M}_{t-1}^{-1/2} M_{t-1}\tilde{M}_{t-1}^{-1/2}\|_2 \leq \frac{1}{10(1+8d\alpha)^2}$ and the fact that $\tilde{M}_{t-1}$ is spectrally close to $\tilde{M}_t$. Now consider the second term in the RHS of Equation (10). Since $\hat{H}_{t-1} = M_{t-1}^{1/2} Z_{t-1} M_{t-1}^{1/2}$, we have

$$\begin{aligned}
&\|\tilde{M}_t^{-1/2}(\hat{H}_{t-1} - H_{t-1})\tilde{M}_t^{-1/2}\|_2 \\
&= \|\tilde{M}_t^{-1/2} M_{t-1}^{1/2}(Z_{t-1} - M_{t-1}^{-1/2} H_{t-1} M_{t-1}^{-1/2}) M_{t-1}^{1/2}\tilde{M}_t^{-1/2}\|_2 \\
&\leq \|\tilde{M}_{t-1}^{-1/2} M_{t-1}^{1/2}(Z_{t-1} - M_{t-1}^{-1/2} H_{t-1} M_{t-1}^{-1/2}) M_{t-1}^{1/2}\tilde{M}_{t-1}^{-1/2}\|_2 \\
&\overset{(a)}{\leq} 2\|Z_{t-1} - M_{t-1}^{-1/2} H_{t-1} M_{t-1}^{-1/2}\|_2 \\
&\overset{(b)}{\leq} 2\left(B + \epsilon\right) d^2 \lambda^{-2}.
\end{aligned}$$

where (a) follows from the fact that $\|I - \tilde{M}_{t-1}^{-1/2} M_{t-1}\tilde{M}_{t-1}^{-1/2}\|_2 \leq \frac{1}{10(1+8d\alpha)^2}$ and (b) follows from Equation (7). Combining the previous two displays shows that for our choice of η_1, $\|I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\|_2 \leq \frac{1}{4(1+8d\alpha)^2}$. This shows that M_t is spectrally close to $\tilde{M}_t$ and

$$\|\mathbf{y}_t - \mathbf{x}_t\|_{\tilde{M}_t} \leq 2\|\mathbf{y}_t - \mathbf{x}_t\|_{M_t} \leq 4\lambda < 1.$$

Since $\|\mathbf{y}_t - \mathbf{x}_t\|_{\tilde{M}_t} \geq \|\mathbf{y}_t - \mathbf{x}_t\|_{\nabla^2 R(\mathbf{x}_t)}$, using the Dikin Ellipsoid property of SCB stated in Section 2, we have $\mathbf{y}_t \in \mathcal{X}$.

2. The focus region update condition of our algorithm always ensures that

$$\mathrm{Vol}(F_t \cap B_{\alpha, M_t}(\mathbf{x}_t)) \geq \frac{1}{2}\mathrm{Vol}(F_t).$$

So, from Lemma 13 we know that for any $\mathbf{x} \in F_t$, $\|\mathbf{x} - \mathbf{x}_t\|_{M_t} \leq 4d\alpha$. Using this, together with the fact that $\|I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\|_2 \leq \frac{1}{4(1+8d\alpha)^2}$, we get $\|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t} \leq 8d\alpha$. By relying on Assumption 3 on SCB R, we then get

$$\forall \mathbf{x} \in F_t,\ \nabla^2 R(\mathbf{x}) \succeq \frac{1}{(1+8d\alpha)^2}\nabla^2 R(\mathbf{x}_t).$$

3. We now show that $R_t(\mathbf{x})$ is strictly convex over interior of F_t. Consider the following for any $\mathbf{x} \in \text{int}(F_t)$

$$\begin{aligned}
\nabla^2 R(\mathbf{x}) + \eta_{1:t-1}\hat{H}_{1:t-1} &\overset{(a)}{\succeq} \frac{1}{(1+8d\alpha)^2}\nabla^2 R(\mathbf{x}_t) + \eta_{1:t-1}\hat{H}_{1:t-1} \\
&\succeq \frac{1}{(1+8d\alpha)^2}\nabla^2 R(\mathbf{x}_t) + \eta_{1:t-1}H_{1:t-1} + (M_t - \tilde{M}_t) \\
&\overset{(b)}{\succeq} \frac{1}{(1+8d\alpha)^2}\tilde{M}_t - \frac{1}{4(1+8d\alpha)^2}\tilde{M}_t \\
&\succ 0,
\end{aligned}$$

where (a) follows from the previous property and (b) follows from the fact that $\|I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\|_2 \leq \frac{1}{4(1+8d\alpha)^2}$. This shows that R_t is strictly convex over F_t.

4. We now prove stability of the iterates. In particular, we show that $\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{\tilde{M}_t} \leq c\eta_t$. If $\iota_{t-1} = 0$, then this trivially holds. So lets consider the case where $\iota_{t-1} = 1$. From the first order optimality conditions, we have

$$\forall \mathbf{x} \in F_t, \quad \langle \nabla R_{t+1}(\mathbf{x}_{t+1}) - \nabla R_{t+1}(\mathbf{x}_t) + \eta_t \hat{g}_t, \mathbf{x} - \mathbf{x}_{t+1} \rangle \geq 0.$$

Note that from our definition of F_t, F_{t-1} we always have $F_t \subseteq F_{t-1}$ and $\mathbf{x}_t \in F_t$. So substituting $\mathbf{x}_t$ in the first equation gives us

$$\left\langle \nabla R(\mathbf{x}_{t+1}) - \nabla R(\mathbf{x}_t) + \eta_t \hat{g}_t + \sum_{s=1}^{t} \eta_s \hat{H}_s(\mathbf{x}_{t+1} - \mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_{t+1} \right\rangle \geq 0.$$

To prove the required result, we show that for any $\mathbf{x}$ such that $\|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t} > c\eta_t$, the following holds

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}) - \nabla R(\mathbf{x}_t) + \eta_{1:t-1}\hat{H}_{1:t-1}(\mathbf{x} - \mathbf{x}_t), \mathbf{x} - \mathbf{x}_t \right\rangle \\
&\qquad > \eta_t \|\hat{g}_t\|_{\tilde{M}_t}^* \|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t} + \eta_t \|\hat{H}_t(\mathbf{x} - \mathbf{x}_t)\|_{\tilde{M}_t}^* \|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t}.
\end{aligned}$$

This would then imply that the above optimality condition doesn't hold. We first lower bound the LHS of the above equation. Consider the following for any $\mathbf{x} \in F_t$ such that $\|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t} > c\eta_t$

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}) - \nabla R(\mathbf{x}_t) + \eta_{1:t-1}\hat{H}_{1:t-1}(\mathbf{x} - \mathbf{x}_t), \mathbf{x} - \mathbf{x}_t \right\rangle \\
&= \int_{s=0}^{1} (\mathbf{x} - \mathbf{x}_t)^T \left[\nabla^2 R(\mathbf{x}_t + s(\mathbf{x} - \mathbf{x}_t)) + \eta_{1:t-1}\hat{H}_{1:t-1}\right] (\mathbf{x} - \mathbf{x}_t) ds \\
&\overset{(a)}{\geq} \int_{s=0}^{\frac{c\eta_t}{\|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t}}} (\mathbf{x} - \mathbf{x}_t)^T \left[\nabla^2 R(\mathbf{x}_t + s(\mathbf{x} - \mathbf{x}_t)) + \eta_{1:t-1}\hat{H}_{1:t-1}\right] (\mathbf{x} - \mathbf{x}_t) ds \\
&\overset{(b)}{\geq} \frac{c\eta_t}{\|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t}} (\mathbf{x} - \mathbf{x}_t)^T \left[(1 - c\eta_t)^2 \nabla^2 R(\mathbf{x}_t) + \eta_{1:t-1}\hat{H}_{1:t-1}\right] (\mathbf{x} - \mathbf{x}_t),
\end{aligned}$$

where (a) uses the fact that $\nabla^2 R(\mathbf{x}) + \eta_{1:t-1}\hat{H}_{1:t-1}$ is a PSD matrix for any $\mathbf{x} \in F_t$ and (b) relies on property P1 of SCB stated in Equation (1). We further lower bound the RHS of the above equation as follows

$$\begin{aligned}
&(1 - c\eta_t)^2 \nabla^2 R(\mathbf{x}_t) + \eta_{1:t-1}\hat{H}_{1:t-1} \\
&\quad = (1 - c\eta_t)^2 \nabla^2 R(\mathbf{x}_t) + \eta_{1:t-1} H_{1:t-1} + M_t - \tilde{M}_t \\
&\quad \succeq (1 - c\eta_t)^2 \tilde{M}_t - \tilde{M}_t^{1/2}\left[I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\right]\tilde{M}_t^{1/2} \\
&\quad = \tilde{M}_t^{1/2}\left[(1 - c\eta_t)^2 I - \left(I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\right)\right]\tilde{M}_t^{1/2}.
\end{aligned}$$

Substituting this in the previous equation gives us

$$\begin{aligned}
&\left\langle \nabla R(\mathbf{x}) - \nabla R(\mathbf{x}_t) + \eta_{1:t-1}\hat{H}_{1:t-1}(\mathbf{x} - \mathbf{x}_t), \mathbf{x} - \mathbf{x}_t \right\rangle \\
&\quad \geq c\eta_t \|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t} \lambda_{min}\left((1 - c\eta_t)^2 I - \left(I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\right)\right) \\
&\quad > \frac{c\eta_t}{2}\|\mathbf{x} - \mathbf{x}_t\|_{\tilde{M}_t},
\end{aligned}$$

where the last inequality follows from the fact that $\|(I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\|_2 \leq \frac{1}{4(1+8d\alpha)^2}$, and our choice of hyper-parameters. Next, consider the following

$$\begin{aligned}
(\|\hat{g}_t\|_{\tilde{M}_t}^*)^2 &= \hat{g}_t^T \tilde{M}_t^{-1} \hat{g}_t \\
&= \lambda^{-2} d^2 f_t^2(\mathbf{y}_t) \mathbf{v}_{1,t}^T M_t^{1/2} \tilde{M}_t^{-1} M_t^{1/2} \mathbf{v}_{1,t} \\
&\leq 2\lambda^{-2} d^2 (B + \epsilon)^2.
\end{aligned}$$

$$\begin{aligned}
(\|\hat{H}_t(\mathbf{x} - \mathbf{x}_t)\|_{\tilde{M}_t}^*)^2 &= (\mathbf{x} - \mathbf{x}_t)^T \hat{H}_t \tilde{M}_t^{-1} \hat{H}_t (\mathbf{x} - \mathbf{x}_t)^T \\
&\leq \left(\frac{d^2 f_t(\mathbf{y}_t)}{\lambda^2}\right)^2 \|\mathbf{x} - \mathbf{x}_t\|_{M_t}^2 \|M_t^{1/2} \tilde{M}_t^{-1} M_t^{1/2}\|_2 \\
&\overset{(a)}{\leq} 32\lambda^{-4} d^6 (B + \epsilon)^2 \alpha^2,
\end{aligned}$$

where (a) follows from the fact that for any $\mathbf{x} \in F_t$, $\|\mathbf{x} - \mathbf{x}_t\|_{M_t} \leq 4d\alpha$. This shows that

$$\eta_t \|\hat{g}_t\|_{\tilde{M}_t}^* + \eta_t \|\hat{H}_t(\mathbf{x} - \mathbf{x}_t)\|_{\tilde{M}_t}^* \leq \frac{c\eta_t}{2}$$

This shows that $\mathbf{x}_{t+1}$ should satisfy $\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{\tilde{M}_t} \leq c\eta_t$.

To shows that $\tilde{M}_t$ and $\tilde{M}_{t+1}$ are spectrally close to each other, we rely on the closeness of $\mathbf{x}_{t+1}$ and $\mathbf{x}_t$ and use the same arguments as in the base case.

5. The last property that remains to be shown is that if $\iota_t = 0$, then $\iota_t = \iota_{t+1} = \cdots = \iota_{\mathcal{T}}$, $\mathbf{x}_t = \mathbf{x}_{t+1} \cdots = \mathbf{x}_{\mathcal{T}}$ and $F_t = F_{t+1} \cdots = F_{\mathcal{T}}$. We assume $\iota_{t-1} = 1$, since otherwise the property is trivially true. In this case, we know that $R_t(\mathbf{x})$ is strictly convex over F_t and so the Newton update in line 19 of Algorithm 1 has a unique minimizer.

   When $\iota_t = 0$, we have $\hat{g}_t = 0, \hat{H}_t = 0$. So the OMD update in line 19 of Algorithm 1 is given by $\mathbf{x}_{t+1} = \text{argmin}_{\mathbf{x} \in F_t} \Phi_{R_{t+1}}(\mathbf{x}, \mathbf{x}_t)$. Since $R_{t+1}(\mathbf{x}) = R_t(\mathbf{x})$ and $\mathbf{x}_t \in F_t$, it is easy to see that $\mathbf{x}_{t+1} = \mathbf{x}_t$. So the algorithm wouldn't make any progress in further rounds.

This finishes the proof of the lemma. ■

We now show that the focus region doesn't get updated more than $12d\log T$ times.

Lemma 22 (Focus region updates) *Consider the setting of Theorem 7. Let $\mathcal{T}$ be the minimum between T and the first time at which the modified algorithm restarts. Then the focus region gets updated no more than $12d\log T$ times before $\mathcal{T}$.*

Proof The proof uses similar arguments as in Lemma 18. We prove the Lemma using contradiction. Assume that the focus region gets updated more than $12d\log T$ times before the algorithm restarts. Let $\tau < \mathcal{T}$ be the iteration where the focus region update happens for $12d\log T^{th}$ time. We now show that the restart condition should have triggered in iteration τ.

We have the following upper bound on the volume of $F_{\tau+1}$:

$$\mathrm{Vol}(F_{\tau+1}) \leq \mathrm{Vol}(F_\tau) \leq \frac{1}{T^{6d}}\mathrm{Vol}(\mathcal{X}_\xi).$$

This follows from the fact that the volume of the focus region reduces by a factor of 1/2 whenever the focus region update condition triggers. In the rest of the proof, we show that if the volume of focus region is less than $\frac{1}{T^{6d}}\mathrm{Vol}(\mathcal{X}_\xi)$, then the restart condition should have triggered.

Step 1. First of all, for our choice of γ, we have $(1+\gamma)^{12d\log T} \leq 10$. Consequently, $\eta_\tau \leq 10\eta_1$. So the properties of the iterates we proved in Lemma 21 apply to our setting here. From this Lemma, we can infer that $\iota_\tau = 1$. Otherwise, we know that the focus region shouldn't have changed in the τ^{th} iteration (recall, in Lemma 21 we showed that if $\iota_\tau = 0$, then $F_\tau = F_{\tau+1}$). Moreoever, from this Lemma we can infer that $\forall t \leq \tau, \iota_t = 1$. So the cumulative loss estimate is close to the true cumulative loss and satisfies

$$\sup_{\mathbf{x}\in F_\tau}\Big|\sum_{s=1}^{\tau-1}(\hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) - q_s(\mathbf{x}) + q_s(\mathbf{x}_s))\Big| \leq \frac{1}{\eta_1}.$$

Step 2. Let $\mathbf{u}_{\tau+1}$ be the minimizer of $\sum_{s=1}^{\tau}\hat{f}_s(\mathbf{x})$ over F_τ. Suppose $B\left(\mathbf{u}_{\tau+1}, \frac{1}{T^2}\right)\cap\mathcal{X}_\xi \subset F_\tau$. Then

$$\mathrm{Vol}(F_\tau) \geq \mathrm{Vol}\left(B\left(\mathbf{u}_{\tau+1}, \frac{1}{T^2}\right)\cap\mathcal{X}_\xi\right).$$

Next, from our assumption that $\mathcal{X}$ contains a euclidean ball of radius 1, we can infer that $\mathcal{X}_\xi = \xi\mathbf{x}_1 + (1-\xi)\mathcal{X}$ contains a ball of radius $(1-\xi)$ in it. Let $\tilde{B}$ be the ball of radius $(1-\xi)$ that lies in $\mathcal{X}_\xi$. By convexity of $\mathcal{X}$ and the fact that the diameter of $\mathcal{X}$ is less than or equal to T, we have

$$\left(1-\frac{1}{T^3}\right)\mathbf{u}_{\tau+1} + \frac{1}{T^3}\tilde{B} \subseteq B\left(\mathbf{u}_{\tau+1}, \frac{1}{T^2}\right)\cap\mathcal{X}_\xi.$$

This shows that $\mathrm{Vol}(F_\tau) \geq T^{-4d}\omega_d$, where ω_d is the volume of unit sphere in $\mathbb{R}^d$. Combining this with the previous upper bound on $\mathrm{Vol}(F_\tau)$, we get

$$T^{-4d}\omega_d, \leq \mathrm{Vol}(F_\tau) \leq T^{-6d}\mathrm{Vol}(\mathcal{X}) \overset{(a)}{\leq} T^{-5d}\omega_d,$$

where (a) follows from the fact that the diameter of $\mathcal{X}$ is upper bounded by T. We arrived at a contradiction. This shows that $B\left(\mathbf{u}_{\tau+1}, \frac{1}{T^2}\right)\cap\mathcal{X}_\xi \not\subset F_\tau$.

Step 3. Since $B\left(\mathbf{u}_{\tau+1}, \frac{1}{T^2}\right) \cap \mathcal{X}_\xi \not\subset F_\tau$, the following holds: $\exists \mathbf{x} \in \partial F_\tau \cap \text{int}(\mathcal{X}_\xi)$ such that $\|\mathbf{x} - \mathbf{u}_{\tau+1}\|_2 \leq \frac{1}{T^2}$. Now, consider the following for such an $\mathbf{x}$

$$\begin{aligned}\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{u}_{\tau+1}) &= \sum_{s=1}^{\tau} q_s(\mathbf{x}) - q_s(\mathbf{u}_{\tau+1}) \\ &\quad + \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{u}_{\tau+1}) - q_s(\mathbf{x}) + q_s(\mathbf{u}_{\tau+1}).\end{aligned}$$

Since each q_s is T-Lipschitz, the first term in the RHS above is upper bounded by 1. Since the cumulative loss estimate is close to the true cumulative loss, the second term can be bounded as

$$\begin{aligned}\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{u}_{\tau+1}) - q_s(\mathbf{x}) + q_s(\mathbf{u}_{\tau+1}) &\leq \frac{2}{\eta_1} + \hat{f}_\tau(\mathbf{x}) - \hat{f}_\tau(\mathbf{u}_{\tau+1}) - q_\tau(\mathbf{x}) + q_\tau(\mathbf{u}_{\tau+1}) \\ &\overset{(a)}{\leq} \frac{2}{\eta_1} + 2B + \hat{f}_\tau(\mathbf{x}) - \hat{f}_\tau(\mathbf{u}_{\tau+1}),\end{aligned}$$

where (a) follows from the fact that q_s is a bounded function. $\hat{f}_\tau(\mathbf{x}) - \hat{f}_\tau(\mathbf{u}_{\tau+1})$ can be bounded as follows

$$\begin{aligned}\hat{f}_\tau(\mathbf{x}) - \hat{f}_\tau(\mathbf{u}_{\tau+1}) &= \frac{1}{2}(\mathbf{x} - \mathbf{x}_\tau)^T \hat{H}_\tau (\mathbf{x} - \mathbf{x}_\tau) + \langle \hat{g}_\tau, \mathbf{x} - \mathbf{x}_\tau \rangle \\ &\quad - \frac{1}{2}(\mathbf{u}_{\tau+1} - \mathbf{x}_\tau)^T \hat{H}_\tau (\mathbf{u}_{\tau+1} - \mathbf{x}_\tau) - \langle \hat{g}_\tau, \mathbf{u}_{\tau+1} - \mathbf{x}_\tau \rangle \\ &\leq \frac{1}{2}\lambda^{-2} d^2 (B+\epsilon)(\|\mathbf{x} - \mathbf{x}_\tau\|_{M_\tau}^2 + \|\mathbf{u}_{\tau+1} - \mathbf{x}_\tau\|_{M_\tau}^2) \\ &\quad + \lambda^{-1} d (B+\epsilon)(\|\mathbf{x} - \mathbf{x}_\tau\|_{M_\tau} + \|\mathbf{u}_{\tau+1} - \mathbf{x}_\tau\|_{M_\tau}).\end{aligned}$$

From Lemma 21, we know that for any $\mathbf{x} \in F_\tau$, $\|\mathbf{x} - \mathbf{x}_\tau\|_{M_\tau} \leq 4d\alpha$. Substituting this in the previous equation we get

$$\hat{f}_\tau(\mathbf{x}) - \hat{f}_\tau(\mathbf{u}_{\tau+1}) \leq 16(B+\epsilon)\left(\lambda^{-2} d^4 \alpha^2 + \lambda^{-1} d^2 \alpha\right) \leq \frac{1}{\eta_1}.$$

This shows that $\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}) - \hat{f}_s(\mathbf{u}_{\tau+1}) \leq \frac{4}{\eta_1}$. We now show that this implies the restart condition should have triggered. Consider the following

$$\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) &= \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{u}_{\tau+1}) \\
&\leq \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x}) \\
&= \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}\rangle - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) \\
&= \frac{4}{\eta_1} + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}_{s+1}\rangle + \langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \\
&\quad - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s) \\
&\overset{(a)}{\leq} \frac{4}{\eta_1} + 10\lambda^{-3}\alpha d^4 (B+\epsilon)^2 \sum_{s=1}^{\tau} \eta_s + \sum_{s=1}^{\tau} \langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \\
&\quad - \sum_{s=1}^{\tau} \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s),
\end{aligned}$$

where (a) follows from the stability of the iterates we proved in Lemma 21. Since $\mathbf{x}_{s+1}$ is the minimizer of $\min_{\mathbf{y}\in F_s} \eta_s \langle \hat{g}_s, \mathbf{y}\rangle + \Phi_{R_{s+1}}(\mathbf{y}, \mathbf{x}_s)$, we have the following from the first order optimality conditions

$$\langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \leq \frac{\Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_{s+1}) - \Phi_{R_{s+1}}(\mathbf{x}_{s+1}, \mathbf{x}_s)}{\eta_s}.$$

Using this in the previous display, we get

$$\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 10\lambda^{-3}\alpha d^4 (B+\epsilon)^2 \sum_{s=1}^{\tau} \eta_s + \sum_{s=1}^{\tau} \frac{\Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x}, \mathbf{x}_{s+1})}{\eta_s} \\
- \sum_{s=1}^{\tau} \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T \hat{H}_s(\mathbf{x}-\mathbf{x}_s).
\end{aligned}$$

Rearranging the terms in the RHS above, we get

$$\begin{aligned}
\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 10\lambda^{-3}\alpha d^4 (B+\epsilon)^2 \sum_{s=1}^{\tau} \eta_s + \frac{\Phi_R(\mathbf{x}, \mathbf{x}_1)}{\eta_1} - \frac{\Phi_{R_{\tau+1}}(\mathbf{x}, \mathbf{x}_{\tau+1})}{\eta_\tau} \\
+ \sum_{s=2}^{\tau} \left(\frac{1}{\eta_s} - \frac{1}{\eta_{s-1}}\right) \Phi_{R_s}(\mathbf{x}, \mathbf{x}_s).
\end{aligned}$$

Recall, $\mathbf{x} \in \partial F_\tau \cap \mathrm{int}(\mathcal{X}_\xi)$. Let τ' be such that $\mathbf{x} \in \partial B_{\alpha, M_{\tau'}}(\mathbf{x}_{\tau'})$. Then

$$\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 10\lambda^{-3}\alpha d^4 (B+\epsilon)^2 \sum_{s=1}^{\tau} \eta_s + \frac{\Phi_R(\mathbf{x}, \mathbf{x}_1)}{\eta_1} - \gamma \frac{\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'})}{\eta_{\tau'}}.$$

Since $M_{\tau'}, \tilde{M}_{\tau'}$ are spectrally close to each other and since $\|\mathbf{x} - \mathbf{x}_{\tau'}\|_{M_{\tau'}} = \alpha$, we have $\|\mathbf{x} - \mathbf{x}_{\tau'}\|_{\tilde{M}_{\tau'}} \geq \alpha/2$. Using this, we now lower bound $\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'})$

$$\begin{aligned}
\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'}) &= \Phi_R(\mathbf{x}, \mathbf{x}_{\tau'}) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_{\tau'})^T \left(\sum_{s=1}^{\tau'-1} \eta_s H_s\right)(\mathbf{x} - \mathbf{x}_{\tau'}) \\
&\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}_{\tau'})^T \left(M_{\tau'} - \tilde{M}_{\tau'}\right)(\mathbf{x} - \mathbf{x}_{\tau'}) \\
&\overset{(a)}{\geq} \frac{\alpha}{2} - \log\left(1 + \frac{\alpha}{2}\right) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_{\tau'})^T \left(M_{\tau'} - \tilde{M}_{\tau'}\right)(\mathbf{x} - \mathbf{x}_{\tau'}) \\
&\overset{(b)}{\geq} \frac{\alpha}{2} - \log\left(1 + \frac{\alpha}{2}\right) - \frac{\alpha}{20(1 + 8d\alpha)^2},
\end{aligned}$$

where (a) follows from property (P6) of SCB stated in Equation (22) and (b) follows from the fact that $M_{\tau'}, \tilde{M}_{\tau'}$ are spectrally close to each other. For our choice of α, $\Phi_{R_{\tau'}}(\mathbf{x}, \mathbf{x}_{\tau'})$ can be lower bounded by $\alpha/4$. We now upper bound $\Phi_R(\mathbf{x}, \mathbf{x}_1)$. Since $\mathbf{x} \in \mathcal{X}_\xi$, using property P8 of SCB stated in Appendix G, we can upper bound $\Phi_R(\mathbf{x}, \mathbf{x}_1)$ as

$$\Phi_R(\mathbf{x}, \mathbf{x}_1) = R(\mathbf{x}) \leq 4\nu \log T.$$

Substituting the above two bounds in the previous display and using the fact that $\eta_\tau \leq 10\eta_1$, we get

$$\sum_{s=1}^{\tau} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y} \in F_\tau} \sum_{s=1}^{\tau} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1} + 100\lambda^{-3}\alpha d^4 (B + \epsilon)^2 \eta_1 T + \frac{4\nu \log T}{\eta_1} - \frac{\alpha\gamma}{20\eta_1} \leq -\frac{\beta}{\eta_1}.$$

This implies, the restart condition should have triggered. This shows that the focus region doesn't get updated more than $12d \log T$ times.

■

E.1. Proof of Proposition 8

In this section, we first show that the cumulative Hessian estimates and cumulative loss function estimates generated by the modified algorithm concentrate well around their expected values. In particular, Lemma 23 is concerned about concentration of the Hessian estimates $\{\hat{H}_t\}_{t=1}^T$, and Lemma 25 is concerned about loss estimates $\{\hat{f}_t\}_{t=1}^T$ of the modifed algorithm. These two Lemmas immediately imply that $\iota_t = 1$ for any $t \leq \mathcal{T}$ w.h.p, where $\mathcal{T}$ is the minimum between T and the first time at which the modified algorithm restarts. Consequently, with high probability, the iterates of the modified and the original algorithms are exactly the same. These two Lemmas together prove Proposition 8.

Before we proceed, note that the focus region gets updated at most $12d \log T$ times before the algorithm restarts. So, for our choice of γ, we have $(1 + \gamma)^{12d \log T} \leq 10$. Consequently, for all $t \leq \mathcal{T}$, $\eta_t \leq 10\eta_1$. So the results of Lemma 21 apply to all the iterates in the first $\mathcal{T}$ iterations of the modified algorithm.

Lemma 23 (Concentration of Hessian estimates) *Let $\mathcal{T}$ be the minimum between T and the first time at which the modified algorithm restarts. Then for any $t \leq \mathcal{T}$, the following statement holds with probability at least $1 - T^{-2}$*

$$\|I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\|_2 = O\left(\alpha^2 \eta_1 \lambda^{-2} d^5 B \sqrt{T \log(dT)}\right).$$

Proof We first try to derive upper and lower bounds for $\tilde{M}_t$. From Lemma 21, we know that for all $s \leq \mathcal{T}$, and for all $\mathbf{x} \in F_s$, $\|\mathbf{x} - \mathbf{x}_s\|_{\tilde{M}_s} \leq 8d\alpha$. So, from Assumption 3 we have $\tilde{M}_t \succeq \frac{1}{(1+8d\alpha)^2} \tilde{M}_s$ for all $s \leq t$. This implies

$$\tilde{M}_t \succeq \frac{1}{(1+8d\alpha)^2} \tilde{M}_1 = \frac{1}{(1+8d\alpha)^2} \nabla^2 R(\mathbf{x}_1).$$

Moreover, from Lemma 19 we have $\tilde{M}_t \preceq T^8(\nu + 2\sqrt{\nu})^2(\nabla^2 R(\mathbf{x}_1) + I)$. Since $\nabla^2 R(\mathbf{x}_1)$ is a fixed quantity, for large enough T we have $\frac{1}{\text{poly}(T)} I \preceq \nabla^2 R(\mathbf{x}_1) \preceq \text{poly}(T) I$. This then shows that there exist positive constants c_l, c_u such that $T^{-c_l} I \preceq \tilde{M}_t \preceq T^{c_u} I$ for any $t \leq \mathcal{T}$.

Next consider the following

$$I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2} = \sum_{s=1}^{t-1} \eta_s \tilde{M}_t^{-1/2} \left(\iota_s H_s - \hat{H}_s\right) \tilde{M}_t^{-1/2}.$$

So we have

$$\begin{aligned}
&\|I - \tilde{M}_t^{-1/2} M_t \tilde{M}_t^{-1/2}\|_2 \\
&\leq \sup_{T^{-c_l} I \preceq A \preceq T^{c_u} I} \bar{\iota}_A \Big\| \sum_{s=1}^{t-1} \eta_s A^{-1/2} \left(\hat{H}_s - \iota_s H_s\right) A^{-1/2} \mathbb{I}\left(A \succeq \frac{1}{(1+8d\alpha)^2} \tilde{M}_s\right) \Big\|_2,
\end{aligned}$$

where $\bar{\iota}_A$ is an indicator random variable which is equal to 1 if and only if

$$\forall s \leq \mathcal{T}, \quad A \succeq \frac{1}{(1+8d\alpha)^2} \tilde{M}_s.$$

We now focus on bounding the RHS of the above equation. We write $\hat{H}_t$ as

$$\hat{H}_t = \hat{H}_{t,1} + \hat{H}_{t,2} = \frac{\lambda^{-2}}{2} d^2 \iota_t (\underbrace{r_t(\mathbf{y}_t)}_{\hat{H}_{t,1}} + \underbrace{q_t(\mathbf{y}_t)}_{\hat{H}_{t,2}}) M_t^{1/2} \left(\mathbf{v}_{1,t} \mathbf{v}_{2,t}^T + \mathbf{v}_{2,t} \mathbf{v}_{1,t}^T\right) M_t^{1/2}$$

Now consider the RHS in the second-to-last display

$$\begin{aligned}
&\bar{\iota}_A \Big\| \sum_{s=1}^{t-1} \eta_s A^{-1/2} \left(\hat{H}_{s,1} + \hat{H}_{s,2} - \iota_s H_s\right) A^{-1/2} \mathbb{I}\left(A \succeq \frac{1}{(1+8d\alpha)^2} \tilde{M}_s\right) \Big\|_2 \\
&\quad \leq \Big\| \sum_{s=1}^{t-1} \eta_s A^{-1/2} \hat{H}_{s,1} A^{-1/2} \mathbb{I}\left(A \succeq \frac{1}{(1+8d\alpha)^2} \tilde{M}_s\right) \Big\|_2 \\
&\quad\quad + \bar{\iota}_A \Big\| \sum_{s=1}^{t-1} \eta_s A^{-1/2} \left(\hat{H}_{s,2} - \iota_s H_s\right) A^{-1/2} \mathbb{I}\left(A \succeq \frac{1}{(1+8d\alpha)^2} \tilde{M}_s\right) \Big\|_2
\end{aligned}$$

First consider the first term in the RHS above. We have

$$\begin{aligned}
&\Big|\Big|\sum_{s=1}^{t-1}\eta_s A^{-1/2}\hat{H}_{s,1}A^{-1/2}\mathbb{I}\left(A\succeq\frac{1}{(1+8d\alpha)^2}\tilde{M}_s\right)\Big|\Big|_2\\
&\quad\leq\sum_{s=1}^{t-1}\Big|\Big|\eta_s A^{-1/2}\hat{H}_{s,1}A^{-1/2}\mathbb{I}\left(A\succeq\frac{1}{(1+8d\alpha)^2}\tilde{M}_s\right)\Big|\Big|_2
\end{aligned}$$

If $\iota_s=0$, then the s^{th} term in the RHs above is 0. On the other hand if $\iota_s=1$, then we know that $M_s,\tilde{M}_s$ are spectrally close to each other. In this case, the s^{th} term above is upper bounded by $20\epsilon\lambda^{-2}\eta_1 d^2(1+8d\alpha)^2$. This follows from the fact that $r_t(\mathbf{y}_t)$ is bounded by ϵ and $A\succeq\frac{1}{(1+8d\alpha)^2}\tilde{M}_s$. So the RHS above is upper bounded by $20\epsilon\lambda^{-2}\eta_1 d^2(1+8d\alpha)^2T$. Now consider the second term

$$\begin{aligned}
&\bar{\iota}_A\Big|\Big|\sum_{s=1}^{t-1}\eta_s A^{-1/2}\left(\hat{H}_{s,2}-\iota_s H_s\right)A^{-1/2}\mathbb{I}\left(A\succeq\frac{1}{(1+8d\alpha)^2}\tilde{M}_s\right)\Big|\Big|_2\\
&\quad=\bar{\iota}_A\Big|\Big|\sum_{s=1}^{t-1}\eta_s A^{-1/2}\left(\hat{H}_{s,2}-\iota_s H_s\right)A^{-1/2}\phi\left((1+8d\alpha)^2\lambda_{min}(\tilde{M}_s^{-1/2}A\tilde{M}_s^{-1/2})\right)\Big|\Big|_2,
\end{aligned}$$

where $\phi:\mathbb{R}_+\to\mathbb{R}_+$ is defined as

$$\phi(x)=\begin{cases}1 & \text{if } x\geq 1\\ 2x-1 & \text{if } 1>x>1/2\\ 0 & \text{if } \frac{1}{2}\geq x\geq 0\end{cases}.$$

Continuing, we get

$$\begin{aligned}
&\bar{\iota}_A\Big|\Big|\sum_{s=1}^{t-1}\eta_s A^{-1/2}\left(\hat{H}_{s,2}-\iota_s H_s\right)A^{-1/2}\mathbb{I}\left(A\succeq\frac{1}{(1+8d\alpha)^2}\tilde{M}_s\right)\Big|\Big|_2\\
&\quad\leq\Big|\Big|\sum_{s=1}^{t-1}\eta_s A^{-1/2}\left(\hat{H}_{s,2}-\iota_s H_s\right)A^{-1/2}\phi\left((1+8d\alpha)^2\lambda_{min}(\tilde{M}_s^{-1/2}A\tilde{M}_s^{-1/2})\right)\Big|\Big|_2.
\end{aligned}$$

So we have

$$\begin{aligned}
&\|I-\tilde{M}_t^{-1/2}M_t\tilde{M}_t^{-1/2}\|_2 &(13)\\
&\quad\leq\sup_{T^{-c_l}I\preceq A\preceq T^{c_u}I}\Big|\Big|\sum_{s=1}^{t-1}\eta_s A^{-1/2}\left(\hat{H}_{s,2}-\iota_s H_s\right)A^{-1/2}\phi\left((1+8d\alpha)^2\lambda_{min}(\tilde{M}_s^{-1/2}A\tilde{M}_s^{-1/2})\right)\Big|\Big|_2 &(14)\\
&\qquad+20\epsilon\lambda^{-2}\eta_1 d^2(1+8d\alpha)^2T. &(15)
\end{aligned}$$

We now bound the first term in the RHS above using standard concentration results for matrix-valued martingales (see Lemma 16). Define random variable $Z_{A,s}$ as follows

$$Z_{A,s}=\begin{cases}\eta_s A^{-1/2}\left(\hat{H}_{s,2}-\iota_s H_s\right)A^{-1/2}\phi\left((1+8d\alpha)^2\lambda_{min}(\tilde{M}_s^{-1/2}A\tilde{M}_s^{-1/2})\right), & \text{if } s\leq\mathcal{T},\\ 0 & \text{if } \mathcal{T}<s\leq T\end{cases}.$$

Note that $\{Z_{A,s}\}_{s=1}^{T}$ is a matrix-valued martingale difference sequence and satisfies $\mathbb{E}_t\left[Z_{A,t}\right] = 0$. Moreover, $Z_{A,s}$ is a bounded random variable which satisfies $\|Z_{A,s}\|_2 = O\left(\eta_1\lambda^{-2}d^2(1+8\alpha d)^2 B\right)$. This is easy to see when $\iota_s = 0$. When $\iota_s = 1$, it follows from the facts that $M_s, \tilde{M}_s$ are spectrally close to each other and $A \succeq \frac{1}{(1+8d\alpha)^2}\tilde{M}_s$ and $q_s(\mathbf{x})$ is bounded by B. By relying on standard concentration results for matrix martingale sequences, we get with probability at least $1-\delta$

$$\forall t \leq T, \quad \|\sum_{s=1}^{t} Z_{A,s}\|_2 \leq O\left(\alpha^2\eta_1\lambda^{-2}d^4 B\sqrt{T\log(2T/\delta)}\right). \tag{16}$$

We now do a union bound over all A such that $T^{-c_l}I \preceq A \preceq T^{c_u}I$. We first construct an Δ-net so that the following holds: for every A, there exists a A_Δ in the Δ-net such that $(1+(Td)^{-1})A_\Delta \succeq A \succeq (1-(Td)^{-1})A_\Delta$. We can show that the size of such an Δ-net is $\tilde{O}\left((Td)^{cd^2}\right)$, for some positive constant c. Moreover, we can show that for every A, there exists an A_Δ in the Δ-net such that

$$\|\sum_{s=1}^{t-1} Z_{A,s} - Z_{A_\Delta,s}\|_2 \leq \tilde{O}\left(\alpha^2\eta_1\lambda^{-2}d^4 B\sqrt{T}\right).$$

This follows from the fact that ϕ is bounded and Lipschitz. Now consider the following

$$\begin{aligned}\sup_{T^{-c_l}I \preceq A \preceq T^{c_u}I}\|\sum_{s=0}^{t-1} Z_{A,s}\|_2 &= \sup_A \|\sum_{s=0}^{t-1} Z_{A_\Delta,s}\|_2 + \sup_A \|\sum_{s=0}^{t-1} Z_{A_\Delta,s} - Z_{A,s}\|_2\\ &\leq \sup_{A_\Delta \text{ in } \Delta\text{-net}} \|\sum_{s=0}^{t-1} Z_{A_\Delta,s}\|_2 + O\left(\alpha^2\eta_1\lambda^{-2}Bd^4\sqrt{T}\right),\end{aligned}$$

where A_Δ is the point in Δ-net which is closest to A. Finally, by relying on the bound in Equation (16) and performing a union bound over all the elements in the Δ-net gives us $\sup_A \|\sum_{s=0}^{t-1} Z_{A,s}\|_2 = \tilde{O}\left(\alpha^2\eta_1\lambda^{-2}d^5 B\sqrt{T}\right)$. Plugging this bound in Equation (13) and using the fact that $\epsilon = O\left(dBT^{-1/2}\right)$ gives us the required result. ■

Remark 24 (Convexifying the restart condition) *We note that a similar argument as above can be used to show that the following two matrices are spectrally close to each other*

$$N_t = \nabla^2 R(\mathbf{x}_t) + \eta_1(d\alpha)^2\sum_{s=1}^{t-1}\hat{H}_s, \quad \tilde{N}_t = \nabla^2 R(\mathbf{x}_t) + \eta_1(d\alpha)^2\sum_{s=1}^{t-1} H_s.$$

In particular, we can show that $\|I - \tilde{N}_t^{-1/2}N_t\tilde{N}_t^{-1/2}\|_2 \leq \frac{1}{2}$. This would entail that N_t is invertible and positive definite. This in turn implies that the following objective is convex

$$\min_{\mathbf{y}\in F_t}\sum_{s=0}^{t}\hat{f}_s(\mathbf{y}) + (d^2\alpha^2\eta_1)^{-1}(\mathbf{y}-\mathbf{x}_t)^T\nabla^2 R(\mathbf{x}_t)(\mathbf{y}-\mathbf{x}_t).$$

Now consider the restart condition stated in line 16 of Algorithm 1. It involves solving $\min_{\mathbf{y}\in F_t}\sum_{s=0}^{t}\hat{f}_s(\mathbf{y})$. Note that this objective itself may not be convex. However, it is point-wise close to the above objective, which is convex. To see this, note that in Lemma 21 we showed that $\forall \mathbf{x}\in F_t$, $\|\mathbf{x}-\mathbf{x}_t\|_{\tilde{M}_t}\le 8d\alpha$. As a result, $\forall \mathbf{x}\in F_t$, $(d^2\alpha^2\eta_1)^{-1}(\mathbf{y}-\mathbf{x}_t)^T\nabla^2 R(\mathbf{x}_t)(\mathbf{y}-\mathbf{x}_t) = O\left(\eta_1^{-1}\right)$. Consequently, the two objectives are $O\left(\frac{1}{\eta_1}\right)$ close to each other. So, one can efficiently check for an "approximate" restart condition by minimizing the above convex objective.

Lemma 25 (Concentration of loss estimates) *Let $\mathcal{T}$ be the minimum between T and the first time at which the modified algorithm restarts. Then for any $t\le\mathcal{T}$, the following statement holds with probability at least $1-T^{-2}$*

$$\sup_{\mathbf{x}\in F_t}\Big|\sum_{s=1}^{t-1}\eta_1(\hat{f}_s(\mathbf{x})-\hat{f}_s(\mathbf{x}_s)-\iota_s q_s(\mathbf{x})+\iota_s q_s(\mathbf{x}_s))\Big| \le \tilde{O}\left(\alpha^2\eta_1\lambda^{-2}Bd^{9/2}\sqrt{T}\right).$$

Proof First note that

$$\hat{f}_s(\mathbf{x})-\hat{f}_s(\mathbf{x}_s)=\frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s)+\langle\hat{g}_s,\mathbf{x}-\mathbf{x}_s\rangle\,.$$

We split $\hat{H}_s,\hat{g}_s$ into two components, one corresponding to r_s and the other corresponding to q_s

$$\begin{aligned}\hat{H}_t &= \frac{\lambda^{-2}}{2}d^2\iota_t(\underbrace{r_t(\mathbf{y}_t)}_{\hat{H}_{t,1}}+\underbrace{q_t(\mathbf{y}_t)}_{\hat{H}_{t,2}})M_t^{1/2}\left(\mathbf{v}_{1,t}\mathbf{v}_{2,t}^T+\mathbf{v}_{2,t}\mathbf{v}_{1,t}^T\right)M_t^{1/2}\\ \hat{g}_t &= \lambda^{-1}d\iota_t(\underbrace{q_t(\mathbf{y}_t)}_{\hat{g}_{t,2}}+\underbrace{r_t(\mathbf{y}_t)}_{\hat{g}_{t,1}})M_t^{1/2}\mathbf{v}_{1,t}.\end{aligned}$$

Similarly, we define $\hat{r}_s(\mathbf{x})$ and $\hat{q}_s(\mathbf{x})$ as follows. These are obtained by splitting $\hat{f}_s(\mathbf{x})$ into two components based on r_s and q_s

$$\begin{aligned}\hat{r}_s(\mathbf{x})-\hat{r}_s(\mathbf{x}_s) &= \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_{s,1}(\mathbf{x}-\mathbf{x}_s)+\langle\hat{g}_{s,1},\mathbf{x}-\mathbf{x}_s\rangle\\ \hat{q}_s(\mathbf{x})-\hat{q}_s(\mathbf{x}_s) &= \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_{s,2}(\mathbf{x}-\mathbf{x}_s)+\langle\hat{g}_{s,2},\mathbf{x}-\mathbf{x}_s\rangle\,.\end{aligned}$$

We first upper bound $|\sum_{s=1}^{t-1}\hat{r}_s(\mathbf{x})-\hat{r}_s(\mathbf{x}_s)|$. First note that from Lemma 21 we know that for any $\mathbf{x}\in F_t$, $\|\mathbf{x}-\mathbf{x}_t\|_{M_t}\le 4d\alpha$. Using this, we have the following for any $\mathbf{x}\in F_t$.

$$\begin{aligned}|\sum_{s=1}^{t-1}\hat{r}_s(\mathbf{x})-\hat{r}_s(\mathbf{x}_s)| &\le \sum_{s=1}^{t-1}|\hat{r}_s(\mathbf{x})-\hat{r}_s(\mathbf{x}_s)| && (17)\\ &\le 16\epsilon T(\lambda^{-2}d^4\alpha^2+\lambda^{-1}d^2\alpha) && (18)\\ &\le 32\alpha^2\epsilon\lambda^{-2}d^4T. && (19)\end{aligned}$$

Next, we upper bound $|\sum_{s=1}^{t-1} \hat{q}_s(\mathbf{x}) - \hat{q}_s(\mathbf{x}_s) - \iota_s q_s(\mathbf{x}) + \iota_s q_s(\mathbf{x}_s)|$. Define random variables $Z_{\mathbf{x},s}$ as

$$Z_{\mathbf{x},s} = \begin{cases} \eta_1(\hat{q}_s(\mathbf{x}) - \hat{q}_s(\mathbf{x}_s) - \iota_s q_s(\mathbf{x}) + \iota_s q_s(\mathbf{x}_s)) & \text{if } s \leq \mathcal{T} \\ 0 \quad \text{otherwise} \end{cases}.$$

It is easy to see that $\{Z_{\mathbf{x},s}\}_{s=1}^{T}$ is a martingale difference sequence. Moreover, $Z_{\mathbf{x},s}$ is a bounded random variable which satisfies

$$|Z_{\mathbf{x},s}| \leq 32\alpha^2\eta_1\lambda^{-2}Bd^4.$$

This again follows from the fact that $\mathbf{x} \in F_t$, $\|\mathbf{x} - \mathbf{x}_t\|_{M_t} \leq 4d\alpha$ which we proved in Lemma 21. By relying on standard concentration bounds for martingale difference sequences (see Lemma 15), we get that with probability at least $1-\delta$,

$$\sup_{t\leq T} |\sum_{s=1}^{t-1} Z_{\mathbf{x},s}| = O\left(\lambda^{-2}d^4\alpha^2B\eta_1\sqrt{T\log T/\delta}\right).$$

Next, we bound $\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=1}^{t-1} Z_{\mathbf{x},s}|$ using Δ-net arguments. Let $\mathcal{N}_\Delta$ be an Δ-net over F_t which satisfies the following: for every $\mathbf{x}$, there exists a $\mathbf{x}_\Delta \in \mathcal{N}_\Delta$ such that $\|\mathbf{x} - \mathbf{x}_\Delta\|_{M_t} \leq \Delta$. Then

$$\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=1}^{t-1} Z_{\mathbf{x},s}| \leq \underbrace{\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=0}^{t-1} Z_{\mathbf{x}_\Delta,s}|}_{T_1} + \underbrace{\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=0}^{t-1} Z_{\mathbf{x}_\Delta,s} - Z_{\mathbf{x},s}|}_{T_2}. \tag{20}$$

Using a simple union bound, T_1 can be bounded as

$$T_1 \leq O\left(\lambda^{-2}d^4\alpha^2B\eta_1\sqrt{T\log T|\mathcal{N}_\Delta|/\delta}\right) \stackrel{(a)}{\leq} O\left(\lambda^{-2}d^{9/2}\alpha^2B\eta_1\sqrt{T\log\frac{\alpha dT}{\Delta\delta}}\right),$$

where the bound holds with probability at least $1-\delta$ and (a) holds since $\forall \mathbf{x} \in F_t, \|\mathbf{x} - \mathbf{x}_t\|_{M_t} \leq 4d\alpha$ and as a result $|\mathcal{N}_\Delta| \leq \left(\frac{4d\alpha}{\Delta}\right)^d$. T_2 can be bounded as follows

$$\begin{aligned}
&\sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=0}^{t-1} Z_{\mathbf{x}_\Delta,s} - Z_{\mathbf{x},s}| \\
&\stackrel{(a)}{\leq} \sup_{\mathbf{x}\in F_t}\sup_{t\leq T}|\sum_{s=0}^{t-1}\eta_1\left\langle \hat{g}_{s,2} - \iota_s\nabla q_s(\mathbf{x}_s), \mathbf{x} - \mathbf{x}_\Delta\right\rangle| \\
&\quad + \sup_{\mathbf{x}\in F_t}\sup_{t\leq T}\left|\sum_{s=0}^{t-1}\eta_1\left\langle \hat{H}_{s,2} - \iota_s H_s, (\mathbf{x}-\mathbf{x}_s)(\mathbf{x}-\mathbf{x}_s)^T - (\mathbf{x}_\Delta - \mathbf{x}_s)(\mathbf{x}_\Delta - \mathbf{x}_s)^T\right\rangle_F\right|
\end{aligned}$$

where (a) follows from the definitions of $Z_{\mathbf{x},s}$ and $q_s(\mathbf{x}), \hat{q}_s(\mathbf{x})$ and $\langle \cdot, \cdot \rangle_F$ is the frobenius inner product. The first term in the RHS above can be bounded as

$$\begin{aligned}
&\sup_{\mathbf{x}\in F_t} \sup_{t\le T} \left| \sum_{s=0}^{t-1} \eta_1 \left\langle \hat{g}_{s,2} - \iota_s \nabla q_s(\mathbf{x}_s), \mathbf{x} - \mathbf{x}_\Delta \right\rangle \right| \\
&\overset{(a)}{\le} 2\eta_1 \lambda^{-1} dB \sup_{\mathbf{x}\in F_t} \sup_{t\le T} \left(\sum_{s=0}^{t-1} \|\mathbf{x} - \mathbf{x}_\Delta\|_{M_s} \right) \\
&\overset{(b)}{\le} 2(1+8d\alpha)^2 \eta_1 \lambda^{-1} dB \sup_{\mathbf{x}\in F_t} \sup_{t\le T} \left(\sum_{s=0}^{t-1} \|\mathbf{x} - \mathbf{x}_\Delta\|_{M_t} \right) = O\left(\lambda^{-1} d^3 \alpha^2 B \eta_1 \Delta T\right),
\end{aligned}$$

where (a) follows from the facts that $\|\hat{g}_{s,2}\|_{M_s}^* \le \lambda^{-1} dB$, $\mathbb{E}_s\left[\hat{g}_{s,2}\right] = \iota_s \nabla q_s(\mathbf{x}_s)$, and (b) follows from Lemma 21 where we showed that $M_s \preceq (1+8d\alpha)^2 M_t$.

Using similar arguments and the fact that $\forall \mathbf{x} \in F_t, \|\mathbf{x} - \mathbf{x}_t\|_{M_t} \le 4d\alpha$, the second term in the RHS of the second-to-last display can be bounded as

$$\begin{aligned}
&\sup_{\mathbf{x}\in F_t} \sup_{t\le T} \left| \sum_{s=0}^{t-1} \eta_1 \left\langle \hat{H}_{s,2} - \iota_s H_s, (\mathbf{x} - \mathbf{x}_s)(\mathbf{x} - \mathbf{x}_s)^T - (\mathbf{x}_\Delta - \mathbf{x}_s)(\mathbf{x}_\Delta - \mathbf{x}_s)^T \right\rangle_F \right| \\
&= O\left(\lambda^{-2} d^5 \alpha^3 B \eta_1 \Delta T\right).
\end{aligned}$$

Choosing $\Delta = \frac{1}{\alpha\sqrt{dT}}$, and plugging the above bounds for T_1, T_2 in Equation (20) gives us $\sup_{\mathbf{x}\in F_t} \sup_{t\le T} \left| \sum_{s=1}^{t-1} Z_{\mathbf{x},s} \right| = \tilde{O}\left(\alpha^2 \eta_1 \lambda^{-2} B d^{9/2} \sqrt{T}\right)$. Finally, combining Equation (17) and Equation (20), and using the fact that $\epsilon = O\left(dBT^{-1/2}\right)$ gives us the requires result. ∎

Remark 26 *For our choice of hyper-parameters, the concentration bounds in Lemmas 23, 25 show that the indicator random variables* $\{\iota_t\}_{t=1}^{\mathcal{T}}$ *are equal to* 1 *with high probability. This entails that the iterates produced by the modified algorithm are exactly equal to the iterates produced by the actual algorithm with high probability. As a result all the properties we showed for the modified algorithm in Lemmas 21, 22, 23, 25 also hold for the original algorithm with high probability.*

E.2. Main argument for Theorem 7

We are now ready to prove Theorem 7. Since we know that with high probability, the iterates of the modified algorithm which relies on indicator variables ι_t are exactly same as the original algorithm, it suffices to prove the regret bound for the modified algorithm. In the sequel, we work with the modified algorithm. Throughout the proof, we let $\mathcal{T}$ be the minimum between T and the first time step at which the algorithm restarts. Let τ be the minimum between $\mathcal{T}$ and the last time step where $\iota_\tau = 1$. Our goal is to bound the following quantity

$$\sum_{s=1}^{\mathcal{T}} \iota_s f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{s=1}^{\mathcal{T}} \iota_s f_s(\mathbf{x}) = \sum_{s=1}^{\tau} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{s=1}^{\tau} f_s(\mathbf{x}).$$

Case 1 ($\mathcal{T} = T$). We first consider the case where the restart condition didn't trigger in the first T iterations (i.e., $\mathcal{T} = T$). In this case, we show that the regret is $\tilde{O}\left(T^{1/2}\right)$. Since the restart condition hasn't triggered, we know that

$$\sum_{s=1}^{\mathcal{T}} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_{\mathcal{T}}} \sum_{s=1}^{\mathcal{T}} \hat{f}_s(\mathbf{y}) \geq -\frac{\beta}{\eta_1}.$$

From the proof of Lemma 22, this implies $\forall \mathbf{x} \in \partial F_{\mathcal{T}} \cap \mathrm{int}(\mathcal{X}_\xi)$

$$\sum_{s=1}^{\mathcal{T}} \hat{f}_s(\mathbf{x}) - \min_{\mathbf{y}\in F_{\mathcal{T}}} \sum_{s=1}^{\mathcal{T}} \hat{f}_s(\mathbf{y}) \geq \frac{4}{\eta_1}.$$

(In Lemma 22, we proved a contrapositive statement. We showed that if $\exists \mathbf{x} \in \partial F_T \cap \mathrm{int}(\mathcal{X}_\xi)$ such that $\sum_{s=0}^{T} \hat{f}_s(\mathbf{x}) - \min_{\mathbf{y}\in F_T} \sum_{s=0}^{T} \hat{f}_s(\mathbf{y}) \leq \frac{4}{\eta_1}$, then $\sum_{s=0}^{T} \hat{f}_s(\mathbf{x}_s) - \min_{\mathbf{y}\in F_T} \sum_{s=0}^{T} \hat{f}_s(\mathbf{y}) \leq -\frac{\beta}{\eta_1}$). Since our cumulative loss estimate concentrates well around the true cumulative loss (i.e., $\iota_{\mathcal{T}} = 1$), this implies

$$\forall \mathbf{x} \in \partial F_{\mathcal{T}} \cap \mathrm{int}(\mathcal{X}_\xi), \quad \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{x}) - \min_{\mathbf{y}\in F_{\mathcal{T}}} \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{y}) \geq \frac{2}{\eta_1}.$$

Since q_s's are convex, this implies the minimizer of $\min_{\mathbf{x}\in\mathcal{X}_\xi} \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{x})$ is in $F_{\mathcal{T}}$. So, the regret of the algorithm can be bounded as follows

$$\begin{aligned}
\mathrm{Reg}_T = \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{x}) &\leq \epsilon T + \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{x}) \\
&\stackrel{(a)}{\leq} 1 + \epsilon T + \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}_\xi} \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{x}) \\
&= 1 + \epsilon T + \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{y}_s) - \min_{\mathbf{x}\in F_{\mathcal{T}}} \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{x}),
\end{aligned}$$

where (a) follows from the definition of $\mathcal{X}_\xi = (1-\xi)\mathcal{X} + \xi \mathbf{x}_1$ and the fact that the loss functions are Lipschitz and the diameter of $\mathcal{X}$ is bounded. Next, consider the following for any $\mathbf{x} \in F_{\mathcal{T}}$

$$\begin{aligned}
\sum_{s=1}^{\mathcal{T}} q_s(\mathbf{y}_s) - \sum_{s=1}^{\mathcal{T}} q_s(\mathbf{x}) = \underbrace{\sum_{s=1}^{\mathcal{T}} \left[q_s(\mathbf{y}_s) - q_s(\mathbf{x}_s)\right]}_{T_1} + \underbrace{\sum_{s=1}^{\mathcal{T}} \left[q_s(\mathbf{x}_s) - q_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) + \hat{f}_s(\mathbf{x})\right]}_{T_2} \\
+ \underbrace{\sum_{s=1}^{\mathcal{T}} \left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right]}_{T_3}.
\end{aligned}$$

Bounding T_1. Consider the following

$$\sum_{s=0}^{T} q_s(\mathbf{y}_s) - q_s(\mathbf{x}_s) \leq \sum_{s=0}^{T} \lambda \left\langle \nabla q_s(\mathbf{x}_s), M_s^{-1/2}(\mathbf{v}_{1,s} + \mathbf{v}_{2,s}) \right\rangle$$
$$+ \lambda^2 \frac{1}{2} (\mathbf{v}_{1,s} + \mathbf{v}_{2,s})^T M_s^{-1/2} H_s M_s^{-1/2} (\mathbf{v}_{1,s} + \mathbf{v}_{2,s}).$$

Let $Z_s = \lambda \left\langle \nabla f_s(\mathbf{x}_s), M_s^{-1/2}(\mathbf{v}_{1,s} + \mathbf{v}_{2,s}) \right\rangle$ if $s \leq \tau$ and 0 if $s > \tau$. Note that $\{Z_s\}_{s=1}^T$ is a martingale difference sequence with each Z_s being bounded: $|Z_s| \leq 2dB$. This follows from the observation that $\nabla q_s(\mathbf{x}_s) = \mathbb{E}_s\left[\hat{g}_s\right]$ and the fact that $M_s^{-1/2}\hat{g}_s$ is a bounded random variable. By relying on standard concentration bounds for martingale difference sequences (see Lemma 15), we get that with probability at least $1-\delta$, $\sum_{s=1}^T Z_s = O\left(dB\sqrt{T \log 1/\delta}\right)$. Next, consider the last term in the RHS

$$\begin{aligned}
&(\mathbf{v}_{1,s} + \mathbf{v}_{2,s})^T M_s^{-1/2} H_s M_s^{-1/2} (\mathbf{v}_{1,s} + \mathbf{v}_{2,s}) \leq 4\|M_s^{-1/2} H_s M_s^{-1/2}\|_2 \\
&\quad \leq 4\|\tilde{M}_{s+1}^{-1/2} H_s \tilde{M}_{s+1}^{-1/2}\|_2 \|M_s^{-1/2} \tilde{M}_{s+1} M_s^{-1/2}\|_2 \\
&\quad \leq 4\|\tilde{M}_{s+1}^{-1/2} H_s \tilde{M}_{s+1}^{-1/2}\|_2 \|M_s^{-1/2} \tilde{M}_s M_s^{-1/2}\|_2 \|\tilde{M}_s^{-1/2} \tilde{M}_{s+1} \tilde{M}_s^{-1/2}\|_2
\end{aligned}$$

Since $\tilde{M}_s, M_s, M_{s+1}$ are spectrally close to each other, we can show that $\|M_s^{-1/2}\tilde{M}_s M_s^{-1/2}\|_2$, $\|\tilde{M}_s^{-1/2}\tilde{M}_{s+1}\tilde{M}_s^{-1/2}\|_2$ are close to 1. So we have

$$(\mathbf{v}_{1,s} + \mathbf{v}_{2,s})^T M_s^{-1/2} H_s M_s^{-1/2} (\mathbf{v}_{1,s} + \mathbf{v}_{2,s}) \leq 8\|\tilde{M}_{s+1}^{-1/2} H_s \tilde{M}_{s+1}^{-1/2}\|_2.$$

Using similar arguments as in the proof of Theorem 9 (see Equation (6)), we get the following upper bound for T_1: $O\left(dB\sqrt{T\log 1/\delta} + \frac{d \log dT}{\eta_1}\right)$.

Bounding T_2. Since $\iota_\tau = 1$, T_2 can be upper bounded as

$$\begin{aligned}
T_2 &\leq \frac{1}{\eta_1} + \left[q_\tau(\mathbf{x}_\tau) - q_\tau(\mathbf{x}) - \hat{f}_\tau(\mathbf{x}_\tau) + \hat{f}_\tau(\mathbf{x})\right] \\
&\leq \frac{1}{\eta_1} + + \langle \hat{g}_\tau - \nabla q_\tau(\mathbf{x}_\tau), \mathbf{x} - \mathbf{x}_\tau \rangle + \frac{1}{2}\left\langle \hat{H}_\tau - H_\tau, (\mathbf{x} - \mathbf{x}_\tau)(\mathbf{x} - \mathbf{x}_\tau)^T \right\rangle_F \\
&\leq \frac{2}{\eta_1},
\end{aligned}$$

where the last inequality follows from the facts that $\|\mathbf{x} - \mathbf{x}_\tau\|_{M_\tau} \leq 4d\alpha$, $\|\hat{g}_\tau\|_{M_\tau}^* \leq \lambda^{-1}d(B+\epsilon)$, $\|M_\tau^{-1/2}\hat{H}_\tau M_\tau^{-1/2}\|_2 \leq \lambda^{-2}d^2(B+\epsilon)$..

Bounding T_3. To bound T_3, we consider the following

$$\begin{aligned}
\sum_{s=0}^{T}\left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right] &= \sum_{s=1}^{T}\langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}\rangle - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s)\\
&= \sum_{s=1}^{T}\langle \hat{g}_s, \mathbf{x}_s - \mathbf{x}_{s+1}\rangle + \langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle - \frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s)\\
&\overset{(a)}{\leq} 10\lambda^{-3}\alpha d^4(B+\epsilon)^2\sum_{s=1}^{T}\eta_s + \sum_{s=1}^{T}\langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle\\
&\quad - \sum_{s=1}^{T}\frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s),
\end{aligned}$$

where (a) follows from the stability of the iterates we proved in Lemma 21. Since $\mathbf{x}_{s+1}$ is the minimizer of $\min_{\mathbf{y}\in F_s}\eta_s\langle \hat{g}_s, \mathbf{y}\rangle + \Phi_{R_{s+1}}(\mathbf{y}, \mathbf{x}_s)$, we have

$$\langle \hat{g}_s, \mathbf{x}_{s+1} - \mathbf{x}\rangle \leq \frac{\Phi_{R_{s+1}}(\mathbf{x},\mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x},\mathbf{x}_{s+1}) - \Phi_{R_{s+1}}(\mathbf{x}_{s+1},\mathbf{x}_s)}{\eta_s}.$$

Using this in the previous display, we get

$$\begin{aligned}
\sum_{s=0}^{T}\left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right] &\leq 10\lambda^{-3}\alpha d^4(B+\epsilon)^2\sum_{s=1}^{T}\eta_s + \sum_{s=1}^{T}\frac{\Phi_{R_{s+1}}(\mathbf{x},\mathbf{x}_s) - \Phi_{R_{s+1}}(\mathbf{x},\mathbf{x}_{s+1})}{\eta_s}\\
&\quad - \sum_{s=1}^{T}\frac{1}{2}(\mathbf{x}-\mathbf{x}_s)^T\hat{H}_s(\mathbf{x}-\mathbf{x}_s).
\end{aligned}$$

Rearranging the terms in the RHS above, we get

$$\begin{aligned}
\sum_{s=0}^{T}\left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right] &\leq 10\lambda^{-3}\alpha d^4(B+\epsilon)^2\sum_{s=1}^{T}\eta_s + \frac{\Phi_R(\mathbf{x},\mathbf{x}_1)}{\eta_1} - \frac{\Phi_{R_{T+1}}(\mathbf{x},\mathbf{x}_{T+1})}{\eta_T}\\
&\quad + \sum_{s=2}^{T}\left(\frac{1}{\eta_s} - \frac{1}{\eta_{s-1}}\right)\Phi_{R_s}(\mathbf{x},\mathbf{x}_s)\\
&\overset{(a)}{\leq} 10\lambda^{-3}\alpha d^4(B+\epsilon)^2\sum_{s=1}^{T}\eta_s + \frac{\Phi_R(\mathbf{x},\mathbf{x}_1)}{\eta_1},
\end{aligned}$$

where (a) follows from the facts that R_s is convex, and $\eta_s \geq \eta_{s-1}$ for all s. Hence the last two terms are negatives and can be ignored. Since $\mathbf{x} \in \mathcal{X}_\xi$, using property P8 of SCB stated in Appendix G, we can upper bound $\Phi_R(\mathbf{x},\mathbf{x}_1)$ as

$$\Phi_R(\mathbf{x},\mathbf{x}_1) = R(\mathbf{x}) \leq 4\nu\log T.$$

Combining the bounds for T_1, T_2, T_3 shows that with probability at least $1 - T^{-2}$ the regret is upper bounded by

$$\tilde{O}\left(\epsilon T + dB\sqrt{T} + \frac{(\nu+d)}{\eta_1} + \lambda^{-3}\alpha d^4(B+\epsilon)^2\eta_1 T\right) = \tilde{O}\left(d^{11}(d+\nu)^5\sqrt{T}\right).$$

Case 2 ($\mathcal{T} < T$). We now consider the case where the restart condition triggered at some iteration $\mathcal{T} < T$. Using the fact that the restart condition hasn't triggered in iteration $\mathcal{T} - 1$ and using similar arguments as in the beginning of Case 1, we can again show that the minimizer of the cumulative loss over the entire domain lies in the focus region $F_{\mathcal{T}}$, and $\iota_{\mathcal{T}} = 1$. So regret until $\mathcal{T}$ is given by

$$\begin{aligned}\mathrm{Reg}_{\mathcal{T}} = \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{x}) &\overset{(a)}{\leq} 1 + \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in\mathcal{X}_\xi} \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{x}) \\ &= 1 + \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{y}_s) - \min_{\mathbf{x}\in F_{\mathcal{T}}} \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{x}),\end{aligned}$$

where (a) follows from the definition of $\mathcal{X}_\xi$. Using the same regret decomposition as in Case 1, for any $\mathbf{x} \in F_{\mathcal{T}}$

$$\begin{aligned}\sum_{s=1}^{\mathcal{T}} f_s(\mathbf{y}_s) - \sum_{s=1}^{\mathcal{T}} f_s(\mathbf{x}) \leq \epsilon T + \underbrace{\sum_{s=1}^{\mathcal{T}} [q_s(\mathbf{y}_s) - q_s(\mathbf{x}_s)]}_{T_1} + \underbrace{\sum_{s=1}^{\mathcal{T}} \left[q_s(\mathbf{x}_s) - q_s(\mathbf{x}) - \hat{f}_s(\mathbf{x}_s) + \hat{f}_s(\mathbf{x})\right]}_{T_2} \\ + \underbrace{\sum_{s=1}^{\mathcal{T}} \left[\hat{f}_s(\mathbf{x}_s) - \hat{f}_s(\mathbf{x})\right]}_{T_3}.\end{aligned}$$

We use the same arguments as in Case 1 to bound T_1, T_2 as

$$T_1 = O\left(dB\sqrt{T\log 1/\delta} + \frac{d\log dT}{\eta_1}\right), \quad T_2 = \frac{2}{\eta_1}.$$

Since the restart condition triggered in round $\mathcal{T}$, T_3 is bounded by $-\frac{\beta}{\eta_1}$. Combining all these bounds, we get the following bound on regret

$$\mathrm{Reg}_{\mathcal{T}} \leq \epsilon T + O\left(dB\sqrt{T\log 1/\delta} + \frac{d\log dT}{\eta_1}\right) + \frac{2}{\eta_1} - \frac{\beta}{\eta_1}.$$

For our choice of hyper-parameters, the above bound is less than 0.

Appendix F. Additional Results

Proposition 27 (Gaussian Smoothing) *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a potentially non-smooth function. Define the smoothed function $\hat{f}$ as $\hat{f}(\mathbf{x}) = \mathbb{E}_{\mathbf{u}\sim\mathcal{N}(0,I)}\left[f(\mathbf{x} + C\mathbf{u})\right]$, for some symmetric positive definite matrix C. Then $\hat{f}$ is twice differentiable with the following gradient and Hessian*

$$\nabla\hat{f}(\mathbf{x}) = \mathbb{E}_{\mathbf{u}\sim\mathcal{N}(0,I)}\left[C^{-1}\mathbf{u}f(\mathbf{x} + C\mathbf{u})\right], \ \nabla^2\hat{f}(\mathbf{x}) = \mathbb{E}_{\mathbf{u}\sim\mathcal{N}(0,I)}\left[C^{-1}(\mathbf{u}\mathbf{u}^T - I)C^{-1}f(\mathbf{x} + C\mathbf{u})\right].$$

Proof

Gradient. Using the expression for probability density function of a multivariate Gaussian, we get

$$
\begin{aligned}
\nabla \hat{f}(\mathbf{x}) &= \frac{\partial}{\partial \mathbf{x}} \int \frac{1}{(2\pi)^{d/2}} f(\mathbf{x}+C\mathbf{u}) e^{-\|\mathbf{u}\|^2/2} d\mathbf{u} \overset{(a)}{=} \frac{\partial}{\partial \mathbf{x}} \int \frac{1}{(2\pi|C|^2)^{d/2}} f(\mathbf{y}) e^{-\|\mathbf{y}-\mathbf{x}\|^2_{C^{-2}}/2} d\mathbf{y} \\
&= \int \frac{\partial}{\partial \mathbf{x}} \frac{1}{(2\pi|C|^2)^{d/2}} f(\mathbf{y}) e^{-\|\mathbf{y}-\mathbf{x}\|^2_{C^{-2}}/2} d\mathbf{y} = \int \frac{C^{-2}(\mathbf{y}-\mathbf{x})}{(2\pi|C|^2)^{d/2}} f(\mathbf{y}) e^{-\|\mathbf{y}-\mathbf{x}\|^2_{C^{-2}}/2} d\mathbf{y} \\
&\overset{(b)}{=} \int \frac{C^{-1}\mathbf{u}}{(2\pi)^{d/2}} f(\mathbf{x}+C\mathbf{u}) e^{-\|\mathbf{u}\|^2/2} d\mathbf{u},
\end{aligned}
$$

where we used change of variables in (a) and (b). This shows that

$$
\nabla \hat{f}(\mathbf{x}) = \mathbb{E}_{\mathbf{u}\sim\mathcal{N}(0,I)} \left[C^{-1}\mathbf{u} f(\mathbf{x}+C\mathbf{u}) \right].
$$

Hessian. We use a similar argument as above to compute the Hessian. From the first display above, we have

$$
\nabla \hat{f}(\mathbf{x}) = \int \frac{C^{-2}(\mathbf{y}-\mathbf{x})}{(2\pi|C|^2)^{d/2}} f(\mathbf{y}) e^{-\|\mathbf{y}-\mathbf{x}\|^2_{C^{-2}}/2} d\mathbf{y}.
$$

Using the definition of Hessian, we get

$$
\begin{aligned}
\nabla^2 \hat{f}(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} \nabla \hat{f}(\mathbf{x}) &= \frac{\partial}{\partial \mathbf{x}} \int \frac{C^{-2}(\mathbf{y}-\mathbf{x})}{(2\pi|C|^2)^{d/2}} f(\mathbf{y}) e^{-\|\mathbf{y}-\mathbf{x}\|^2_{C^{-2}}/2} d\mathbf{y} \\
&= \int \frac{\partial}{\partial \mathbf{x}} \frac{C^{-2}(\mathbf{y}-\mathbf{x})}{(2\pi|C|^2)^{d/2}} f(\mathbf{y}) e^{-\|\mathbf{y}-\mathbf{x}\|^2_{C^{-2}}/2} d\mathbf{y} \\
&= \int \frac{C^{-2}(\mathbf{y}-\mathbf{x})(\mathbf{y}-\mathbf{x})^T C^{-2} - C^{-2}}{(2\pi|C|^2)^{d/2}} f(\mathbf{y}) e^{-\|\mathbf{y}-\mathbf{x}\|^2_{C^{-2}}/2} d\mathbf{y} \\
&\overset{(a)}{=} \int \frac{C^{-1}\mathbf{u}\mathbf{u}^T C^{-1} - C^{-2}}{(2\pi|C|^2)^{d/2}} f(\mathbf{x}+C\mathbf{u}) e^{-\|u\|^2/2} d\mathbf{u}
\end{aligned}
$$

where we used change of variables in (a). This shows that

$$
\mathbb{E}_{\mathbf{u}\sim\mathcal{N}(0,I)} \left[C^{-1}(\mathbf{u}\mathbf{u}^T - I) C^{-1} f(\mathbf{x}+C\mathbf{u}) \right].
$$

Appendix G. Review of Self Concordant Barriers

This section reviews some useful properties of Self Concordant (SC) functions and Self Concordant Barriers (SCBs). Most of the content in this section is from Nemirovski (2004); Nesterov (2018).

- **(P3)** *Non-degeneracy*: If $\mathcal{X}$ doesn't contain straight lines, then the Hessian $\nabla^2 R(\mathbf{x})$ is nondegenerate (*i.e.,* $\nabla^2 R(\mathbf{x}) \succ 0$) at all points $\mathbf{x} \in \mathrm{int}(\mathcal{X})$.
- **(P4)** For any $\mathbf{x} \in \mathrm{int}(\mathcal{X})$, we have

$$
\mathcal{X} \cap \{\mathbf{y} : \langle \nabla R(\mathbf{x}), \mathbf{y}-\mathbf{x} \rangle \geq 0\} \subseteq B_{\nu+2\sqrt{\nu}, \nabla^2 R(\mathbf{x})}(\mathbf{x}). \tag{21}
$$

- **(P5)** *Semiboundedness*: For any $\mathbf{x} \in \text{int}(\mathcal{X}), \mathbf{y} \in \mathcal{X}$, $\langle \nabla R(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq \nu$.
- **(P6)** For any $\mathbf{x}, \mathbf{y} \in \text{int}(\mathcal{X})$,

$$R(\mathbf{y}) - R(\mathbf{x}) - \langle \nabla R(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq \|\mathbf{y} - \mathbf{x}\|_{\nabla^2 R(\mathbf{x})} - \log(1 + \|\mathbf{y} - \mathbf{x}\|_{\nabla^2 R(\mathbf{x})}). \tag{22}$$

- **(P7)** For any $\mathbf{x}, \mathbf{y} \in \text{int}(\mathcal{X})$, we have

$$\langle \nabla R(\mathbf{y}) - \nabla R(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq \frac{\|\mathbf{y} - \mathbf{x}\|^2_{\nabla^2 R(\mathbf{x})}}{1 + \|\mathbf{y} - \mathbf{x}\|_{\nabla^2 R(\mathbf{x})}}. \tag{23}$$

- **(P8)** Define the Minkowsky function of $\mathcal{X}$ with the pole at $\mathbf{x}$ as

$$\pi_{\mathbf{x}}(\mathbf{y}) = \inf\{t > 0 | \mathbf{x} + t^{-1}(\mathbf{y} - \mathbf{x}) \in \mathcal{X}\}.$$

  Then for any $\mathbf{x}, \mathbf{y} \in \text{int}(\mathcal{X})$

$$R(\mathbf{y}) \leq R(\mathbf{x}) + \nu \log \frac{1}{1 - \pi_{\mathbf{x}}(\mathbf{y})} \tag{24}$$

$$\nabla^2 R(\mathbf{y}) \preceq \left(\frac{\nu + 2\sqrt{\nu}}{1 - \pi_{\mathbf{x}}(\mathbf{y})} \right)^2 \nabla^2 R(\mathbf{x}). \tag{25}$$

■