Dynamic Regret in Outlier-Oblivious Online Optimization using Nonconvex Robust Losses

Adarsh Barik[1] **Anand Krishna**[2] **Vincent Y. F. Tan**[3]

[1]Department of Computer Science and Engineering, Indian Institute of Technology Delhi, India
[2]Poiro, Evam Labs, Bengaluru, India
[3]Department of Mathematics, Department of Electrical and Computer Engineering, National University of Singapore

Abstract

We study a robust online convex optimization framework, where an adversary can introduce outliers by corrupting loss functions in an arbitrary number of rounds k, unknown to the learner. In contrast to prior works, we consider both bounded and unbounded domains and allow for large gradients for the losses without relying on a Lipschitz assumption or any prior knowledge of k. We introduce the Log Exponential Adjusted Robust and iNvex (LEARN) loss, a non-convex (invex) robust loss function to mitigate the effects of outliers and develop a robust variant of the online gradient descent algorithm by leveraging the LEARN loss. We establish dynamic regret guarantees with respect to the uncorrupted rounds and conduct experiments to validate our theory. Our upper bound matches the existing lower bound in the bounded domains and is tight (up to logarithmic factors) in the unbounded domains. Furthermore, we present a unified analysis framework for developing online optimization algorithms for non-convex (invex) losses, utilizing it to provide regret bounds with respect to the LEARN loss, which may be of independent interest.

1 INTRODUCTION

In the mercurial era of digital information, the proliferation of online data streams has become ubiquitous, with applications spanning from financial markets and healthcare to e-commerce and social media [D. Hendricks and Wilcox, 2016, Zhang et al., 2015, Fedoryszak et al., 2019, Kejariwal et al., 2015]. The surge in real-time data generation across these diverse fields has underscored the pressing need for adaptive optimization strategies. These strategies need to navigate dynamic and uncertain environments, especially as organizations rely more on online decision-making processes. In this context, the demand for efficient and robust optimization techniques in the presence of outliers has become more pronounced.

Outlier Robustness: Pioneering works in robust statistics [Huber, 1964, Beaton and Tukey, 1974] emphasized mitigating outliers' impact on algorithm performance. In online optimization originally designed for uncorrupted streams, outliers pose challenges leading to suboptimal solutions [Feng et al., 2017]. This motivates developing outlier-robust online methodologies, an area of recent focus [Diakonikolas et al., 2022, van Erven et al., 2021, Bhaskara and Ruwanpathirana, 2020, Zhang and Cutkosky, 2025].

Online Convex Optimization (OCO): The OCO framework [Hazan, 2023, Orabona, 2023] offers a useful mathematical framework for studying sequential predictions, especially in the presence of adversarial interventions. The adversary chooses a deterministic sequence of convex losses f_t for each $t \in \{1, 2, \ldots, T\}$, where T denotes the time horizon. In every round, the learner picks an action θ_t in a convex set Θ for which she suffers the loss $f_t(\theta_t)$ for the round t and additionally gets to observe f_t. The learner's goal is to minimize the cumulative loss, i.e., $\sum_{t=1}^{T} f_t(\theta_t)$, and her performance is evaluated using the notion of regret.

OCO with Robustness: The OCO framework serves as a valuable starting point for studying algorithms that dynamically adapt to incoming data points while considering potential outliers' influence. Drawing inspiration from the model proposed by van Erven et al. [2021], this work considers the robust OCO framework, albeit in a dynamic setting, where the adversary is allowed to choose any $\mathcal{S} \subseteq [T]$ of rounds to be 'clean' and inject outliers into the remaining rounds. This adversarial influence is exerted by tampering with the side information s_t integral to constructing the loss function f_t. Considering the example of the online ridge regression problem, where $s_t = (x_t, y_t)$ represents the incoming data point for round t, and the loss is defined as $f_t(s_t, \theta_t) = \frac{1}{2}(\langle \theta_t, x_t \rangle - y_t)^2$, the adversary can corrupt either x_t, y_t or both to corrupt the round t. The corrupted

Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence (UAI 2026), PMLR 337:328–363.

round is considered an outlier. We use c_t to denote the clean (uncorrupted) side information and o_t for the oultlier (corrupted) side information. The learner's performance is measured using regret with respect to only the clean rounds, and we refer to this as clean regret. In this paper, we explore the notion of dynamic regret – a performance measure suitable for dynamic environments characterized by evolving data and shifting optimal decision parameters. Dynamic regret entails adapting strategies to compete effectively with evolving comparators, thereby mitigating the impact of changing environments on decision-making processes. We adopt the notion of dynamic regret from Besbes et al. [2015], Yang [2016], Mokhtari et al. [2016] where the performance of the online learner is evaluated against a clairvoyant adversary who has complete foresight of the loss functions for all rounds and selects the optimal action at each round accordingly. Let $\boldsymbol{\theta} \coloneqq (\theta_t)_{t\in[T]}$ be a sequence of learner's actions and $(\theta_t^*)_{t\in[T]}$ be a sequence of benchmark points such that $\theta_t^* = \arg\min_{\theta\in\Theta} f_t(c_t, \theta)$. Then, the clean dynamic regret is defined to be

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \coloneqq \sum_{t\in\mathcal{S}} \left(f_t(s_t, \theta_t) - f_t(s_t, \theta_t^*)\right) .$$

Our goal is to develop robust online algorithms to achieve optimal regret bounds with respect to T and the number of outliers k, which is unknown a priori.

Problem Setting: Diverging from van Erven et al. [2021]'s static setting, which assumes the knowledge of the number of outliers k, we introduce novel OCO techniques for outlier scenarios in a dynamic setting when the number of outliers k is unknown. In contrast to common bounded domain and Lipschitz loss assumptions [Hazan, 2023, Orabona, 2023], our approach allows unbounded domains and non-Lipschitz losses. Following Jacobsen and Cutkosky [2023], we permit large gradients away from f_t's minimizer, extending applicability to outlier scenarios. Our work focuses on m-strongly convex losses. Many widely used regularization techniques in machine learning, such as ℓ_2-regularization, naturally induce strong convexity in the loss function. This is particularly valuable in high-dimensional settings, where regularization enhances solution stability and mitigates overfitting. They have also been considered in an OCO setup by Besbes et al. [2015], Mokhtari et al. [2016]. Importantly, this assumption does not trivialize the problem. In the static setting, strongly convex losses facilitate a logarithmic regret-bound relative to the time horizon T. However, in the dynamic setting, Zhang et al. [2017] showed that linear regret is unavoidable in the worst-case scenario, even for the strongly convex functions. Following their construction, we extend their result to show that the dynamic regret is lower bounded by $\Omega\left(\sqrt{V_T T}\right)$ where optimal path variation V_T is defined as $V_T = \sum_{t=1}^T \|\theta_t^* - \theta_{t+1}^*\|$ (See Theorem B.1).

Main Contributions: In this paper, we study a crucial question within the robust OCO framework:

Can we achieve optimal clean dynamic regret with respect to time horizon T and the outlier count k, without its prior knowledge, when the adversary can corrupt an arbitrary subset of rounds?

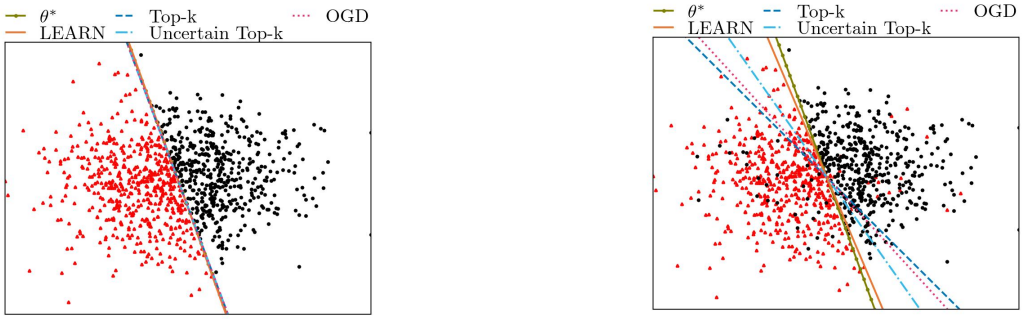
We answer the aforementioned question in the affirmative with a linear dependence on the number of outliers, which is known to be unavoidable in the robust OCO setting [van Erven et al., 2021]. Our main contributions in this work are summarized below:

(1) **Introduction of robust invex loss:** At the core of our work lies the introduction of a robust, invex loss function, which we refer to as the LEARN loss (see Equation (3)). Invex functions are a special class of nonconvex functions where every stationary point is a global minimum. The limitations of convex losses in handling outliers have been well documented in the literature [Chen et al., 2022]. As an alternative solution, researchers have investigated non-convex robust losses [Barron, 2019]. However, given the nonconvexity, this transition poses challenges in the analysis. Striking a judicious balance, our LEARN loss addresses the outlier-handling requirements and lends itself to rigorous theoretical analysis. Subsequently, we introduce LEARN-OGD, a robust variant of online gradient descent leveraging the outlier-handling LEARN loss (Algorithm 1).

(2) **Optimal clean dynamic regret in bounded domain:** We study the clean regret for strongly convex loss functions within the robust OCO framework (Section 3.1). Our algorithm can handle potentially unbounded gradients while remaining agnostic to the number of outliers. Our upper bound exhibits optimal growth rate, specifically as $\mathcal{O}(\sqrt{(1+V_T)T} + k)$. Note that a linear dependence on k is unavoidable even for the static case as shown in van Erven et al. [2021].

(3) **An expert framework with outliers and novel clean dynamic regret in unbounded domain:** Extending dynamic regret bounds to unbounded domains is nontrivial. The existing methods, such as Jacobsen and Cutkosky [2023] utilize experts to tackle this problem. Introducing outliers poses new challenges, hindering a straightforward extension to our setup. We address this by constructing an expert framework with N experts that can handle outliers, incurring $\mathcal{O}(\sqrt{T\log N} + k)$ expert regret (Lemma 3.2). The construction and analysis of this expert framework could be of independent interest. Utilizing it, we propose a parameter-free algorithm EXP-LEARN-OGD (Algorithm 2) and derive clean dynamic regret bounds as $\mathcal{O}(\sqrt{(1+V_T)T} + V_T + \sqrt{T\log T} + k)$. This matches existing bounds without outliers in terms of T and V_T dependence.

(4) **Online optimization with invex losses:** As a complementary contribution, we provide a unifying analytical



(a) $k = 0$ (b) $k = T^{\frac{2}{3}}$

Figure 1: We compared LEARN-OGD (Algorithm 1) and baselines (Section 5) on binary classification with hinge loss and sequential data. Figure 1a shows the true decision boundary with no outliers and all the methods recover it exactly. Figure 1b shows how decision boundaries for baseline methods deviate from the ground truth in the presence of outliers while LEARN-OGD remains robust to outliers.

Table 1: Best known upper and lower bounds on the dynamic regret of strongly convex loss functions f_t in the presence of outliers (with respect to time-horizon T and number of outliers k).

Setting	Best-known upper bound	Our upper bound	Lower bound w/o corruption
Bounded domain	$\mathcal{O}(\sqrt{(1+V_T)T}+k)$ [van Erven et al., 2021] + Appendix C	$\mathcal{O}(\sqrt{(1+V_T)T}+k)$	$\Omega(\sqrt{V_T T})$ Zhang et al. [2017] + Theorem B.1
Unbounded domain	Not Available	$\mathcal{O}\left(\sqrt{(1+V_T)T}+V_T +\sqrt{T\log T}+k\right)$	$\Omega(\sqrt{V_T T}+V_T)$ Jacobsen and Cutkosky [2023] + Theorem B.3

framework for analyzing the regret of LEARN-OGD (Algorithm 1), with invex losses (Section 4). This framework extends to geodesically convex [Wang et al., 2023] and weakly pseudo-convex [Gao et al., 2018] losses, enhancing the applicability of our approach.

Table 1 presents a concise summary of our results and contrasts them with the best existing results. In the bounded domain, the best-known upper bound comes from a straightforward extension of van Erven et al. [2021]'s results to the dynamic setting (See Appendix C). However, unlike van Erven et al. [2021]'s approach, our algorithm does not require prior knowledge of k. In the unbounded domain, to the best of our knowledge, our work presents the first-known results in the dynamic setting. We further validate our theoretical results through comprehensive numerical experiments conducted in Section 5. Figure 1 shows results from our online SVM experiments (Appendix J.2) where LEARN-OGD consistently predicts decision boundaries closely matching ground truth despite outliers (details are provided in Appendix J.2).

Related Works: Our starting point is the robust OCO framework with outliers proposed by van Erven et al. [2021], though their analysis is focused on static regret with known outlier counts. The foundational works of Huber [1964] and Beaton and Tukey [1974] motivate the importance of handling outliers from a robust statistics perspective. Incorporating robust loss functions like Huber, Tukey's biweight, etc., has been explored by Barron [2019], Belagiannis et al. [2015] to enhance algorithm resilience against outliers. We build upon the expert framework approach of Jacobsen and Cutkosky [2023] for unconstrained domains, extending it to handle outliers. Our robust approach aligns with the broader literature on robust optimization in theoretical computer science and machine learning problems, as comprehensively surveyed by Diakonikolas and Kane [2023]. A detailed section on the related works appears in the Appendix A.

2 PRELIMINARIES AND NOTATIONS

We consider an online learning protocol delineated by the subsequent framework. At each iteration $t \in [T] := \{1, 2, \ldots, T\}$, the learner picks an action $\theta_t \in \Theta$ where $\Theta \subseteq \mathbb{R}^d$ is a convex set. Subsequently, the adversary reveals an m-strongly convex, non-negative loss function $f_t(\theta) := f_t(s_t, \theta)$ where s_t denotes the side information for round t, resulting in the learner incurring a loss of $f_t(s_t, \theta_t)$.

The side information s_t is also revealed along with the function f_t to the learner in every round t. For fixed s_t, $\nabla f_t(s_t, \theta_t)$ denotes the gradient of $f_t(s_t, \theta)$ with respect to θ at $\theta = \theta_t$. While we assume throughout that the gradient $\nabla f_t(s_t, \theta_t)$ exists, our analysis still holds when f_t is not smooth. We allow for the existence of potentially large gradients by following the quadratically bounded gradient assumption of Jacobsen and Cutkosky [2023] with unknown non-negative constants G and L, i.e.,

$$\|\nabla f_t(s_t, \theta_t)\| \leq G + L\|\theta_t - \omega_t^*\| \,, \tag{1}$$

where

$$\omega_t^* \coloneqq \arg\min_{\theta \in \Theta} f_t(s_t, \theta)$$

acts as the reference point. Inequality (1) can be further relaxed to admit unbounded gradients in corrupted rounds, especially within a bounded domain, by devising a filter mechanism, provided the learner has access to k or problem-specific parameters such as G, L and m; see Appendix L for details. The round t is said to be an *outlier round* if s_t is corrupted and the corruption can be arbitrary. We use $\mathcal{S} \subseteq [T]$ to denote the set of clean rounds. The loss incurred by the learner is assumed to be smaller than B in the clean rounds, i.e., $\max_{t \in \mathcal{S}} f_t(s_t, \theta_t) \leq B$, while it may be unbounded in outlier rounds. Next, we provide a formal definition of *invex functions*, a generalization of convex functions:

Definition 2.1 (Invex functions)**.** A differentiable function $g : \mathbb{R}^d \to \mathbb{R}$ is called a *ζ-invex function* on set Θ if for any $\vartheta_1, \vartheta_2 \in \Theta$ and a vector-valued function $\zeta : \Theta \times \Theta \to \mathbb{R}^d$,

$$g(\vartheta_2) \geq g(\vartheta_1) + \langle \nabla g(\vartheta_1), \zeta(\vartheta_2, \vartheta_1) \rangle \tag{2}$$

where $\nabla g(\vartheta_1)$ is the gradient of g at ϑ_1. When $\zeta(\vartheta_1, \vartheta_2) = \vartheta_2 - \vartheta_1$, the inequality in (2) reduces to the characterization of convex functions for differentiable functions.

LEARN loss: We introduce a non-convex robust loss, Log Exponential Adjusted Robust and iNvex (LEARN) loss $g : \Theta \to \mathbb{R}$ that transforms the output of a convex function $f : \Theta \to \mathbb{R}$ as follows

$$g(\theta) = -a \log\left(\exp\left(-\frac{1}{a} f(\theta) \right) + b \right) , \tag{3}$$

where $a, b > 0$ are constants that can be tuned. Robust losses have been widely used in various real-world problems to reduce the sensitivity of algorithms to outliers [Barron, 2019]. The structure of LEARN loss shares similarities with the log-sum-exp function, used previously for robustness [Lozano et al., 2016]. To understand how LEARN loss works to mitigate its susceptibility to outliers, we examine its behavior by comparing how g resembles the square loss $f(r) = r^2$ where r denotes the residual, and the minimum is attained at 0. The LEARN loss g closely follows the behavior of f in the vicinity of the origin and gradually levels off as we move away from it (Figure 3 in Appendix D). Importantly, the minimizers for both g and f are the same. Observe that for LEARN loss, the norm of the gradient monotonically decreases with the distance from the minimum. This reduces the adverse influence of outliers during a gradient descent-style update. This property, known as the *redescending property* [Hampel et al., 2011] in the M-estimation literature, enhances the robustness of the model. It is worth noting that empirical studies have shown that non-convex losses offer better robustness when compared to the convex alternatives [Maronna et al., 2006].

In the following sections, we present our results and discuss their implications, focusing on key ideas and intuition. Detailed and formal proofs are provided in the appendix.

3 ROBUST ONLINE CONVEX OPTIMIZATION

In this section, we develop LEARN-OGD, a robust variant of the online gradient descent (OGD) algorithm aimed at minimizing regret within the robust (OCO) framework. Given a convex set $\Theta \subseteq \mathbb{R}^d$, the algorithm starts by picking an arbitrary action $\theta_1 \in \Theta$. For each round $t \in [T]$, the learner is revealed a convex loss function f_t in conjunction with the side-information s_t. Note that the adversary is capable of corrupting the side information for any k out of the T rounds and the learner does not see whether s_t is corrupted. To mitigate the challenge presented by the corruption introduced by the adversary, our algorithm leverages LEARN loss. The algorithm proceeds to construct g_t based on the original loss function f_t (Step 5) and executes the subsequent action through a projected gradient descent update on g_t.

To streamline the presentation of the results in this paper, we introduce the following notations:

$$\begin{aligned}
\psi &\coloneqq G + \max\left\{ \frac{mL}{2ab}, \frac{4a^2 L}{m^2 b} \right\}, \\
\phi &\coloneqq \frac{1}{b} \max\left\{ \frac{m}{2a}, \frac{4a^2}{m^2} \right\}, \\
\kappa &\coloneqq \frac{1}{b} \max\left\{ \frac{m}{2a}, \frac{2a}{m} \right\}, \\
\nu &\coloneqq \frac{1}{b} \max\left\{ a, \frac{1}{a} \right\}, \\
\xi &\coloneqq 1 + b \exp\left(\frac{B}{a} \right) .
\end{aligned}$$

The learner has the flexibility to fine-tune these constants by carefully selecting parameters a and b. For corrupted rounds, we introduce an additional quantity

$$\delta_{\mathcal{S}} \coloneqq \max_{t \in [T] \setminus \mathcal{S}} \|\omega_t^* - \theta_t^*\|$$

that measures the adversary's capability to influence the optimal action of round t, denoted by θ_t^*. Note that for uncorrupted rounds, $\omega_t^* = \theta_t^*$.

3.1 CLEAN DYNAMIC REGRET ANALYSIS

In this section, we study how LEARN-OGD performs within the robust OCO framework when evaluated using the clean regret.

Algorithm 1 LEARN-OGD

Initialize: $\theta_1 \in \Theta$
for $t = 1, \dots, T$ **do**
  Play: θ_t
  Observe: Side information s_t and strongly convex loss function $f_t(s_t, \theta)$
  Construct: For $a, b > 0$, let $g_t(\theta) \coloneqq g_t(s_t, \theta) = -a \log(\exp(-\frac{1}{a} f_t(s_t, \theta)) + b)$
  Update: Set $\theta_{t+1} \leftarrow \Pi_\Theta(\theta_t - \alpha_t \nabla g_t(\theta_t))$, where $\Pi_\Theta(\cdot)$ is the projection onto the convex set Θ
end for

3.1.1 Bounded domain

In the following theorem, we present upper bounds on the clean dynamic regret for m-strongly convex functions $\{f_t\}_{t=1}^T$ in a bounded domain. We take $\|\theta\| \leq D$ for all $\theta \in \Theta$. The learner determines the actions to be taken by following Algorithm 1. We provide an $\mathcal{O}(\sqrt{(1+V_T)T}+k)$ regret bound tightly matching the lower bound presented in Theorem B.1 with respect to V_T and T.

Theorem 3.1. *For each $t \in [T]$, consider a sequence of m-strongly convex loss functions $\{f_t(s_t, \cdot)\}_{t=1}^T$. Losses during uncorrupted rounds are bounded in $[0, B]$, while losses during corrupted rounds may be unbounded. Let $\boldsymbol{\theta}$ be the sequence of actions chosen by the learner using Algorithm 1. If the learner chooses $\alpha_t = \alpha = \sqrt{\frac{4D^2+6DV_T}{\psi^2 T}}$, then*

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \xi \Bigg(\psi \sqrt{(4D^2 + 6DV_T)T} + k(G\phi + L\kappa) + k(G + L\phi)\delta_{\mathcal{S}} \Bigg). \tag{4}$$

Let us briefly discuss key features of the above result.

- Optimal growth rates: The clean dynamic regret in Eq. (4) has optimal dependence on the total number of rounds T and the number of outlier rounds k (see van Erven et al. [2021] and Theorem B.1).
- Sensitivity to adversary strength: The regret bound depends on $\delta_{\mathcal{S}}$, which quantifies how strongly the adversary can perturb the optimal action during corrupted rounds. This term highlights the regret's sensitivity to adversarial influence in the presence of outliers.
- Tunable constants: Constants in the regret bound that do not scale with T can be improved via careful choice of parameters a and b, offering a practical avenue for further optimization of learner performance.

LEARN-OGD with the recommended step-size from Theorem 3.1 already encapsulates the key insights from our approach for handling corrupted rounds. However, the choice of α_t in Theorem 3.1 requires knowledge of V_T (and of G and L). This dependence can be removed using an expert-style approach. We give an adaptive version of the algorithm in Algorithm 2 and analyze it in Section 3.1.2; here we present the non-adaptive version to emphasize core ideas. The expert framework – deploying multiple experts with different step sizes and tracking the best performer – is a standard method for achieving adaptivity in online learning [Zhang et al., 2018]. Extending the expert framework to handle corrupted rounds is nontrivial; we address this extension and bound the resulting expert-regret in Lemma 3.2.

3.1.2 Unbounded domain via an expert framework

The scenario becomes notably more intricate when dealing with an unbounded domain, introducing an additional layer of complexity exacerbated by the presence of corrupted rounds. Recent work by Jacobsen and Cutkosky [2023] has addressed the challenges associated with unbounded domains through the development of an expert framework. In alignment with this approach, we propose an expert framework tailored to handle outlier rounds. Our framework encompasses N experts, each executing an instance of Algorithm 1 with a fixed stepsize $\alpha^\tau > 0$ and operating within a bounded domain $D^\tau > 0$ where τ refers to a given expert. The learner strategically selects their action by carefully constructing a weighted average of the actions taken by the individual experts. We present the experts' version of the LEARN-OGD, named EXP-LEARN-OGD, in Algorithm 2.

Algorithm 2 EXP-LEARN-OGD

Initialize: For all $\tau \in [N]$, select $\theta_1^\tau \in \Theta^\tau$, where $\Theta^\tau \coloneqq \{\theta \mid \|\theta\|_2 \leq D^\tau\} \cap \Theta$ and set $\rho_1^\tau = 1$
Define: $Z_t \coloneqq \sum_{\tau=1}^N \rho_t^\tau$
for $t = 1, \dots, T$ **do**
  Play: $\theta_t \leftarrow \frac{1}{Z_t} \sum_{\tau=1}^N \rho_t^\tau \theta_t^\tau$
  Observe: Side information s_t and m-strongly convex loss function f_t
  for $\tau = 1, 2, \dots, N$ **do**
    Query: $f_t(s_t, \theta_t^\tau)$
    Construct: For $a, b > 0$, set $g_t(\theta) \coloneqq g_t(s_t, \theta) = -a \log(\exp(-\frac{1}{a} f_t(s_t, \theta)) + b)$
    Update: $\theta_{t+1}^\tau \leftarrow \Pi_{\Theta^\tau}(\theta_t^\tau - \alpha^\tau \nabla g_t(\theta_t^\tau))$
  end for
  Define: $\widetilde{\eta}_t \coloneqq \min_{\tau \in [N]} \eta_t(s_t, \theta_t^\tau)$
  Update: For all $\tau \in [N]$ where $\beta > 0$, update $\rho_{t+1}^\tau \leftarrow \rho_t^\tau \exp(-\beta \widetilde{\eta}_t f_t(s_t, \theta_t^\tau))$
end for

Next, we provide a proof sketch for bounding the clean dynamic regret for Algorithm 2. Consider parameters $A_{\max} \geq C\sqrt{T}$ for sufficiently large $C > 0$. An expert τ selects parameters (α^τ and D^τ) from a set $\mathcal{E}$, defined as:

$$\mathcal{E} \coloneqq \left\{(\alpha, D) : \alpha \in \mathcal{E}_1, D \in \mathcal{E}_2\right\}, \tag{5}$$

where $\mathcal{E}_1 \coloneqq \left\{\alpha_i = \min\left\{\frac{2^i}{\sqrt{T}}, \frac{A_{\max}}{\sqrt{T}}\right\}\right\}_{i\in\mathbb{N}}$ and $\mathcal{E}_2 \coloneqq \left\{D_j = \min\left\{\frac{\epsilon 2^j}{T}, \frac{\epsilon 2^T}{T}\right\}\right\}_{j\in\mathbb{N}}$ for some $\epsilon > 0$. Clearly, $N \coloneqq |\mathcal{E}| \leq T\log_2 A_{\max}$. Let $\boldsymbol{\theta}^\tau \coloneqq (\theta_1^\tau, \cdots, \theta_T^\tau)$ be a sequence of actions chosen by the expert τ. Observe that for any expert τ,

$$\begin{aligned}\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) = &\underbrace{\sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^\tau) - \sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^*)}_{\text{Meta Regret } \mathfrak{R}_{\mathrm{M}}^T(\boldsymbol{\theta}^\tau, \mathcal{S})} \\ &+ \underbrace{\sum_{t\in\mathcal{S}} f_t(s_t, \theta_t) - \sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^\tau)}_{\text{Expert Regret } \mathfrak{R}_{\mathrm{E}}^T(\boldsymbol{\theta}, \boldsymbol{\theta}^\tau, \mathcal{S})}\end{aligned}$$

Taking $\mathfrak{D} \coloneqq \max_{t\in[T]} \|\theta_t^*\|$, we bound $\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S})$ in four major steps. First, if $\mathfrak{D} \leq D^\tau$ for any specific expert τ, then we establish that $\mathfrak{R}_{\mathrm{M}}^T(\boldsymbol{\theta}^\tau, \mathcal{S})$ is bounded (Lemma H.2). Second, we demonstrate that $\mathfrak{R}_{\mathrm{E}}^T(\boldsymbol{\theta}, \boldsymbol{\theta}^\tau, \mathcal{S})$ remains bounded when $\mathfrak{D} \leq D^\tau$. To that end, we prove the following bound on expert regret under the presence of outliers.

Lemma 3.2. *By choosing $\beta = \sqrt{\frac{8\log N}{T\nu^2}}$, the following regret guarantee holds with respect to any expert $\tau \in [N]$ following Algorithm 2:*

$$\mathfrak{R}_{\mathrm{E}}^T(\boldsymbol{\theta}, \boldsymbol{\theta}^\tau, \mathcal{S}) \leq \xi\left(k\nu + \nu\sqrt{\frac{T\log N}{2}}\right).$$

Third, we establish the bound on $\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S})$ when $\mathfrak{D} \geq D_{\max} \coloneqq \max_{\tau\in[N]} D^\tau$ (Lemma H.3). Finally, we show that $\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S})$ is bounded when $\mathfrak{D} \leq D_{\min} \coloneqq \min_{\tau\in[N]} D^\tau$ (Lemma H.4). Combining these, we have the following result. A formal version is presented in Appendix H.

Theorem 3.3 (Informal)**.** *For each $t \in [T]$, let $\{f_t(s_t, \cdot)\}_{t=1}^T$ be a sequence of m-strongly convex loss functions with the possibility of at most k of them corrupted by outliers. We assume that the loss for uncorrupted rounds falls within the range $[0, B]$, whereas the loss for corrupted rounds can be unbounded. Let $\boldsymbol{\theta}$ be a sequence of actions chosen by the learner using Algorithm 2 with parameters for the experts chosen from $\mathcal{E}$ as defined in Eq.* (5)*. Then,*

$$\mathfrak{R}_{\mathrm{RD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \mathcal{O}\left(\sqrt{(\mathfrak{D}^2 + \mathfrak{D}V_T)T} + k + \mathfrak{D}V_T + \sqrt{T\log T}\right).$$

A few remarks are in order.

- Matching lower bound: The result matches the lower bound in T and V_T from Jacobsen and Cutkosky [2023] for the no-outlier setting. It is shown in Appendix B that for any online learning algorithm $\mathcal{A}$, there exists a sequence of strongly convex and smooth functions $f_1, \ldots, f_T$, such that

$$\sum_{t=1}^T (f_t(\theta_t) - f_t(\theta_t^*)) = \Omega\left(\sqrt{V_T T}\right).$$

  where $\theta_1, \ldots, \theta_T$ are the actions generated by $\mathcal{A}$.
- Bounded-domain via a single set of experts: Algorithm 2 can be run in a bounded domain using only the set of experts $\mathcal{E}_1$ (defined below Eq. (5)) to choose the step size.
- Parameter-free variant: The bounded-domain modification yields a parameter-free version of Algorithm 1 that requires only $O(\log T)$ experts.
- Removing horizon dependence: The dependence on the time horizon T can be eliminated via the standard doubling trick (see Appendix K).

Challenges in the analysis. We outline the main analytical challenges and the ideas used to overcome them, deferring full proofs to the Appendix.

1. A generic dynamic regret decomposition. For standard OCO algorithms (e.g., online gradient descent), the dynamic clean regret typically admits a decomposition of the form

$$\begin{aligned}\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \sigma_1\frac{\|\theta_1 - \theta_1^*\|}{\alpha} &+ \underbrace{\sigma_2\alpha\sum_{t\in\mathcal{S}} \|\nabla f_t(s_t, \theta_t)\|}_{Q_1} \\ &+ \underbrace{\sigma_3 \sum_{t\in[T]\setminus\mathcal{S}} \|\nabla f_t(s_t, \theta_t)\|\|\theta_t - \theta_t^*\|}_{Q_2} \\ &+ \underbrace{\sigma_4 \sum_{t=1}^T \frac{\|2\theta_{t+1} - \theta_t^* - \theta_{t+1}^*\|\|\theta_t^* - \theta_{t+1}^*\|}{\alpha}}_{Q_3},\end{aligned}$$

for positive constants $\sigma_i, i \in [3]$. Each term reflects a distinct source of difficulty: Q_1 aggregates gradient magnitudes over uncorrupted rounds. Q_2 captures the interaction between gradient size and deviation from the comparator during corrupted rounds. Q_3 depends on the temporal variability of the comparator sequence and scales with the domain geometry. This decomposition makes clear that controlling dynamic regret requires simultaneously bounding gradient growth, corruption effects, and comparator drift.

2. Why standard bounded-domain arguments fail. When gradients and domains are uniformly bounded, choosing α appropriately yields $Q_1, Q_3 = \mathcal{O}(\sqrt{T})$ and $Q_2 = \mathcal{O}(k)$ leading to near-optimal guarantees. Existing robust OCO

analyses (e.g., van Erven et al. [2021]) achieve this by: assuming bounded gradients on uncorrupted rounds (to control Q_1), and filtering out rounds with large gradients (to control Q_2), which requires prior knowledge of k.

However, in our setting, gradient norms may be unbounded, even on uncorrupted rounds, and the corruption budget k is unknown. As a result, neither Q_1 nor Q_2 can be controlled using standard truncation or filtering arguments. Moreover, if the domain itself is unbounded, Q_3 may grow arbitrarily large. These issues fundamentally break existing analyses.

3. Controlling gradient growth via adaptive scaling. Our first key idea is to redesign the LEARN loss so that the analysis operates on the scaled gradients $\eta_t \nabla f_t(s_t, \theta_t)$ rather than on $\nabla f_t(s_t, \theta_t)$ directly. By carefully choosing η_t and exploiting the fact that exponential growth dominates polynomial growth (Lemmas I.1 and I.2), we show that the cumulative contribution of these scaled gradients remains controlled even when the raw gradients are unbounded. This mechanism simultaneously stabilizes Q_1 and Q_2 without requiring bounded gradients or knowledge of k.

4. Handling unbounded domains via an expert framework. The remaining challenge is Q_3, which diverges when the decision domain is unbounded. To address this, we introduce an expert framework in which each expert operates within a carefully chosen bounded domain and uses a fixed step size, and the learner aggregates expert decisions using a meta-algorithm. We prove (in Lemma 3.2 and Appendix H) that this construction effectively controls the domain-dependent terms and prevents uncontrolled growth of Q_3.

5. Resulting guarantee. Combining adaptive gradient scaling with the expert framework yields a dynamic regret bound of $\mathcal{O}(\sqrt{(1+V_T)T} + V_T + \sqrt{T \log T} + k)$ without assuming bounded gradients, bounded domains, or prior knowledge of the corruption level.

4 ONLINE OPTIMIZATION WITH NON-CONVEX LEARN LOSS

As an independent contribution, we present regret bound associated with LEARN losses g_t in an online optimization setup. This is the case when the learner directly observes the invex loss g_t. Note that by construction $\omega_t^* = \arg\min_{\theta \in \Theta} g_t(\theta)$. This section, in isolation, does not consider corruption, so we omit the effect of the side information s_t. In this scenario, the dynamic regret $\mathfrak{R}_{\mathrm{ID}}^T$ associated with the invex losses g_t is defined as

$$\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) \coloneqq \sum_{i=1}^{T} g_t(\theta_t) - \sum_{i=1}^{T} g_t(\omega_t^*) \,. \tag{6}$$

We introduce a versatile framework designed to bound the dynamic regret in the context of online invex optimization. Notably, this framework is not confined to LEARN loss; it extends seamlessly to encompass other invex losses, including but not limited to geodesically convex loss and weakly pseudo-convex loss. We assume that Θ is a closed and bounded convex set and *all* the convex losses are bounded, i.e., $f_t(s_t, \theta) \leq B < \infty$ for all $t \in [T]$. We use the invexity of the LEARN loss and combine it with the results of our unified analysis framework (See Appendix F) to prove an upper bound on the dynamic regret associated with the robust losses g_t. To this end, we show the invexity of g_t constructed at each step $t \in [T]$ by LEARN-OGD (Step 5). We denote

$$\eta_t \coloneqq \eta_t(s_t, \theta_t) = \frac{\exp(-\frac{1}{a} f_t(s_t, \theta_t))}{b + \exp(-\frac{1}{a} f_t(s_t, \theta_t))} \,. \tag{7}$$

Lemma 4.1. *The robust loss $g_t(\theta) \coloneqq g_t(s_t, \theta)$ constructed at each round $t \in [T]$ is a ζ_t-invex function. Moreover, for actions θ_t generated in Algorithm 1 and ω_t^* as defined in Equation* (6)*, we may take ζ_t as*

$$\zeta_t(\omega_t^*, \theta_t) = \frac{\omega_t^* - \theta_t}{\eta_t}.$$

Theorem 4.2 establishes an upper bound on the dynamic regret $\mathfrak{R}_{\mathrm{ID}}^T$ when the learner employs Algorithm 1 and executes the actions $\boldsymbol{\theta}$ by constructing ζ_t-invex robust losses g_t for each $t \in [T]$.

Theorem 4.2. *Consider a sequence of m-strongly convex loss functions $\{f_t(s_t, \cdot)\}_{t=1}^T$. Let $\{g_t\}_{t=1}^T$ be the corresponding sequence of ζ_t-invex robust losses generated via Algorithm 1 (Step 5) and $\boldsymbol{\theta}$ be the sequence of actions chosen by the learner using Algorithm 1. If the learner selects $\alpha_t = \alpha = \sqrt{\frac{4D^2 + 6DV_T}{\psi^2 T}}$, then the dynamic regret*

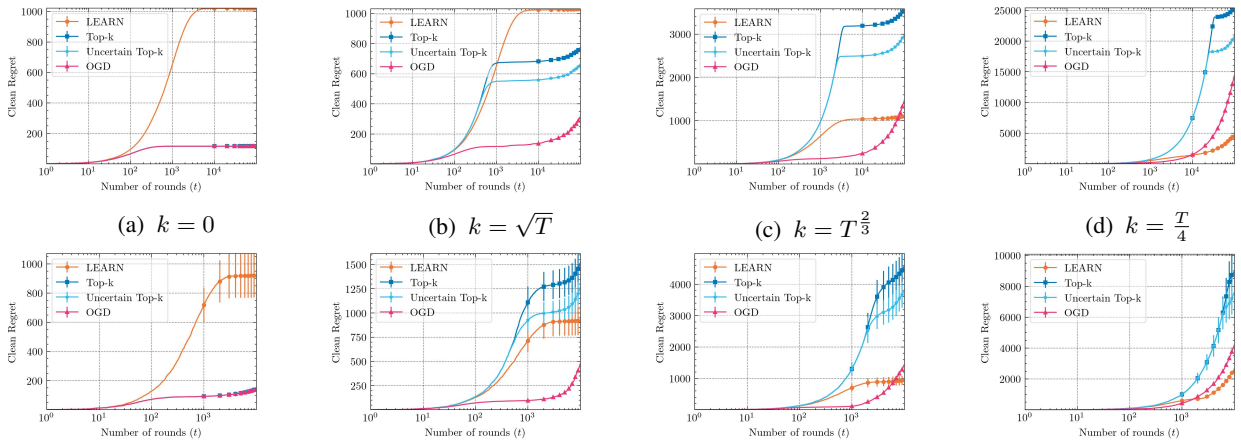
$$\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) = \mathcal{O}\left(\xi\psi\sqrt{(D^2 + DV_T)T}\right).$$

By adopting a similar construction as Algorithm 2 – an adaptive variant of Algorithm 1 – the dependence of α_t on problem parameters such as V_T, G, and L in Theorem 4.2 can be eliminated through the use of an expert framework.

5 EXPERIMENTAL VALIDATION

We focus on two quintessential machine learning problems, online linear regression and classification, to validate our theory. In particular, we focus on online ridge regression and online support vector machine (SVM).

We compare LEARN-OGD (marked LEARN in figures) with the Top-k filter algorithm (Top-k) of van Erven et al. [2021] and vanilla online gradient descent (OGD). OGD is used as the standard learning algorithm within Top-k filter algorithm. Note that Top-k has access to the number of outliers k, but LEARN-OGD does not know k. We also implemented a



(a) $k = 0$ (b) $k = \sqrt{T}$ (c) $k = T^{\frac{2}{3}}$ (d) $k = \frac{T}{4}$

(e) $k = 0$ (f) $k = \sqrt{T}$ (g) $k = T^{\frac{2}{3}}$ (h) $k = \frac{T}{4}$

Figure 2: Clean dynamic regret plots for different algorithms running on the Online Ridge Regression (Top row, Figures 2a-2d) and Online SVM (Bottom row, Figures 2e-2h).

version of Top-k, labeled "Uncertain Top-k", which only has access to an estimate of k fixed at $0.75k$. We conduct experiments in an unbounded domain with potentially large gradients of m-strongly convex loss functions. Note that none of the baselines provide theoretical guarantees for unbounded domains with outliers.

All methods employ the same stepsize $\alpha = 1/\sqrt{T}$. Given that the losses are m-strongly convex, it is possible to set $\alpha_t = 1/(mt)$ for the baselines, although such a choice does not ensure a better regret bound in a dynamic setting. Besides, in our experiments, the value of m is very small ($m = \lambda = 10^{-4}$), which results in a large step size. Consequently, the expected regret bound of $\mathcal{O}(\frac{\log T}{m})$ becomes quite large. A numerical comparison revealed that using $\alpha = 1/\sqrt{T}$ yields smaller regret for the baselines (See Appendix J.5 for additional details). Therefore, we opted for this more favorable step size in our experimental setup.

We report the results for the clean dynamic regrets in Figure 2, averaged across 30 independent runs with the standard errors. We provide additional experimental details in Appendix J.2 and Appendix J.3.

Online linear regression: In this experiment, the learner encounters the following loss function f_t at each round for a fixed $\lambda = 10^{-4}$:

$$f_t\big((x_t, y_t), \theta\big) = \frac{\lambda}{2}\|\theta\|^2 + (y_t - \langle x_t\,, \theta\rangle)^2\,.$$

The response for uncorrupted rounds is generated using the linear model $y_t = \langle \theta^*\,, x_t\rangle + e_t$ with a unit norm $\theta^* \in [-1, 1]^{100}$. The entries of x_t are drawn independently from the normal distribution $\mathcal{N}(0, 1)$. The additive noise e_t is independently chosen from $\mathcal{N}(0, 10^{-6})$. For corrupted rounds, y_t is chosen uniformly at random from the interval $[0, 1]$. Figures 2a-2d show the results from our experiments with varying numbers of outliers over $T = 10^5$ rounds.

Online SVM: Following Shalev-Shwartz et al. [2007], we address the primal version of the online SVM problem with $\lambda = 10^{-4}$. The loss at each round t is

$$f_t\big((x_t, y_t), \theta\big) = \frac{\lambda}{2}\|\theta\|^2 + \max(0, 1 - y_t\langle x_t\,, \theta\rangle)\,.$$

The labels for the uncorrupted rounds are generated according to $y_t = \operatorname{sign}\big(\langle \theta^*\,, x_t\rangle\big)$, where $\theta^* \in [1, 11]^2$. Mislabeling occurs with probability 0.05 when $|\langle \theta^*\,, x_t\rangle| \leq 0.1$. The entries of x_t are drawn independently from $\mathcal{N}(0, 100)$. For corrupted rounds, the adversary flips the sign of y_t. Figures 2e-2h show the experimental results with varying numbers of outliers over $T = 10^4$ rounds.

In the absence of outliers ($k = 0$), the results for both online linear regression (Figure 2a) and Online SVM (Figure 2e) illustrate a significant gap in the cumulative clean regret of LEARN-OGD compared to other baselines. Our approach is intentionally designed to anticipate outliers in the data, leading to cautious and slow updates initially, resulting in the accumulation of regret in the beginning. However, it levels off after the initial few rounds. As the number of outliers increases (Figures 2b and 2c for online linear regression and Figures 2f and 2g for online SVM), the performance of all methods, except ours, deteriorates significantly. Although OGD initially incurs small regret, its growth rate is the largest, implying that it will eventually surpass other methods. LEARN-OGD, on the other hand, maintains roughly the same regret until $k = T^{\frac{2}{3}}$. This demonstrates LEARN-OGD's robustness to outliers, even with higher outlier proportions. These observations corroborate our result in Eq. (17) of Theorem 3.1. Lastly, recall that with $k = T/4$, Eq. (17) does not assure sublinear regret. Correspondingly, all methods exhibit increasing regrets in this regime. Appendix J.4 presents additional experiments conducted in a more dynamic environment [Mokhtari et al., 2016], where $V_T = \mathcal{O}(\sqrt{T})$.

6 CONCLUSION AND FUTURE WORK

In this work, we established tight dynamic regret guarantees within the robust OCO framework. Introducing the LEARN-OGD algorithm, we showcased its effectiveness in handling outliers even when the outlier count k remained unknown, particularly when the loss functions are m-strongly convex. Moreover, our algorithm worked with unbounded domains and accommodated large gradients. Central to our algorithm, we employ LEARN loss, an invex robust loss designed to diminish the influence of outliers during gradient descent-style updates. Additionally, we formulated a unified analysis framework to develop online optimization algorithms tailored for a class of non-convex losses referred to as invex losses.

Introducing $\mathcal{O}(T)$ experts in an unbounded domain can result in significant computational overhead. Previous work, such as Jacobsen and Cutkosky [2023], has acknowledged this challenge, particularly due to the use of the *artificial domain* trick. However, to the best of our knowledge, no effective workaround has been proposed to address this issue. Investigating approaches to reduce the number of experts or eliminate the reliance on the artificial domain trick presents an intriguing direction for future research.

Acknowledgements

V. Y. F. Tan is supported by a Singapore Ministry of Education (MOE) AcRF Tier 1 grant (A-8002934-00-00) and an AcRF Tier 2 grant (A-8004062-00-00).

References

Jonathan T Barron. A general and adaptive robust loss function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4331–4339, 2019.

Albert E. Beaton and John W. Tukey. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, 16(2):147–185, 1974. ISSN 00401706. URL http://www.jstor.org/stable/1267936.

Vasileios Belagiannis, Christian Rupprecht, Gustavo Carneiro, and Nassir Navab. Robust optimization for deep regression. In *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE, December 2015. doi: 10.1109/iccv.2015.324. URL http://dx.doi.org/10.1109/ICCV.2015.324.

Omar Besbes, Yonatan Gur, and Assaf Zeevi. Non-stationary stochastic optimization. *Operations research*, 63(5):1227–1244, 2015.

Aditya Bhaskara and Aravinda Kanchana Ruwanpathirana. Robust algorithms for online k-means clustering. In *Proceedings of the 31st International Conference on Algorithmic Learning Theory*, volume 117 of *Proceedings of Machine Learning Research*, pages 148–173. PMLR, 08 Feb–11 Feb 2020. URL https://proceedings.mlr.press/v117/bhaskara20a.html.

Michael J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, 1996. ISSN 1077-3142. doi: https://doi.org/10.1006/cviu.1996.0006. URL https://www.sciencedirect.com/science/article/pii/S1077314296900065.

Le Chang, Steven Roberts, and Alan Welsh. Robust LASSO regression using Tukey's biweight criterion. *Technometrics*, 60(1):36–47, 2018.

Sitan Chen, Frederic Koehler, Ankur Moitra, and Morris Yau. Online and distribution-free robustness: Regression and contextual bandits with Huber contamination. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 684–695. IEEE, 2022.

T. Gebbie D. Hendricks and D. Wilcox. Detecting intraday financial market states using temporal clustering. *Quantitative Finance*, 16(11):1657–1678, 2016. doi: 10.1080/14697688.2016.1171378.

John E. Dennis and Roy E. Welsch. Techniques for nonlinear least squares and robust regression. *Communications in Statistics - Simulation and Computation*, 7:345–359, 1978. URL https://api.semanticscholar.org/CorpusID:121036822.

Ilias Diakonikolas and Daniel M Kane. *Algorithmic high-dimensional robust statistics*. Cambridge university press, 2023.

Ilias Diakonikolas, Daniel M Kane, Ankit Pensia, and Thanasis Pittas. Streaming algorithms for high-dimensional robust statistics. In *International Conference on Machine Learning*, pages 5061–5117. PMLR, 2022.

Mateusz Fedoryszak, Brent Frederick, Vijay Rajaram, and Changtao Zhong. Real-time event detection on social data streams. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, page 2774–2782, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450362016. doi: 10.1145/3292500.3330689. URL https://doi.org/10.1145/3292500.3330689.

Jiashi Feng, Huan Xu, and Shie Mannor. Outlier robust online learning. *arXiv preprint arXiv:1701.00251*, 2017.

Xiand Gao, Xiaobo Li, and Shuzhong Zhang. Online learning with non-convex losses and non-stationary regret. In *International Conference on Artificial Intelligence and Statistics*, pages 235–243. PMLR, 2018.

Kaan Gokcesu and Hakan Gokcesu. Generalized Huber loss for robust learning and its efficient minimization for a robust statistics. *arXiv preprint arXiv:2108.12627*, 2021.

F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel. *Robust Statistics: The Approach Based on Influence Functions*. Wiley Series in Probability and Statistics. Wiley, 2011. ISBN 9781118150689. URL https://books.google.com.sg/books?id=XK3uhrVefXQC.

Elad Hazan. *Introduction to Online Convex Optimization*. 2023.

Peter J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73 – 101, 1964. doi: 10.1214/aoms/1177703732. URL https://doi.org/10.1214/aoms/1177703732.

Peter J Huber. Robust regression: asymptotics, conjectures and Monte Carlo. *The annals of statistics*, pages 799–821, 1973.

Andrew Jacobsen and Ashok Cutkosky. Unconstrained online learning with unbounded losses. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 14590–14630. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/jacobsen23a.html.

Arun Kejariwal, Sanjeev Kulkarni, and Karthik Ramasamy. Real time analytics: algorithms and systems. *Proceedings of the VLDB Endowment*, 8(12):2040–2041, August 2015. ISSN 2150-8097. doi: 10.14778/2824032.2824132. URL http://dx.doi.org/10.14778/2824032.2824132.

Aurélie C Lozano, Eunho Yang, and Nicolai Meinshausen. Minimum distance estimation for robust high-dimensional regression. *Electronic Journal of Statistics*, 2016.

R.A. Maronna, D.R. Martin, and V.J. Yohai. *Robust Statistics: Theory and Methods*. Wiley Series in Probability and Statistics. Wiley, 2006. ISBN 9780470010921. URL https://books.google.com.sg/books?id=iFVjQgAACAAJ.

Aryan Mokhtari, Shahin Shahrampour, Ali Jadbabaie, and Alejandro Ribeiro. Online optimization in dynamic environments: Improved regret rates for strongly convex problems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 7195–7201. IEEE, 2016.

Francesco Orabona. *A Modern Introduction to Online Learning*. 2023.

Shai Shalev-Shwartz, Yoram Singer, and Nathan Srebro. Pegasos: Primal estimated sub-gradient solver for SVM. In *Proceedings of the 24th International Conference on Machine learning*, pages 807–814, 2007.

Tim van Erven, Sarah Sachs, Wouter M Koolen, and Wojciech Kotlowski. Robust online convex optimization in the presence of outliers. In *Conference on Learning Theory*, pages 4174–4194. PMLR, 2021.

Xi Wang, Zhipeng Tu, Yiguang Hong, Yingyi Wu, and Guodong Shi. Online optimization over Riemannian manifolds. *Journal of Machine Learning Research*, 24 (84):1–67, 2023.

Tianbao Yang. Online accelerated gradient descent and variational regret bounds. *Technical note*, 2016.

Tianbao Yang, Lijun Zhang, Rong Jin, and Jinfeng Yi. Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient. In *International Conference on Machine Learning*, pages 449–457. PMLR, 2016.

Fan Zhang, Junwei Cao, Samee U. Khan, Keqin Li, and Kai Hwang. A task-level adaptive mapreduce framework for real-time streaming data in healthcare applications. *Future Generation Computer Systems*, 43-44:149–160, 2015. ISSN 0167-739X. doi: https://doi.org/10.1016/j.future.2014.06.009.

Jiujia Zhang and Ashok Cutkosky. Unconstrained robust online convex optimization. *International Conference on Machine Learning*, 2025.

Lijun Zhang, Tianbao Yang, Jinfeng Yi, Rong Jin, and Zhi-Hua Zhou. Improved dynamic regret for non-degenerate functions. *Advances in Neural Information Processing Systems*, 30, 2017.

Lijun Zhang, Shiyin Lu, and Zhi-Hua Zhou. Adaptive online learning in dynamic environments. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 1330–1340, Red Hook, NY, USA, 2018. Curran Associates Inc.

Dynamic Regret in Outlier-Oblivious Online Optimization using Nonconvex Robust Losses (Supplementary Material)

Adarsh Barik[1] **Anand Krishna**[2] **Vincent Y. F. Tan**[3]

[1]Department of Computer Science and Engineering, Indian Institute of Technology Delhi, India
[2]Poiro, Evam Labs, Bengaluru, India
[3]Department of Mathematics, Department of Electrical and Computer Engineering, National University of Singapore

A RELATED WORK

This section provides a detailed overview of closely related lines of research that inform our work.

Outlier Robust Online Learning Our work builds on the work of van Erven et al. [2021] where they consider the OCO framework when outliers are present. They specifically address the scenario where k, the number of outliers, is known in advance, and the adversary can arbitrarily corrupt any of the k rounds. The model assumes that inliers have a bounded range (not known), while outliers can extend to an unbounded range. Notably, learner performance is quantified through clean static regret. The core of their algorithm involves maintaining top-k gradients, triggering updates to the online learning algorithm only when the norm of the gradient is less than 2 times the minimum norm within the top-k gradients. They establish robust regret guarantees, offering $\mathcal{O}(G(\mathcal{S})(\sqrt{T}+k))$ for convex functions and $\mathcal{O}(\log T + G(\mathcal{S})k)$ for strongly convex functions, where $G(\mathcal{S})$ represents the norm of the largest gradient considering only clean rounds. Their work, while providing valuable insights, assumes known values for k, a bounded domain, and the existence of a Lipschitz bound $G(\mathcal{S})$ for clean rounds. In contrast, our work extends beyond these assumptions, particularly focusing on dynamic environments where the number of outliers is unknown, the domain is unbounded, and the losses may not be Lipschitz.

Robust Losses A pivotal element of our investigation lies in formulating and examining a non-convex robust loss function. To contextualize our work, we highlight pertinent robust losses featured in the existing literature. Notable examples include Huber loss [Huber, 1964], Tukey's biweight loss [Beaton and Tukey, 1974], Welsch loss [Dennis and Welsch, 1978] and Cauchy loss [Black and Anandan, 1996] among others, each tailored to address different applications. Huber loss, for instance, strikes a balance between mean squared error and mean absolute error, rendering it suitable for robust regression [Huber, 1973]. Similarly, Tukey's biweight loss, known for its efficacy in robust regression, provides resistance against outliers through a truncated quadratic penalty [Chang et al., 2018]. It is worth noting that Tukey's biweight loss, along with Cauchy and Welsch losses, are non-convex and satisfy the *redescending* property [Hampel et al., 2011]. LEARN loss also shares this attribute. For a detailed comparison of LEARN loss with different robust losses see Appendix D. Various generalizations and extensions of these losses have been studied [Barron, 2019, Gokcesu and Gokcesu, 2021], reflecting the extensive exploration and adaptation of robust loss functions within the broader literature.

OCO with unbounded domains and gradients It was only in the recent work of Jacobsen and Cutkosky [2023] that a dynamic regret guarantee was achieved by avoiding the assumption that there exists $\|\nabla f_t\| \leq G$ for all rounds t. They obtain dynamic regret guarantees in the standard OCO framework when the domain is unbounded, and the gradients grow large with the norm of actions ($\|\theta_t\|$). They provide the first algorithm for dynamic regret minimization in the unbounded domain and gradient setup. In this work, we adopt an expert scheme akin to theirs to handle unbounded domains and gradients that can grow large in the robust outlier framework.

Related Applications The work of Chen et al. [2022] studies online linear regression within a realizable context, accounting for small noise using the Huber contamination model [Huber, 1964]. However, our model (similarly the model of van Erven et al. [2021]) differs from theirs significantly. Notably, we do not impose any assumptions of realizability or

control the corruption mechanism through probabilistic assumptions. Although distinct in the techniques employed, it is important to acknowledge a substantial body of related work addressing robust optimization in various theoretical computer science and machine learning problems. For a comprehensive overview, the textbook authored by Diakonikolas and Kane [2023] stands out as an excellent reference.

B LOWER BOUND ON DYNAMIC REGRET

Theorem B.1 (Restatement of Theorem 5 from Zhang et al. [2017])**.** *For any online learning algorithm $\mathcal{A}$, there always exists a sequence of strongly convex and smooth functions $f_1, \ldots, f_T$, such that*

$$\sum_{t=1}^{T} \left(f_t(\theta_t) - f_t(\theta_t^*)\right) = \Omega\left(\sqrt{V_T T}\right) .$$

where $\theta_1, \ldots, \theta_T$ is the actions generated by $\mathcal{A}$.

Proof. Theorem B.1 restates the result of Theorem 5 from Zhang et al. [2017] result in terms of V_T. In their construction, they use $\theta_t^* = \epsilon_t \tau$, where $\epsilon_t \underset{i.i.d.}{\sim} \mathcal{N}(0, I) \in \mathbb{R}^d, \forall t \in \{1, \ldots, T\}$. Verifying that V_T is linear in T is straightforward. In fact,

$$T\sqrt{2(d-1)}|\tau| \leq \mathbb{E}\big[V_T\big] \leq T\sqrt{2d}|\tau|$$

Combining the above with Jensen's inequality gives:

$$\mathbb{E}\left[\sqrt{V_T T}\right] \leq \sqrt{T\mathbb{E}\left[V_T\right]} \leq T(2d)^{\frac{1}{4}}|\tau|^{\frac{1}{2}}$$

Therefore,

$$\mathbb{E}\left[\sum_{t=1}^{T} \left(f_t(\theta_t) - f_t(\theta_t^*)\right)\right] \geq 2dT\tau^2 \geq (2d)^{\frac{3}{4}}|\tau|^{\frac{3}{2}}\mathbb{E}\left[\sqrt{V_T T}\right]$$

If d is treated as constant, it follows that dynamic regret is $\Omega(\sqrt{V_T T})$ even for strongly convex functions. □

Remark B.2. We provide a few comments on the lower bound presented above:

1. The lower bound for dynamic regret can be expressed in various forms. Notably, Zhang et al. [2017] present the lower bound in terms of $S_T \coloneqq \sum_{t=1}^{T} \|\theta_t^* - \theta_{t+1}^*\|^2$. In contrast, we adopt a formulation that incorporates V_T and T, which aligns more closely with the results presented in this work.
2. Our lower bound accounts for worst-case scenarios where V_T may scale linearly with T. Improved bounds, such as those proposed by Yang et al. [2016], can be achieved in more restricted settings by imposing additional constraints on V_T.

Theorem 4.2 in Jacobsen and Cutkosky [2023] establishes a general lower bound on dynamic regret for non-Lipschitz losses under the quadratically bounded gradient assumption (1). Their proof considers a specific sequence of comparators, which does not necessarily correspond to $(\theta_t^*)_{t\in[T]}$. However, extending their results to our setting, where the comparator sequence is explicitly $(\theta_t^*)_{t\in[T]}$, is straightforward. For completeness, we provide the detailed proof below.

Theorem B.3 (Restatement of Theorem 4.2 from Jacobsen and Cutkosky [2023])**.** *For any $\mathfrak{D} > 0$ there is a sequence of functions satisfying* (1) *with $\frac{G}{L} \leq \mathfrak{D}$ such that for any $\gamma \in [0, \frac{1}{2}]$,*

$$\sum_{t=1}^{T} \left(f_t(\theta_t) - f_t(\theta_t^*)\right) \geq \frac{G}{8}\mathfrak{D}^{1-\gamma}[V_T T]^{\gamma} + \frac{L}{16}\mathfrak{D}^{2-\gamma}[V_T T]^{\gamma} .$$

where $\mathfrak{D} \geq \max_t \|\theta_t^*\|$.

Proof. We follow the exact same construction from the proof of Theorem 4.2 in Jacobsen and Cutkosky [2023]. Our task is to convert the bound presented in terms of $(u_t)_{t\in[T]}$ to that of $(\theta_t^*)_{t\in[T]}$. There are two major steps in the proof:

- A lower bound on $\sum_{t=1}^{T}(f_t(\theta_t) - f_t(\theta_t^*))$,
- and an upper bound on V_T.

Observe that the first step comes for free by using the definition of θ_t^*:

$$\sum_{t=1}^{T}(f_t(\theta_t) - f_t(\theta_t^*)) \geq \sum_{t=1}^{T}(f_t(\theta_t) - f_t(u_t)) \geq \frac{1}{2}G\sigma T + \frac{L}{4}T\sigma^2 \ .$$

Next, we only need to show that $V_T = \sum_{t=2}^{T}\|\theta_t^* - \theta_{t-1}^*\| \leq C\sigma T$ where $C > 0$ is an absolute constant. Note that by simply making the gradient of $f_t(\theta)$ zero, we get

$$\theta_t^* = \frac{G + L\sigma}{L}\xi_t,$$
$$\|\theta_t^*\| = \frac{G + L\sigma}{L} \ .$$

Thus,

$$\sum_{t=2}^{T}\|\theta_t^* - \theta_{t-1}^*\| \leq 2\frac{G + L\sigma}{L}T \leq 4\sigma T \ .$$

The last inequality follows due to $\frac{G}{L} \leq \sigma$. The rest of the arguments are the same as Jacobsen and Cutkosky [2023]. In particular, for $\mu \in [0, \frac{1}{2}]$ and $\sigma = \mathfrak{D}T^{-\mu}$, the following inequalities hold true:

$$\sum_{t=1}^{T}(f_t(\theta_t) - f_t(\theta_t^*)) \geq \frac{1}{2}G\mathfrak{D}T^{1-\mu} + \frac{L}{4}T^{1-2\mu}\mathfrak{D}^2$$
$$V_T \leq 4\mathfrak{D}T^{1-\mu}$$

Following a similar algebraic manipulation as Jacobsen and Cutkosky [2023], we show that

$$\sum_{t=1}^{T}(f_t(\theta_t) - f_t(\theta_t^*)) \geq \frac{G}{8}\mathfrak{D}^{1-\gamma}[V_T T]^{\gamma} + \frac{L}{16}\mathfrak{D}^{2-\gamma}[V_T T]^{\gamma} \ .$$

□

C DYNAMIC REGRET IN A BOUNDED DOMAIN USING TOP-K FILTER ALGORITHM

van Erven et al. [2021] established a tight bound on static regret within a bounded domain, assuming prior knowledge of k. For convex losses, they achieved an optimal static regret upper bound of $\mathcal{O}(\sqrt{T} + k)$. In the case of strongly convex losses, this bound is further improved to $\mathcal{O}(\log T + k)$, which remains optimal. In this section, we extend their results to the dynamic setting and establish a clean dynamic regret bound of $\mathcal{O}(\sqrt{V_T T} + k)$ within a bounded domain. Notably, this bound remains tight even for strongly convex functions [Zhang et al., 2017].

We make the following assumptions for our extension:

1. The Lipschitz-adaptive algorithm ALG used by van Erven et al. [2021] is online gradient descent (OGD).
2. The domain Θ is bounded.

Let $\mathcal{S} \subseteq [T]$ be the set of uncorrupted rounds. Then, we are interested in bounding the clean dynamic regret defined as:

$$\mathfrak{R}_{RD}^{T}(\theta, \mathcal{S}) = \sum_{t \in \mathcal{S}}\Big(f_t(s_t, \theta_t) - f_t(s_t, \theta_t^*)\Big) \ ,$$

Let $\mathcal{F} \subset [T]$ denote the rounds flitered out by Algorithm 1 in van Erven et al. [2021], and let $\mathcal{P} = [T] \setminus \mathcal{F}$ denote the rounds that are passed on to ALG. Then,

$$\mathfrak{R}^T_{RD}(\theta, \mathcal{S}) = \underbrace{\sum_{t \in \mathcal{S} \cap \mathcal{P}} \Big(f_t(s_t, \theta_t) - f_t(s_t, \theta^*_t) \Big)}_{T_1} + \underbrace{\sum_{t \in \mathcal{S} \cap \mathcal{F}} \Big(f_t(s_t, \theta_t) - f_t(s_t, \theta^*_t) \Big)}_{T_2}$$

First, we bound T_1.

$$T_1 = \underbrace{\sum_{t \in \mathcal{P}} \Big(f_t(s_t, \theta_t) - f_t(s_t, \theta^*_t) \Big)}_{T_{11}} - \underbrace{\sum_{t \in \mathcal{P} \setminus \mathcal{S}} \Big(f_t(s_t, \theta_t) - f_t(s_t, \theta^*_t) \Big)}_{T_{12}}$$

The bound for the term T_{12} follows directly from van Erven et al. [2021]'s proof of their Theorem 1:

$$T_{12} \leq \sum_{t \in \mathcal{P} \setminus \mathcal{S}} \nabla f_t(s_t, \theta_t)^T (\theta_t - \theta^*_t) \leq 2DG(\mathcal{S})k \ ,$$

where $D = \max_{a,b \in \Theta} \|a - b\|$ and $G(\mathcal{S}) = \max_{t \in \mathcal{S}} \|\nabla f_t(s_t, \theta_t)\|$.

Similarly, we take the bound for T_2 directly from van Erven et al. [2021]'s proof of Theorem 1:

$$T_2 \leq \sum_{t \in \mathcal{S} \cap \mathcal{F}} \nabla f_t(s_t, \theta_t)^T (\theta_t - \theta^*_t) \leq 2DG(\mathcal{S})(k+1)$$

At this point, it only remains to bound T_{11}. Recall that OGD (instantiation of ALG) only sees the rounds in $\mathcal{P}$, and by reindexing them as $t \in \{1, \ldots, |\mathcal{P}|\}$, we can write for all $t \in [|\mathcal{P}|]$:

$$\|\theta_{t+1} - \theta^*_{t+1}\|^2 = \|\theta_{t+1} - \theta^*_t + \theta^*_t - \theta^*_{t+1}\|^2$$

Expanding RHS and by definition of D, we have:

$$\|\theta_{t+1} - \theta^*_{t+1}\|^2 \leq \|\theta_{t+1} - \theta^*_t\|^2 + 3D\|\theta^*_t - \theta^*_{t+1}\|$$

Now, substituting $\theta_{t+1} = \Pi_{\theta \in \Theta}\Big(\theta_t - \alpha_t \nabla f_t(s_t, \theta_t)\Big)$ and using the contractive property of the projection on a convex set, we can write:

$$\|\theta_{t+1} - \theta^*_{t+1}\|^2 \leq \|\theta_t - \alpha_t \nabla f_t(s_t, \theta_t) - \theta^*_t\|^2 + 3D\|\theta^*_t - \theta^*_{t+1}\|$$

Simplifying,

$$\|\theta_{t+1} - \theta^*_{t+1}\|^2 \leq \|\theta_t - \theta^*_t\|^2 + \alpha_t^2 \|\nabla f_t(s_t, \theta_t)\|^2 - 2\alpha_t \nabla f_t(s_t, \theta_t)^T (\theta_t - \theta^*_t) + 3D\|\theta^*_t - \theta^*_{t+1}\|$$

Taking $\alpha_t = \alpha > 0$ and using the convexity of $f_t(s_t, \theta_t)$, we get:

$$2\alpha\Big(f_t(s_t, \theta_t) - f_t(s_t, \theta^*_t) \Big) \leq \|\theta_t - \theta^*_t\|^2 - \|\theta_{t+1} - \theta^*_{t+1}\|^2 + \alpha^2 \|\nabla f_t(s_t, \theta_t)\|^2 + 3D\|\theta^*_t - \theta^*_{t+1}\|$$

Summing it across all $t \in [|\mathcal{P}|]$:

$$T_{11} \leq \frac{\|\theta_1 - \theta^*_1\|^2}{2\alpha} + \frac{\alpha}{2} \sum_{t \in \mathcal{P}} \|\nabla f_t(s_t, \theta_t)\|^2 + \frac{3D}{2\alpha} \sum_{t \in \mathcal{P}} \|\theta^*_t - \theta^*_{t+1}\|$$

We make two observations:

1. $\|\nabla f_t(s_t, \theta_t)\| \leq 2G(\mathcal{S}), \forall t \in \mathcal{P}$ - due to van Erven et al. [2021]
2. $\sum_{t\in\mathcal{P}} \|\theta_t^* - \theta_{t+1}^*\| \leq \sum_{t\in[T]} \|\theta_t^* - \theta_{t+1}^*\| = V_T$

We choose $\alpha = \frac{\sqrt{3DV_T+D^2}}{2G(\mathcal{S})\sqrt{T}}$ to get:

$$T_{11} \leq 2\sqrt{(3DV_T + D^2)T}G(\mathcal{S}) \, .$$

Combining all the above results together, we get

$$\mathfrak{R}_{RD}^T(\theta, \mathcal{S}) \leq 2\sqrt{(3DV_T + D^2)T}G(\mathcal{S}) + 2DG(\mathcal{S})k + 2DG(\mathcal{S})(k+1) = \mathcal{O}(\sqrt{V_T T} + k) \, .$$

D COMPARISON OF ROBUST LOSSES

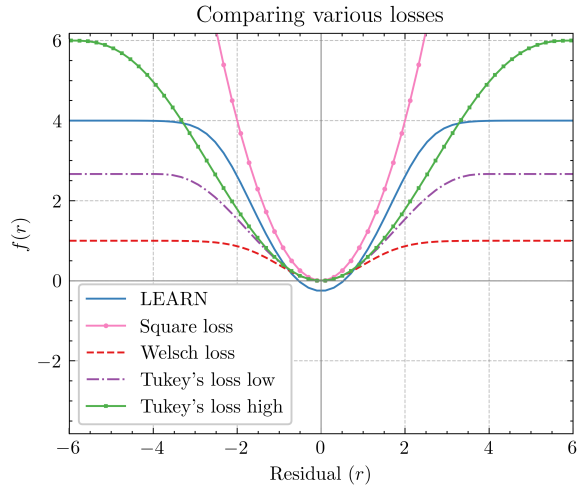


Figure 3: The robust loss closely follows the square loss and flattens out as we move away from the minima.

To understand how the robust loss works to mitigate its susceptibility to outliers, we examine the behavior of different non-convex losses, including our robust loss. We compare how these losses resemble the square loss f near the origin and observe their behavior as we deviate from it.

Let us consider the square loss function $f(r) = r^2$ where r denotes the residual, and the minimum is attained at 0. Our robust loss, denoted by g, instantiated with $f(r)$ with constants $a = 2$ and $b = e^{-2}$, is represented in the Figure 3. Note that g closely follows the behavior of the function f in the vicinity of the origin and gradually levels off as we move away from it. Importantly, the minimizers for both g and f are the same.

In Figure 3, we plot two other non-convex losses that are similar in appearance to our robust loss. First, we examine Tukey's biweight loss, defined as

$$\ell(r) = \begin{cases} \frac{c^2}{6}\left(1 - \left[1 - \left(\frac{r}{c}\right)^2\right]^3\right) & \text{if } |r| \leq c \\ \frac{c^2}{6} & \text{otherwise} \end{cases}$$

where r denotes the residual and c is a tunable parameter. We observe that, even when tuning the parameter c to both high($c = 6$) and low($c = 4$) values, the curve fails to closely follow the curvature of the square loss, contrasting with the behavior of our robust loss. Moving away from the origin, Tukey's loss flattens out at a slower when compared to g.

Next, we consider the Welsch loss defined as

$$h(r) = 1 - \exp\left(-\frac{1}{2}\left(\frac{r}{c}\right)^2\right)$$

Here, the tunable parameter is set to $c = 2$ in Figure 3. However, its limitation of being capped at one impedes its ability to track the underlying square loss precisely. This causes it to flatten out early, making it less suitable for online learning. These observations shed light on the distinctive characteristics of these losses and underscore the efficacy of our proposed robust loss in capturing the desired behavior.

E INVEXITY OF THE ROBUST LOSS

E.1 PROOF OF LEMMA 4.1

Proof. To prove that g_t is a ζ_t-invex function, it suffices to show that all its stationary points are the global minima. Using the monocity of $\log(\cdot)$ and $\exp(\cdot)$ functions, it follows that

$$\overline{\theta}_t \in \arg\min_{\theta \in \mathbb{R}^d} f_t(s_t, \theta) = \arg\min_{\theta \in \mathbb{R}^d} g_t(s_t, \theta)$$

Due to the convexity of f_t, either $\nabla f_t(\overline{\theta}) = 0$ or f_t does not have any stationary point. Now,

$$\nabla g_t(\theta) = \frac{\exp(-\frac{1}{a} f_t(s_t, \theta))}{b + \exp(-\frac{1}{a} f_t(s_t, \theta))} \nabla f_t(\theta) \tag{8}$$

It is obvious that $\nabla g_t(\theta) = 0$ if and only if $\nabla f_t(\theta) = 0$. This immediately implies that g_t is invex. Equivalently, there exists a function $\zeta_t(\vartheta_2, \vartheta_1)$ such that

$$g_t(s_t, \vartheta_2) \geq g_t(s_t, \vartheta_1) + \langle \nabla g_t(s_t, \vartheta_1), \zeta_t(\vartheta_2, \vartheta_1) \rangle .$$

Again, employing the monocity of $\log(\cdot)$ and $\exp(\cdot)$ functions,

$$\omega_t^* \coloneqq \arg\min_{\theta \in \Theta} f_t(s_t, \theta) = \arg\min_{\theta \in \Theta} g_t(s_t, \theta)$$

Let

$$\eta_t = \frac{\exp(-\frac{1}{a} f_t(s_t, \theta_t))}{b + \exp(-\frac{1}{a} f_t(s_t, \theta_t))}, \ \eta_t^* = \frac{\exp(-\frac{1}{a} f_t(s_t, \omega_t^*))}{b + \exp(-\frac{1}{a} f_t(s_t, \omega_t^*))}$$

Observe that $\eta_t^* \geq \eta_t$ as a consequence of $f_t(s_t, \omega_t^*) \leq f_t(s_t, \theta_t)$. Next, we analyze the following quantity:

$$\begin{aligned}
g_t(s_t, \theta_t) - g_t(s_t, \omega_t^*) &= -a \log(\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b) + a \log(\exp(-\frac{1}{a} f_t(s_t, \omega_t^*)) + b) \\
&= a \log \left(\frac{\exp(-\frac{1}{a} f_t(s_t, \omega_t^*)) + b}{\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b} \right) \\
&= a \log \left(\frac{\eta_t}{\eta_t^*} \frac{\exp(-\frac{1}{a} f_t(s_t, \omega_t^*))}{\exp(-\frac{1}{a} f_t(s_t, \theta_t))} \right) \\
&= a \log \frac{\eta_t}{\eta_t^*} + f_t(s_t, \theta_t) - f_t(s_t, \omega_t^*) \\
&\leq f_t(s_t, \theta_t) - f_t(s_t, \omega_t^*) \qquad (9) \\
&\leq \langle \nabla f_t(s_t, \theta_t), \theta_t - \omega_t^* \rangle \qquad (10) \\
&= \frac{1}{\eta_t} \langle \nabla g_t(s_t, \theta_t), \theta_t - \omega_t^* \rangle . \qquad (11)
\end{aligned}$$

Note that the inequality (9) is due to $\eta_t \leq \eta_t^*$, inequality (10) is due to convexity of f_t and Equation (11) is by observing that $\nabla g_t(s_t, \theta_t) = \eta_t \nabla f_t(s_t, \theta_t)$. This completes the proof.

A note on the nondifferentiability of f_t In the scenario of a nondifferentiable yet convex function f_t, the first-order optimality condition implies $0 \in \partial f_t(\theta)$, where $\partial f_t(\cdot)$ denotes the subdifferential set. It's worth noting that the assertion immediately following equation (8) remains valid, taking into account the condition $0 \in \partial f_t(\overline{\theta}_t)$. □

E.2 PROOF OF THEOREM 4.2

Proof. We establish that our problem's setting satisfies the necessary conditions to apply the unified analysis framework detailed in Appendix F. Using the Euclidean distance as our distance function Δ, it is straightforward to verify that Assumption F.8 holds, with $\gamma_1 = 1$ and $\gamma_2 = 1$. The assumption of a bounded domain, Assumption F.9, remains valid when substituting R with $2D$.

Next, we verify Assumption F.10.

$$\begin{aligned}
\|\theta_{t+1} - \theta_t^*\|^2 &= \|\Pi_\Theta(\theta_t - \alpha_t \nabla g_t(s_t, \theta_t)) - \theta_t^*\|^2 \\
&\leq \|\theta_t - \alpha_t \nabla g_t(s_t, \theta_t) - \theta_t^*\|^2 \qquad (12) \\
&= \|\theta_t - \theta_t^*\|^2 + \alpha_t^2 \|\nabla g_t(s_t, \theta_t)\|^2 - 2\alpha_t \langle -\nabla g_t(s_t, \theta_t)\,, \theta_t^* - \theta_t \rangle \\
&= \|\theta_t - \theta_t^*\|^2 + \alpha_t^2 \|\nabla g_t(s_t, \theta_t)\|^2 - 2\eta_t \alpha_t \langle -\nabla g_t(s_t, \theta_t)\,, \frac{\theta_t^* - \theta_t}{\eta_t} \rangle\,,
\end{aligned}$$

where inequality (12) follows from the contraction of the projection on a convex set and η_t is defined in the proof of Lemma 4.1. Using the results from Lemma 4.1, we can verify that the Assumption F.10 holds with $\gamma_3^t = 1$ and $\gamma_4^t = \frac{1}{\eta_t}$. Now, observe that $\max_{t=1}^T \gamma_4^t \leq (1 + b\exp(\frac{B}{a}))$. By choosing a constant step size $\alpha_t = \alpha$ for all $t \in [T]$ and using the results from Theorem F.11, we obtain the following regret bound:

$$\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) \leq \left(1 + b\exp\left(\frac{B}{a}\right)\right)\left(\frac{4D^2}{2\alpha} + \frac{\alpha}{2}\sum_{t=1}^T \|\nabla g_t(\theta_t)\|^2 + \frac{6D\sum_{t=1}^T \|\theta_t^* - \theta_{t+1}^*\|}{2\alpha}\right)\,.$$

We observe that $\nabla g_t(\theta_t) = \eta_t \nabla f_t(\theta_t)$ and then apply Lemma I.2 to write:

$$\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) \leq \left(1 + b\exp\left(\frac{B}{a}\right)\right)\left(\frac{4D^2}{2\alpha} + \frac{\alpha}{2}\psi^2 T + \frac{6DV_T}{2\alpha}\right)\,,$$

where $C_{m,G,L}(a,b)$ is defined in Lemma I.2 and optimal path variation $V_T \coloneqq \sum_{t=1}^T \|\theta_t^* - \theta_{t+1}^*\|$. By choosing,

$$\alpha = \sqrt{\frac{4D^2 + 6DV_T}{\psi^2 T}}\,,$$

we get

$$\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) \leq \left(1 + b\exp\left(\frac{B}{a}\right)\right)\psi\sqrt{(4D^2 + 6DV_T)T}\,.$$

□

F A UNIFIED FRAMEWORK TO BOUND DYNAMIC REGRET FOR ONLINE INVEX OPTIMIZATION

In this section, we develop a unified framework to bound dynamic regret for invex losses using a first-order algorithm. First, we present some definitions and assumptions which will be used later to construct our framework.

Definition F.1 (Invex functions)**.** A differentiable function $g : \mathbb{R}^d \to \mathbb{R}$ is called a ζ-invex function on set Θ if for any $\vartheta_1, \vartheta_2 \in \Theta$ and a vector-valued function $\zeta : \Theta \times \Theta \to \mathbb{R}^d$,

$$g(\vartheta_2) \geq g(\vartheta_1) + \langle \nabla g(\vartheta_1)\,, \zeta(\vartheta_2, \vartheta_1) \rangle$$

where $\nabla g(\vartheta_1)$ is the gradient of g at ϑ_1.

Remark F.2. Conventionally, the notation $\eta(\cdot,\cdot)$ is employed instead of $\zeta(\cdot,\cdot)$ within the definition of an invex function. The preference for ζ in our work arises from the prior use of η for a distinct purpose elsewhere in our paper.

Remark F.3. When working in a non-Euclidean Riemannian manifold the inner product $\langle \cdot, \cdot \rangle$ will be defined in the tangent space of θ_1 induced by the Riemannian metric.

Remark F.4. The choice of $\zeta(\cdot, \cdot)$ may not be unique.

Definition F.5 (First-order Algorithm)**.** We denote an algorithm as $\mathcal{A} := \mathcal{A}(\theta_t, \nabla g(\theta_t), \alpha_t, \zeta)$, where, at round t, it accepts the current iterate $\theta_t \in \Theta$, gradient $\nabla g(\theta_t) \in \mathbb{R}^d$, stepsize $\alpha > 0$, and the function ζ as inputs, and produces the next iterate $\theta_{t+1} \in \Theta$ as its output.

Definition F.6 (Distance function)**.** The function $\Delta : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is called a distance function.[1]

Remark F.7. Although Δ can represent a metric associated with a metric space, we do not explicitly impose this requirement.

We consider an online invex optimization (OIO) framework akin to OCO where the learner engages in a sequential decision-making process. Specifically, in each round $t \in [T]$, the learner selects an action $\theta_t \in \Theta$ and observes a ζ_t-invex loss denoted as $g_t : \Theta \to \mathbb{R}$. Subsequently, the learner employs a first-order optimization algorithm, denoted as $\mathcal{A}$, to perform the next action $\theta_{t+1} \in \Theta$. All the procedural details of this algorithm are presented in the following section.

Algorithm 3 First-order algorithm for OIO

Initialize: $\theta_1 \in \Theta$
for $t = 1, \ldots, T$ **do**
  Play: θ_t
  Observe: ζ_t-invex loss g_t
  Update: $\theta_{t+1} \leftarrow \mathcal{A}(\theta_t, \nabla g_t(\theta_t), \alpha_t, \zeta_t)$
end for

The learner aims to minimize the dynamic regret as defined below:

$$\mathfrak{R}_{\mathrm{ID}}^{T}(\boldsymbol{\theta}) := \sum_{i=1}^{T} g_t(\theta_t) - \sum_{i=1}^{T} g_t(\theta_t^*) \ ,$$

where $\theta_t^* = \arg\min_{\theta \in \Theta} g_t(\theta)$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T)$.

In this section, we use the variable γ with subscripts and superscripts to signify positive and finite ($0 \leq \Gamma < \infty$) parameters. With this notational clarification, we are prepared to enumerate the assumptions essential for our setup.

Assumption F.8 (Generalized Law of Cosines)**.** For all $\vartheta_1, \vartheta_2, \vartheta_3 \in \mathbb{R}^d$, the distance function satisfies the following inequality:

$$\Delta(\vartheta_2, \vartheta_1)^2 \leq \Delta(\vartheta_2, \vartheta_3)^2 + \gamma_1 \Delta(\vartheta_3, \vartheta_1)^2 + 2\gamma_2 \Delta(\vartheta_2, \vartheta_3)\Delta(\vartheta_3, \vartheta_1) \ .$$

Assumption F.9 (Bounded domain)**.** For all $\vartheta_1, \vartheta_2 \in \Theta$,

$$\Delta(\vartheta_1, \vartheta_2) \leq R \ .$$

Assumption F.10 (First-order update property)**.** Let $\theta_{t+1} := \mathcal{A}(\theta_t, \nabla g_t, \alpha_t, \zeta_t)$ be the chosen action at round $t+1$ and $\theta_t^* = \arg\min_{\theta \in \Theta} g_t(\theta)$, then

$$\Delta(\theta_{t+1}, \theta_t^*)^2 \leq \Delta(\theta_t, \theta_t^*)^2 + \gamma_3^t \alpha_t^2 \|\nabla g_t(\theta_t)\|^2 - 2\frac{1}{\gamma_4^t}\alpha_t \langle -\nabla g_t(\theta_t) \, , \zeta_t(\theta_t^*, \theta_t) \rangle \ .$$

Now, we are ready to state our theoretical result.

Theorem F.11. *Let $g_t : \Theta \to \mathbb{R}$ for all $t \in [T]$ be a sequence of ζ_t-invex losses and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T)$ be a sequence of actions generated by a first-order algorithm $\mathcal{A}$. If there exists a distance function $\Delta : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that*

[1] This concept can be extended to encompass non-Euclidean manifolds. Specifically, one can engage with Hadamard manifolds, taking into account geodesically convex functions.

Assumptions F.8, F.9 and F.10 are satisfied for a sequence of stepsizes $(\alpha_1, \ldots, \alpha_T)$, then the following holds:

$$\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) \leq \sum_{t=1}^{T} \frac{\gamma_4^t(\Delta(\theta_t, \theta_t^*)^2 - \Delta(\theta_{t+1}, \theta_{t+1}^*)^2)}{2\alpha_t} + \frac{1}{2}\sum_{t=1}^{T} \alpha_t \gamma_4^t \gamma_3^t \|\nabla g_t(\theta_t)\|^2 + \frac{R\sum_{t=1}^{T} \gamma_4^t(\gamma_1^t + 2\gamma_2^t)\Delta(\theta_t^*, \theta_{t+1}^*)}{2\alpha_t} .$$

Proof. By substituting, $\vartheta_1 = \theta_{t+1}^*, \vartheta_2 = \theta_{t+1}$ and $\vartheta_3 = \theta_t^*$ in Assumption F.8, we get:

$$\begin{aligned}\Delta(\theta_{t+1}, \theta_{t+1}^*)^2 &\leq \Delta(\theta_{t+1}, \theta_t^*)^2 + \gamma_1^t \Delta(\theta_t^*, \theta_{t+1}^*)^2 + 2\gamma_2^t \Delta(\theta_{t+1}, \theta_{t+1}^*)\Delta(\theta_t^*, \theta_{t+1}^*) \\ &\leq \Delta(\theta_{t+1}, \theta_t^*)^2 + R(\gamma_1^t + 2\gamma_2^t)\Delta(\theta_t^*, \theta_{t+1}^*) \qquad (13)\end{aligned}$$

where inequality (13) follows from Assumption F.9. Next, we bound $\Delta(\theta_{t+1}, \theta_t^*)^2$ using Assumption F.10 and get:

$$\begin{aligned}\Delta(\theta_{t+1}, \theta_{t+1}^*)^2 \leq \Delta(\theta_t, \theta_t^*)^2 + \gamma_3^t \alpha_t^2 \|\nabla g_t(\theta_t)\|^2 - 2\frac{1}{\gamma_4^t}\alpha_t \langle -\nabla g_t(\theta_t), \zeta_t(\theta_t^*, \theta_t)\rangle \\ + R(\gamma_1^t + 2\gamma_2^t)\Delta(\theta_t^*, \theta_{t+1}^*)\end{aligned}$$

We get the following after rearranging the terms:

$$\begin{aligned}\langle -\nabla g_t(\theta_t), \zeta_t(\theta_t^*, \theta_t)\rangle \leq \frac{\gamma_4^t(\Delta(\theta_t, \theta_t^*)^2 - \Delta(\theta_{t+1}, \theta_{t+1}^*)^2)}{2\alpha_t} + \frac{\gamma_4^t \gamma_3^t \alpha_t \|\nabla g_t(\theta_t)\|^2}{2} \\ + \frac{R\gamma_4^t(\gamma_1^t + 2\gamma_2^t)\Delta(\theta_t^*, \theta_{t+1}^*)}{2\alpha_t} \qquad (14)\end{aligned}$$

Recall that due to ζ_t-invexity of g_t, we have

$$g_t(\theta_t) - g_t(\theta_t^*) \leq \langle -\nabla g_t(\theta_t), \zeta_t(\theta_t^*, \theta_t)\rangle \qquad (15)$$

Combining inequalities (14) and (15), we get:

$$\begin{aligned}g_t(\theta_t) - g_t(\theta_t^*) \leq \frac{\gamma_4^t(\Delta(\theta_t, \theta_t^*)^2 - \Delta(\theta_{t+1}, \theta_{t+1}^*)^2)}{2\alpha_t} + \frac{\gamma_4^t \gamma_3^t \alpha_t \|\nabla g_t(\theta_t)\|^2}{2} \\ + \frac{R\gamma_4^t(\gamma_1^t + 2\gamma_2^t)\Delta(\theta_t^*, \theta_{t+1}^*)}{2\alpha_t} \qquad (16)\end{aligned}$$

We obtain the desired result by summing inequality (16) from $t = 1$ to T. □

The corollary below shows an explicit bound on $\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta})$ by appropriately choosing α_t.

Corollary F.12. *Let $\gamma_4^{\max} \coloneqq \max_{t=1}^T \gamma_4^t$, $\gamma_{12}^{\max} = \max_{t=1}^T \gamma_1^t + 2\gamma_2^t$ and $\|\nabla g_t(\theta_t)\| \leq \mathcal{G}$ for all $t \in [T]$ in Theorem F.11. Furthermore, define $V_T \coloneqq \sum_{t=1}^T \Delta(\theta_t^*, \theta_{t+1}^*)$. If the learner chooses $\alpha_t = \alpha$ such that*

$$\alpha = \sqrt{\frac{R^2 + R\gamma_{12}^{\max} V_T}{\mathcal{G}^2 \sum_{t=1}^T \gamma_3^t}} ,$$

then the following regret guarantee holds:

$$\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) \leq \gamma_4^{\max}\left(\sqrt{(R^2 + R\gamma_{12}^{\max} V_T)\mathcal{G}^2 \sum_{t=1}^T \gamma_3^t}\right) .$$

Observe that if $\sum_{t=1}^T \gamma_3^t = \mathcal{O}(T)$, then $\mathfrak{R}_{\mathrm{ID}}^T(\boldsymbol{\theta}) = \mathcal{O}(\sqrt{V_T T})$.

G BOUNDS ON THE CLEAN DYNAMIC REGRET IN A BOUNDED DOMAIN

G.1 PROOF OF THEOREM 3.1

Proof. Consider $\|\theta\|_2 \leq D, \forall \theta \in \Theta$. We start our proof by analyzing the following quantity:

$$\begin{aligned}\|\theta_{t+1} - \theta^*_{t+1}\|^2 &= \|\theta_{t+1} - \theta^*_t + \theta^*_t - \theta^*_{t+1}\|^2 \\ &= \|\theta_{t+1} - \theta^*_t\|^2 + \|\theta^*_t - \theta^*_{t+1}\|^2 + 2\langle \theta_{t+1} - \theta^*_t\,, \theta^*_t - \theta^*_{t+1}\rangle \\ &\leq \underbrace{\|\theta_{t+1} - \theta^*_t\|^2}_{\text{Q1}} + \underbrace{6D\|\theta^*_t - \theta^*_{t+1}\|}_{\text{Q2}}\end{aligned} \tag{17}$$

Observe that Q1 can be further decomposed as below.

$$\begin{aligned}\|\theta_{t+1} - \theta^*_t\|^2 &= \|\Pi_\Theta(\theta_t - \alpha_t \nabla g_t(s_t, \theta_t)) - \theta^*_t\|^2 \\ &\leq \|\theta_t - \alpha_t \nabla g_t(s_t, \theta_t) - \theta^*_t\|^2 \qquad (18) \\ &= \|\theta_t - \alpha_t \eta_t \nabla f_t(s_t, \theta_t) - \theta^*_t\|^2 \\ &= \|\theta_t - \theta^*_t\|^2 + \alpha_t^2 \eta_t^2 \|\nabla f_t(s_t, \theta_t)\|^2 - 2\alpha_t \eta_t \langle \theta_t - \theta^*_t\,, \nabla f_t(s_t, \theta_t)\rangle \qquad (19)\end{aligned}$$

The first inequality (18) follows due to the contraction of the projection on the convex sets. We substitute inequality (19) in inequality (17) while keeping track of corrupted and uncorrupted rounds.

First, using convexity of f_t for uncorrupted samples, i.e, $t \in \mathcal{S}$, we get

$$\begin{aligned}2\alpha_t \eta_t (f_t(s_t, \theta_t) - f_t(s_t, \theta^*_t)) \leq \|\theta_t - \theta^*_t\|^2 - \|\theta_{t+1} - \theta^*_{t+1}\|^2 + \alpha_t^2 \eta_t^2 \|\nabla f_t(s_t, \theta_t)\|^2 \\ + 6D\|\theta^*_t - \theta^*_{t+1}\|\,,\end{aligned} \tag{20}$$

and using Cauchy–Schwarz inequality for corrupted samples, i.e., for $t \in [T] \backslash \mathcal{S}$, we get

$$\begin{aligned}0 \leq \|\theta_t - \theta^*_t\|^2 - \|\theta_{t+1} - \theta^*_{t+1}\|^2 + \alpha_t^2 \eta_t^2 \|\nabla f_t(s_t, \theta_t)\|^2 + 2\alpha_t \eta_t \|\theta_t - \theta^*_t\| \|\nabla f_t(s_t, \theta_t)\| \\ + 6D\|\theta^*_t - \theta^*_{t+1}\|\,.\end{aligned} \tag{21}$$

Using the results from Lemma I.2, we can bound:

$$\eta_t \|\nabla f_t(s_t, \theta_t)\| \leq G + \max(\frac{mL}{2ab}, \frac{4a^2 L}{m^2 b}) \coloneqq \psi\,.$$

To bound the last term of inequality (21) consider the term $\eta_t \|\theta_t - \theta^*_t\| \|\nabla f_t(s_t, \theta_t)\|$ for some $t \in [T] \setminus \mathcal{S}$.

$$\begin{aligned}\eta_t \|\theta_t - \theta^*_t\| \|\nabla f_t(s_t, \theta_t)\| &\leq \eta_t G \|\theta_t - \theta^*_t\| + \eta_t L \|\theta_t - \omega^*_t\| \|\theta_t - \theta^*_t\| \\ &\leq \eta_t G(\|\theta_t - \omega^*_t\| + \|\omega^*_t - \theta^*_t\|) + \eta_t L \|\theta_t - \omega^*_t\| (\|\theta_t - \omega^*_t\| + \|\omega^*_t - \theta^*_t\|) \\ &\leq \eta_t G \|\theta_t - \omega^*_t\| + \eta_t G \|\omega^*_t - \theta^*_t\| + \eta_t L \|\theta_t - \omega^*_t\|^2 + \eta_t L \|\omega^*_t \\ &\quad - \theta^*_t\| \|\theta_t - \omega^*_t\|\,.\end{aligned}$$

Next, we will bound the term $\eta_t \|\theta_t - \omega^*_t\|^r, r \in \{1, 2\}$.

$$\begin{aligned}\eta_t \|\theta_t - \omega^*_t\|^r &= \frac{\exp(-\frac{1}{a} f_t(s_t, \theta_t))}{\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b} \|\theta_t - \omega^*_t\|^r \\ &\leq \frac{1}{\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b} \exp(-\frac{1}{a}(f_t(s_t, \omega^*_t))) \exp(-\frac{1}{a}\frac{m}{2}\|\theta_t - \omega^*_t\|^2) \|\theta_t - \omega^*_t\|^r \\ &\leq \frac{1}{b} \exp(-\frac{1}{a}\frac{m}{2}\|\theta_t - \omega^*_t\|^2) \|\theta_t - \omega^*_t\|^r\,.\end{aligned}$$

Using the results from lemma I.1 with $x = \|\theta_t - \omega_t^*\|, c = \frac{m}{2a}$ and $s = 2$,

$$\eta_t \|\theta_t - \omega_t^*\| \leq \frac{1}{b} \max(\frac{m}{2a}, \frac{4a^2}{m^2}) \coloneqq \phi \,,$$
$$\eta_t \|\theta_t - \omega_t^*\|^2 \leq \frac{1}{b} \max(\frac{m}{2a}, \frac{2a}{m}) \coloneqq \kappa \,.$$

Thus,

$$\eta_t \|\theta_t - \theta_t^*\| \|\nabla f_t(s_t, \theta_t)\| \leq G\phi + G\|\omega_t^* - \theta_t^*\| + L\kappa + L\|\omega_t^* - \theta_t^*\|\phi \,.$$

Now we are ready to establish bound on the clean dynamic regret.

We consider $\alpha_t = \alpha$. Observe that, $\eta_t \geq \frac{1}{1+b\exp(\frac{B}{a})}$ and $f_t(s_t, \theta_t) - f_t(s_t, \theta_t^*) \geq 0$ for uncorrupted rounds. Summing up and telescoping inequalities (20) and (21) from $t = 1, \ldots, T$, we get

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \left(1 + b\exp\left(\frac{B}{a}\right)\right) \left(\frac{\|\theta_1 - \theta_1^*\|^2}{2\alpha} + \frac{\alpha}{2}\psi^2 T \right.$$
$$\left. + kG\phi + kG \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\| + kL\kappa + kL\phi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\| + \frac{6DV_T}{2\alpha}\right) \,.$$

Furthermore, observe that $\|\theta_1 - \theta_1^*\| \leq 2D$. Thus,

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \left(1 + b\exp\left(\frac{B}{a}\right)\right) \left(\frac{4D^2}{2\alpha} + \frac{\alpha}{2}\psi^2 T + kG\phi \right.$$
$$\left. + kG \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\| + kL\kappa + kL\phi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\| + \frac{6DV_T}{2\alpha}\right) \,.$$

By choosing,

$$\alpha = \sqrt{\frac{4D^2 + 6DV_T}{\psi^2 T}} \,,$$

we get

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \left(1 + b\exp\left(\frac{B}{a}\right)\right) \left(\psi\sqrt{(4D^2 + 6DV_T)T} + k(G\phi + L\kappa) \right.$$
$$\left. + k(G + L\phi) \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\right) \,.$$

□

H BOUNDS ON THE CLEAN DYNAMIC REGRET IN UNBOUNDED DOMAIN

H.1 FORMAL STATEMENT FOR THEOREM 3.3

Theorem H.1. *For each $t \in [T]$, let $\{f_t(s_t, \cdot)\}_{t=1}^T$ be a sequence of m-strongly convex loss functions with possible k of them corrupted by outliers. We assume that the loss for uncorrupted rounds falls within the range $[0, B]$, whereas the loss for corrupted rounds can be unbounded. Let $\boldsymbol{\theta}$ be a sequence of actions chosen by the learner using Algorithm 2 with parameters for the experts chosen from $\mathcal{E}$ as defined in equation* (5)*. Let $\alpha^* = \sqrt{\frac{32\mathfrak{D}^2 + 24\mathfrak{D}V_T}{T\psi^2}}$. The following regret guarantees hold:*

- *If $\mathfrak{D} \in [\frac{2\epsilon}{T}, \frac{\epsilon 2^T}{T}]$ and $\alpha^* \in [\frac{2}{\sqrt{T}}, \frac{A_{\max}}{\sqrt{T}}]$, then*

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(3\sqrt{\frac{T\psi^2}{2}(16\mathfrak{D}^2 + 12\mathfrak{D}V_T)} + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + k\nu$$
$$+ \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) \,.$$

- *If* $\mathfrak{D} \in [\frac{2\epsilon}{T}, \frac{\epsilon 2^T}{T}]$ *and* $\alpha^* \leq \frac{2}{\sqrt{T}}$*, then*

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(8\mathfrak{D}^2\sqrt{T} + \psi^2\sqrt{T} + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \psi\sqrt{(32\mathfrak{D}^2 + 24\mathfrak{D}V_T)T} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) .$$

- *If* $\mathfrak{D} \in [\frac{2\epsilon}{T}, \frac{\epsilon 2^T}{T}]$ *and* $\alpha^* \geq \frac{A_{\max}}{\sqrt{T}}$*, then*

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(\frac{16\mathfrak{D}^2}{C} + \frac{\psi}{2}\sqrt{(32\mathfrak{D}^2 + 24\mathfrak{D}V_T)T} + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{12\mathfrak{D}V_T}{C} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) .$$

- *If* $\mathfrak{D} < \frac{2\epsilon}{T}$*, then*

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(8\frac{\epsilon^2}{T\sqrt{T}} + \sqrt{T}\psi^2 + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{6\epsilon V_T}{\sqrt{T}} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) .$$

- *If* $\mathfrak{D} > \frac{\epsilon 2^T}{T}$*, then*

$$\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) \leq 2\mathfrak{D}\big(G + 2\mathfrak{D}\big)\log_2 \frac{\mathfrak{D}T}{\epsilon} .$$

H.2 PROOF OF THEOREM 3.3

Proof. Let $\boldsymbol{\theta}^\tau = (\theta_1^\tau, \ldots, \theta_T^\tau)$ be a sequence of actions chosen by the expert τ. Observe that for any expert τ,

$$\begin{aligned}\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S}) &= \sum_{t\in\mathcal{S}} f_t(s_t, \theta_t) - \sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^*) \\ &= \underbrace{\sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^\tau) - \sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^*)}_{\text{Meta Regret } \mathfrak{R}_{\mathrm{M}}^T(\boldsymbol{\theta}^\tau, \mathcal{S})} + \underbrace{\sum_{t\in\mathcal{S}} f_t(s_t, \theta_t) - \sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^\tau)}_{\text{Expert Regret } \mathfrak{R}_{\mathrm{E}}^T(\boldsymbol{\theta}, \boldsymbol{\theta}^\tau, \mathcal{S})}\end{aligned}$$

The proof unfolds in four major steps:

1. Define $\mathfrak{D} \coloneqq \max_{t=1}^T \|\theta_t^*\|$. If $\mathfrak{D} \leq D^\tau$ for any specific expert τ, we establish that $\mathfrak{R}_{\mathrm{M}}^T(\boldsymbol{\theta}^\tau, \mathcal{S})$ is bounded.
2. Demonstrate that $\mathfrak{R}_{\mathrm{E}}^T(\boldsymbol{\theta}, \boldsymbol{\theta}^\tau, \mathcal{S})$ remains bounded when $\mathfrak{D} \leq D^\tau$.
3. Establish the bound on $\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S})$ when $\mathfrak{D} \geq D_{\max} \coloneqq \max_{\tau=1}^N D^\tau$.
4. Conclude the proof by deriving a bound on $\mathfrak{R}_{\mathrm{CD}}^T(\boldsymbol{\theta}, \mathcal{S})$ when $\mathfrak{D} \leq D_{\min} \coloneqq \min_{\tau=1}^N D^\tau$.

First, we bound the meta regret in a bounded domain using the following lemma.

Lemma H.2. *Let* $\mathfrak{D} \leq D^\tau$ *for some expert* τ*. Then,*

$$\mathfrak{R}_{\mathrm{M}}^T(\boldsymbol{\theta}^\tau, \mathcal{S}) \leq \xi\Big(\frac{(D^\tau + \|\theta_1^*\|)^2}{\alpha^\tau} + \frac{\alpha^\tau}{2}\sum_{t=1}^T \big(\eta_t^\tau \|\nabla f_t(s_t, \theta_t^\tau)\|\big)^2 + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{6D^\tau V_T}{\alpha^\tau}\Big)$$

We use the result from Lemma 3.2 to bound the expert regret. In the next lemma, we will bound the clean dynamic regret when $\mathfrak{D}$ is too large.

Lemma H.3. *Let $\mathfrak{D} \geq \max_{\tau=1}^{N} D^\tau := D_{\max}$ and f_t are a sequence of functions satisfying inequality* (1)*, then*

$$\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq 2\mathfrak{D}\big(G + 2\mathfrak{D}\big) \log_2 \frac{\mathfrak{D}T}{\epsilon} \,.$$

Finally, we will show that clean dynamic regret is bounded even when $\mathfrak{D}$ is too small.

Lemma H.4. *Let $\mathfrak{D} \leq D_{\min} = \min_{\tau=1}^{N} D^\tau$, then the following regret guarantee holds:*

$$\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(8\frac{\epsilon^2}{T\sqrt{T}} + \sqrt{T}\psi^2 + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{6\epsilon V_T}{\sqrt{T}} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big)$$

The final step is to bound the dynamic regret when $D_{\min} \leq \mathfrak{D} \leq D_{\max}$, then it must be that for some expert τ,

$$\frac{D^\tau}{2} \leq \mathfrak{D} \leq D^\tau \implies D^\tau \leq 2\mathfrak{D} \,.$$

Using the analysis of Lemma H.2 and Lemma 3.2, we can write:

$$\begin{aligned}
\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) &\leq \xi\Big(\frac{(D^\tau + \|\theta_1^*\|)^2}{\alpha^\tau} + \frac{\alpha^\tau}{2}\sum_{t=1}^{T}\big(\eta_t^\tau \|\nabla f_t(s_t, \theta_t^\tau)\|\big)^2 + k\big(G\phi + L\kappa \\
&\quad + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{6D^\tau V_T}{\alpha^\tau} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) \\
&\leq \xi\Big(\frac{(16\mathfrak{D}^2}{\alpha^\tau} + \frac{\alpha^\tau}{2}T\psi^2 + k\big(G\phi + L\kappa \\
&\quad + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{12\mathfrak{D}V_T}{\alpha^\tau} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big)
\end{aligned}$$

The optimal α^τ would be

$$\alpha^* = \sqrt{\frac{32\mathfrak{D}^2 + 24\mathfrak{D}V_T}{T\psi^2}}$$

If $\alpha^* \leq \alpha_{\min}$, we choose $\alpha^\tau = \alpha_{\min}$, which results in the dynamic regret:

$$\begin{aligned}
\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(&8\mathfrak{D}^2\sqrt{T} + \psi^2\sqrt{T} + k\big(G\phi + L\kappa \\
&+ \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \psi\sqrt{(32\mathfrak{D}^2 + 24\mathfrak{D}V_T)T} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) \,.
\end{aligned}$$

Similarly, if $\alpha^* \geq \alpha_{\max}$, we choose $\alpha^\tau = \alpha_{\max}$, which results in the dynamic regret:

$$\begin{aligned}
\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(&\frac{16\mathfrak{D}^2}{C} + \frac{\psi}{2}\sqrt{(32\mathfrak{D}^2 + 24\mathfrak{D}V_T)T} + k\big(G\phi + L\kappa \\
&+ \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{12\mathfrak{D}V_T}{C} + k\nu + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) \,.
\end{aligned}$$

If α^* falls between $[\alpha_{\min}, \alpha_{\max}]$, then there exists an α^τ such that $\alpha^\tau \leq \alpha^* \leq \alpha^{\tau+1} = 2\alpha^\tau$. We pick this α^τ and bound the dynamic regret as:

$$\begin{aligned}
\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(&3\sqrt{\frac{T\psi^2}{2}(16\mathfrak{D}^2 + 12\mathfrak{D}V_T)} + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + k\nu \\
&+ \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big) \,.
\end{aligned}$$

□

H.3 PROOF OF LEMMA H.2

Proof. We proceed with the proof by analyzing the following quantity.

$$\begin{aligned}
\|\theta_{t+1}^{\tau} - \theta_{t+1}^{*}\|^2 &= \|\theta_{t+1}^{\tau} - \theta_{t}^{*} + \theta_{t}^{*} - \theta_{t+1}^{*}\|^2 \\
&= \|\theta_{t+1}^{\tau} - \theta_{t}^{*}\|^2 + \|\theta_{t}^{*} - \theta_{t+1}^{*}\|^2 + 2\langle \theta_{t+1}^{\tau} - \theta_{t}^{*}, \theta_{t}^{*} - \theta_{t+1}^{*} \rangle \\
&\leq \|\theta_{t+1}^{\tau} - \theta_{t}^{*}\|^2 + 6D^{\tau}\|\theta_{t}^{*} - \theta_{t+1}^{*}\| \\
&= \|\Pi_{\Theta^{\tau}}(\theta_{t}^{\tau} - \alpha_{t}^{\tau}\nabla g_t(s_t, \theta_{t}^{\tau})) - \theta_{t}^{*}\|^2 + 6D^{\tau}\|\theta_{t}^{*} - \theta_{t+1}^{*}\| \\
&\leq \|\theta_{t}^{\tau} - \alpha_{t}^{\tau}\nabla g_t(s_t, \theta_{t}^{\tau}) - \theta_{t}^{*}\|^2 + 6D^{\tau}\|\theta_{t}^{*} - \theta_{t+1}^{*}\| \\
&= \|\theta_{t}^{\tau} - \alpha_{t}^{\tau}\eta_{t}^{\tau}\nabla f_t(s_t, \theta_{t}^{\tau}) - \theta_{t}^{*}\|^2 + 6D^{\tau}\|\theta_{t}^{*} - \theta_{t+1}^{*}\| \\
&= \|\theta_{t}^{\tau} - \theta_{t}^{*}\|^2 + \alpha_{t}^{\tau 2}\eta_{t}^{\tau 2}\|\nabla f_t(s_t, \theta_{t}^{\tau})\|^2 - 2\alpha_{t}^{\tau}\eta_{t}^{\tau}\langle \theta_{t}^{\tau} - \theta_{t}^{*}, \nabla f_t(s_t, \theta_{t}^{\tau})\rangle \\
&\quad + 6D\|\theta_{t}^{*} - \theta_{t+1}^{*}\|
\end{aligned}$$

Using convexity of f_t for uncorrupted samples, i.e, for $t \in \mathcal{S}$,

$$\begin{aligned}
2\alpha_{t}^{\tau}\eta_{t}^{\tau}(f_t(x_t, \theta_{t}^{\tau}) - f_t(x_t, \theta_{t}^{*})) \leq \|\theta_{t}^{\tau} - \theta_{t}^{*}\|^2 - \|\theta_{t+1}^{\tau} - \theta_{t}^{*}\|^2 + \alpha_{t}^{\tau 2}\big(\eta_{t}^{\tau}\|\nabla f_t(x_t, \theta_{t}^{\tau})\|\big)^2 \\
+ 6D^{\tau}\|\theta_{t}^{*} - \theta_{t+1}^{*}\|
\end{aligned}$$

Similarly, for corrupted samples, i.e., $t \in [T]\backslash\mathcal{S}$,

$$\begin{aligned}
0 \leq \|\theta_{t}^{\tau} - \theta_{t}^{*}\|^2 - \|\theta_{t+1}^{\tau} - \theta_{t}^{*}\|^2 + \alpha_{t}^{\tau 2}\big(\eta_{t}^{\tau}\|\nabla f_t(y_t, \theta_{t}^{\tau})\|\big)^2 + 2\alpha_{t}^{\tau}\eta_{t}^{\tau}\|\theta_{t}^{\tau} - \theta_{t}^{*}\|\|\nabla f_t(y_t, \theta_{t}^{\tau})\| \\
+ 6D^{\tau}\|\theta_{t}^{*} - \theta_{t+1}^{*}\|
\end{aligned}$$

Now, we follow the same argument as the dynamic regret bound in Theorem 3.1 and end up with the following bound:

$$\begin{aligned}
\mathfrak{R}_{\mathrm{M}}^{T}(\boldsymbol{\theta}^{\tau}, \mathcal{S}) \leq \Big(1 + b\exp\Big(\frac{B}{a}\Big)\Big)\Big(\frac{\|\theta_{1}^{\tau} - \theta_{1}^{*}\|^2}{\alpha^{\tau}} + \frac{\alpha^{\tau}}{2}\sum_{t=1}^{T}\big(\eta_{t}^{\tau}\|\nabla f_t(s_t, \theta_{t}^{\tau})\|\big)^2 + k\big(G\phi + L\kappa \\
+ \psi \max_{t\in[T]\backslash\mathcal{S}}\|\omega_{t}^{*} - \theta_{t}^{*}\|\big) + \frac{3D^{\tau}V_T}{\alpha^{\tau}}\Big)
\end{aligned}$$

where $V_T = \sum_{t=1}^{T}\|\theta_{t}^{*} - \theta_{t+1}^{*}\|$. The above expression can be further simplified to:

$$\begin{aligned}
\mathfrak{R}_{\mathrm{M}}^{T}(\boldsymbol{\theta}^{\tau}, \mathcal{S}) \leq \xi\Big(\frac{(D^{\tau} + \|\theta_{1}^{*}\|)^2}{\alpha^{\tau}} + \frac{\alpha^{\tau}}{2}\sum_{t=1}^{T}\big(\eta_{t}^{\tau}\|\nabla f_t(s_t, \theta_{t}^{\tau})\|\big)^2 + k\big(G\phi + L\kappa \\
+ \psi \max_{t\in[T]\backslash\mathcal{S}}\|\omega_{t}^{*} - \theta_{t}^{*}\|\big) + \frac{3D^{\tau}V_T}{\alpha^{\tau}}\Big)
\end{aligned}$$

□

H.4 PROOF OF LEMMA 3.2

Proof. Recall from Algorithm 2 that,

$$\begin{aligned}
\rho_{T+1}^{\tau} &= \rho_{T}^{\tau}\exp\big(-\beta\tilde{\eta}_T f_T(s_T, \theta_{T}^{\tau})\big) \\
&= \rho_{1}^{\tau}\exp\big(-\beta\sum_{t=1}^{T}\tilde{\eta}_t f_t(s_t, \theta_{t}^{\tau})\big) \\
&= \exp\big(-\beta\sum_{t\in\mathcal{S}}\tilde{\eta}_t f_t(s_t, \theta_{t}^{\tau})\big)\exp\big(-\beta\sum_{t\in[T]\backslash\mathcal{S}}\tilde{\eta}_t f_t(s_t, \theta_{t}^{\tau})\big)
\end{aligned}$$

Furthermore, using the definition of Z_t:

$$\begin{aligned}
\log \frac{Z_{T+1}}{Z_1} &= \log \frac{\sum_{\tau=1}^N \rho_{T+1}^\tau}{N} \\
&= \log \Big(\sum_{\tau=1}^N \exp\big(-\beta \sum_{t\in\mathcal{S}} \tilde{\eta}_t f_t(s_t, \theta_t^\tau)\big) \exp\big(-\beta \sum_{t\in[T]\setminus\mathcal{S}} \tilde{\eta}_t f_t(s_t, \theta_t^\tau)\big)\Big) - \log N \\
&\geq \log \Big(\exp\big(-\beta \sum_{t\in\mathcal{S}} \tilde{\eta}_t f_t(s_t, \theta_t^\tau)\big)\Big) + \log\Big(\exp(-\beta k\nu)\Big) - \log N \qquad (22) \\
&\geq -\beta \sum_{t\in\mathcal{S}} \tilde{\eta}_t f_t(s_t, \theta_t^\tau) - \beta k\nu - \log N\ , \qquad (23)
\end{aligned}$$

where $\nu := \frac{1}{b}\max(a, \frac{1}{a})$. The inequality (22) follows from Lemma I.3.

Now, consider

$$\begin{aligned}
\log \frac{Z_{t+1}}{Z_t} &= \log \frac{\sum_{\tau=1}^N \rho_{t+1}^\tau}{\sum_{\tau=1}^N \rho_t^\tau} \\
&= \log \frac{\sum_{\tau=1}^N \exp\big(-\beta\tilde{\eta}_t f_t(s_t, \theta_t^\tau)\big)\rho_t^\tau}{\sum_{\tau=1}^N \rho_t^\tau}
\end{aligned}$$

Note that the following inequality holds for any random variable $X \in [a, b]$ and $s > 0$:

$$\log \mathbb{E}[\exp(sX)] \leq s\mathbb{E}[X] + \frac{s^2(b-a)^2}{8}$$

Consider $X = -\tilde{\eta}_t f_t(s_t, \theta_t^\tau)$, $s = \beta$. Also, observe that $X \in [-\nu, 0]$. Thus,

$$\log \frac{Z_{t+1}}{Z_t} \leq -\beta \frac{\sum_{\tau=1}^N \tilde{\eta}_t f_t(s_t, \theta_t^\tau)\rho_t^\tau}{Z_t} + \beta^2 \frac{\nu^2}{8}$$

Using the convexity of $f_t(s_t, \theta_t^\tau)$,

$$\log \frac{Z_{t+1}}{Z_t} \leq -\beta\tilde{\eta}_t f_t(s_t, \theta_t) + \beta^2 \frac{\nu^2}{8}$$

If round t is corrupted, then

$$\log \frac{Z_{t+1}}{Z_t} \leq \beta^2 \frac{\nu^2}{8}$$

Summing through $t = 1$ to T, we get

$$\log \frac{Z_{T+1}}{Z_1} \leq -\beta \sum_{t\in\mathcal{S}} \tilde{\eta}_t f_t(s_t, \theta_t) + \beta^2 T \frac{\nu^2}{8} \qquad (24)$$

Using equations (23) and (24):

$$-\beta \sum_{t\in\mathcal{S}} \tilde{\eta}_t f_t(s_t, \theta_t^\tau) - \beta k\nu - \log N \leq -\beta \sum_{t\in\mathcal{S}} \tilde{\eta}_t f_t(s_t, \theta_t) + \beta^2 T \frac{\nu^2}{8}$$

It follows that,

$$\mathfrak{R}_{\mathrm{E}}^T(\boldsymbol{\theta}, \boldsymbol{\theta}^\tau, \mathcal{S}) \leq \Big(1 + b\exp\Big(\frac{B}{a}\Big)\Big)\Big(k\nu + \frac{\log N}{\beta} + \beta T \frac{\nu^2}{8}\Big)$$

By picking $\beta = \sqrt{\frac{8 \log N}{T\nu^2}}$, we get

$$\mathfrak{R}_{\mathrm{E}}^{T}(\boldsymbol{\theta}, \boldsymbol{\theta}^{\tau}, \mathcal{S}) \leq \Big(1 + b \exp\Big(\frac{B}{a}\Big)\Big)\Big(k\nu + \nu\sqrt{\frac{T \log N}{2}}\Big)$$

□

H.5 PROOF OF LEMMA H.3

Proof. In this setting, we provide a trivial bound on dynamic regret (similar to Jacobsen and Cutkosky [2023]).

$$\begin{aligned}\sum_{t\in\mathcal{S}} f_t(s_t, \theta_t) - \sum_{t\in\mathcal{S}} f_t(s_t, \theta_t^*) &\leq \sum_{t\in\mathcal{S}} \|\nabla f_t(\cdot, \theta_t)\| \|\theta_t - \theta_t^*\| \\ &\leq \sum_{t\in\mathcal{S}} \|\nabla f_t(s_t, \theta_t)\| (D_{\max} + \mathfrak{D}) \\ &\leq 2\mathfrak{D}(G + 2\mathfrak{D})T\end{aligned}$$

Note that due to our construction of set $\mathcal{E}_2$, $D_{\max} = \frac{\epsilon 2^T}{T}$. Clearly, $T \leq \log_2 \frac{\mathfrak{D}T}{\epsilon}$. Thus,

$$\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq 2\mathfrak{D}(G + 2\mathfrak{D}) \log_2 \frac{\mathfrak{D}T}{\epsilon}$$

□

H.6 PROOF OF LEMMA H.4

Proof. Now we consider the setting when $\mathfrak{D} \leq D_{\min}$. Note that since for all $t \in [T]$ we have $\|\theta_t^*\| \leq D_{\min}$, we can use the analysis of Lemma H.2 along with Lemma 3.2 for the expert with parameters $(\alpha_{\min}, D_{\min})$. Formally,

$$\begin{aligned}\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(\frac{(D_{\min} + \|\theta_1^*\|)^2}{\alpha_{\min}} + \frac{\alpha_{\min}}{2}\sum_{t=1}^{T}\big(\eta_t^{\tau}\|\nabla f_t(s_t, \theta_t^{\tau})\|\big)^2 + k\big(G\phi + L\kappa \\ + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{6D_{\min}V_T}{\alpha_{\min}} + k\nu + \nu\sqrt{\frac{T\log N}{2}}\Big)\end{aligned}$$

Recall that,

$$N \leq T \log_2 A_{\max}$$

and

$$\frac{D_{\min}}{\alpha_{\min}} = \frac{\epsilon}{\sqrt{T}}$$

Using the above inequalities, we bound the dynamic regret as

$$\begin{aligned}\mathfrak{R}_{\mathrm{CD}}^{T}(\boldsymbol{\theta}, \mathcal{S}) \leq \xi\Big(8\frac{\epsilon^2}{T\sqrt{T}} + \sqrt{T}\psi^2 + k\big(G\phi + L\kappa + \psi \max_{t\in[T]\setminus\mathcal{S}} \|\omega_t^* - \theta_t^*\|\big) + \frac{6\epsilon V_T}{\sqrt{T}} + k\nu \\ + \nu\sqrt{\frac{T\log(T\log_2 A_{\max})}{2}}\Big)\end{aligned}$$

□

I USEFUL BOUNDS

Within this section, we derive some critical bounds integral to our analytical framework.

Lemma I.1. *For all $c, r, s > 0$ and $x \geq c^{\frac{1}{r}}$, $\exp(-cx^s)x^r \leq \frac{1}{x^s} \leq \frac{1}{c^{\frac{s}{r}}}$.*

Proof. We will show that for $x \geq c^{\frac{1}{r}}$, $x^r \exp(-cx^s) \leq \frac{1}{x^s}$ holds. Below, we assume $x^r \exp(-cx^s) \leq \frac{1}{x^s}$ and find a range of x that satisfies this assumption.

$$x^r \exp(-cx^s) \leq \frac{1}{x^s} \implies r \log x - cx^s \leq -s \log x \implies cx^s \geq \log x^{r+s}$$

Take $x^r \geq c$. As long as $x > 0$, the above inequality is valid. □

Lemma I.2. *Let $f_t(s_t, \theta)$ be an m-strongly convex function whose gradient follows inequality* (1)*. Then,*

$$\eta_t \|\nabla f_t(s_t, \theta_t)\| \leq \psi \coloneqq G + \max(\frac{mL}{2ab}, \frac{4a^2 L}{m^2 b}) \,,$$

where η_t is defined according to Equation (7)*.*

Proof. We begin our proof by analyzing the following quantity.

$$\eta_t \|\nabla f_t(s_t, \theta_t)\| = \frac{\exp(-\frac{1}{a} f_t(s_t, \theta_t))}{\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b} \|\nabla f_t(s_t, \theta_t)\|$$

Note that due to m-strong convexity of $f_t(s_t, \cdot)$, we have

$$f_t(s_t, \omega_t^*) + \langle \nabla f_t(s_t, \omega_t^*) \,, \theta_t - \omega_t^* \rangle + \frac{m}{2} \|\theta_t - \omega_t^*\|^2 \leq f_t(s_t, \theta_t)$$

where $\omega_t^* = \arg\min_{\theta \in \Theta} f_t(s_t, \theta)$. Observe that $\langle \nabla f_t(s_t, \omega_t^*) \,, \theta_t - \omega_t^* \rangle \geq 0$. Thus,

$$f_t(s_t, \omega_t^*) + \frac{m}{2} \|\theta_t - \omega_t^*\|^2 \leq f_t(s_t, \theta_t)$$

This leads to the following inequality:

$$\begin{aligned}
\eta_t \|\nabla f_t(s_t, \theta_t)\| &\leq \eta_t G + L \frac{\exp(-\frac{1}{a}(f_t(s_t, w_t^*))) \exp(-\frac{1}{a}\frac{m}{2}\|\theta_t - \omega_t^*\|^2) \|\theta_t - \omega_t^*\|}{\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b} \\
&\leq G + \frac{L}{\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b} \exp(-\frac{1}{a}\frac{m}{2}\|\theta_t - \omega_t^*\|^2) \|\theta_t - \omega_t^*\| \\
&\leq G + \frac{L}{b} \exp(-\frac{1}{a}\frac{m}{2}\|\theta_t - \omega_t^*\|^2) \|\theta_t - \omega_t^*\|
\end{aligned}$$

As the first case, let $\|\theta_t - \omega_t^*\| < \frac{m}{2a}$, then

$$\eta_t \|\nabla f_t(s_t, \theta_t)\| \leq G + \frac{mL}{2ab} \,.$$

Now, let $\|\theta_t - \omega_t^*\| \geq \frac{m}{2a}$, then using the results from Lemma I.1, we have

$$\eta_t \|\nabla f_t(s_t, \theta_t)\| \leq G + \frac{4a^2 L}{m^2 b}$$

It follows that,

$$\eta_t \|\nabla f_t(s_t, \theta_t)\| \leq \psi \coloneqq G + \max(\frac{mL}{2ab}, \frac{4a^2 L}{m^2 b}) \,.$$

□

Lemma I.3. *Let $f_t(s_t, \theta)$ be a function whose gradient follows inequality* (1)*. Then,*

$$\eta_t f_t(s_t, \theta_t) \leq \frac{1}{b} \max(a, \frac{1}{a}) \coloneqq \nu \ ,$$

where η_t is defined according to Equation (7)*.*

Proof. We follow a similar approach as Lemma I.2:

$$\begin{aligned} \eta_t f_t(s_t, \theta_t) &= \frac{\exp(-\frac{1}{a} f_t(s_t, \theta_t))}{\exp(-\frac{1}{a} f_t(s_t, \theta_t)) + b} f_t(s_t, \theta_t) \\ &\leq \frac{1}{b} \exp(-\frac{1}{a} f_t(s_t, \theta_t)) f_t(s_t, \theta_t) \end{aligned}$$

Note that if $f_t(s_t, \theta_t) \leq \frac{1}{a}$, then

$$\eta_t f_t(s_t, \theta_t) \leq \frac{1}{ab} \ .$$

Using the results from Lemma I.1, if $f_t(s_t, \theta_t) \geq \frac{1}{a}$, then

$$\eta_t f_t(s_t, \theta_t) \leq \frac{a}{b} \ .$$

Thus,

$$\eta_t f_t(s_t, \theta_t) \leq \frac{1}{b} \max(a, \frac{1}{a}) \coloneqq \nu \ .$$

□

J EXPERIMENTAL VALIDATION

In this section, we perform numerical experiments to substantiate our theoretical findings. The experiments were conducted on a MacBook Pro running macOS 14.4.1, equipped with 32 GB of memory and an Apple M2 Max chip.

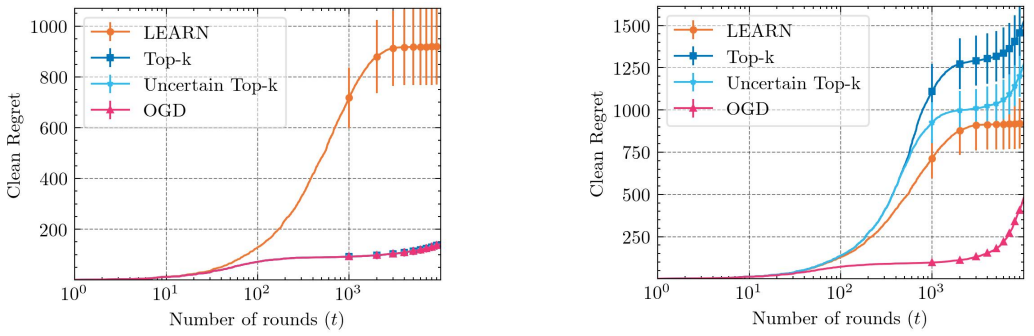
J.1 BASELINE ALGORITHMS

We conduct a comparative analysis, pitting LEARN against several baseline algorithms, each serving as a benchmark in our experiments.

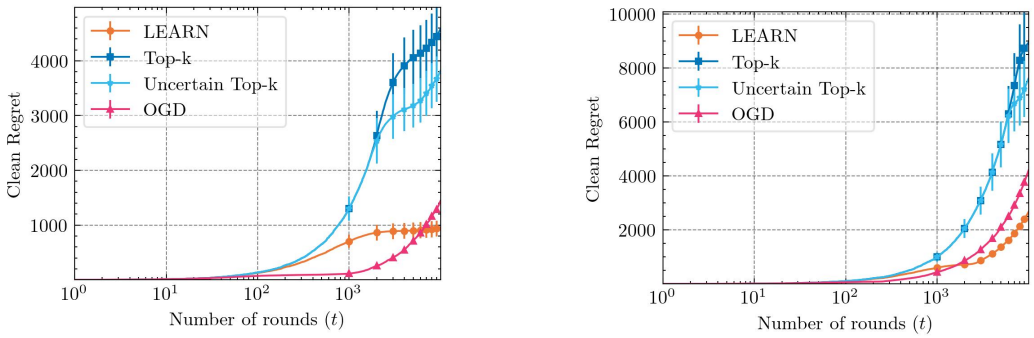
- Top-k Filter Algorithm: Developed by van Erven et al. [2021], this algorithm operates under the assumption that the domain of the Online Convex Optimization (OCO) problem and the gradients of the uncorrupted loss functions f_t are bounded. The core concept involves filtering out k rounds with the largest gradients, assuming they might be outliers. If any outlier round is not filtered, it does not influence the update too much due to the bounded gradient assumption on the uncorrupted loss. The algorithm requires precise knowledge of k and provides an $\mathcal{O}(\sqrt{T} + k)$ bound on clean static regret under the bounded domain and Lipschitz gradient assumption. While this algorithm is analyzed for static regret, it can be extended easily to the setting of dynamic regret. We use the online gradient descent (OGD) as the standard learning algorithm within Top-k filter algorithm.
- Uncertain Top-k Filter Algorithm: We devise a variant of the Top-k filter algorithm equipped with an estimate of the number of outliers. This estimate was fixed at $0.75k$. It is important to emphasize that this manufactured algorithm does not come with any accompanying theoretical guarantees.
- Online Gradient Descent (OGD) Algorithm: As our third baseline, we select OGD, a workhorse for many machine learning problems. OGD operates without explicit consideration of the presence of outliers.

In the absence of outliers, all baselines converge to Online Gradient Descent (OGD). It is imperative to note that none of these methods provide theoretical guarantees in an unbounded domain with outliers. Despite the absence of theoretical guarantees, we selected these algorithms as baselines due to a shortage of alternatives that effectively address both unbounded domains and robustness to outliers.

J.2 ONLINE SUPPORT VECTOR MACHINE



(a) $k = 0$ (b) $k = \sqrt{T}$



(c) $k = T^{\frac{2}{3}}$ (d) $k = \frac{T}{4}$

Figure 4: Clean Regret for Online SVM in Presence of Varying Number of Outliers (k).

In our initial experimental configuration, we conducted a comparative analysis of LEARN against other baseline methods for a classification problem, employing soft-margin Support Vector Machines (SVM) in an online setting. The specifics of the experimental setup are detailed below:

Uncorrupted Data Generation In our experiments, the data for uncorrupted rounds ($t \in \mathcal{S}$) was generated following a specific data generation process:

$$y_t = \text{sign}\big(\langle \theta^* , x_t \rangle\big)$$

Here, the entries of $\theta^* \in \mathbb{R}^2$ were uniformly drawn at random from the interval $[1, 11]$. Note that $s_t = (x_t, y_t)$ is the side information in this case. The entries of x_t for uncorrupted rounds were independently drawn from a normal distribution $\mathcal{N}(0, 100)$. To allow for some mislabeling, the label of a round was flipped with a probability of 0.05 if $|\ \langle \theta^* , x_t \rangle\ | \leq 0.1$.

Corrupted Data Generation In the experimental setup, we executed a total of 10^4 rounds, allowing the adversary the flexibility to corrupt any k out of the T rounds, with the choice made uniformly at random. The experiments encompassed scenarios where k took on values from the set $\left\{0, \sqrt{T}, T^{\frac{2}{3}}, \frac{T}{4}\right\}$.

For the corrupted rounds, the labeling process involved flipping the label, specifically by setting $y_t = -\text{sign}\big(\langle \theta^* , x_t \rangle\big)$.

Strongly Convex Loss Function Drawing inspiration from Shalev-Shwartz et al. [2007], we adopted a fixed loss function f_t for each round, defined as:

$$f_t\big((x_t, y_t), \theta\big) = \frac{\lambda}{2}\|\theta\|^2 + \max(0, 1 - y_t \langle x_t , \theta \rangle)\ .$$

Similar to Shalev-Shwartz et al. [2007], we set the parameter λ to 10^{-4}. To maintain an unbounded domain for θ, we considered it as belonging to $\mathbb{R}^2$. It should be noted that in a noiseless setting with sufficiently small λ, dynamic optimal actions θ_t^* match the ground truth θ^*.

Choice of Parameters In the case of LEARN, the parameter a was set to 10^4, while b was maintained at 10. This selection was determined through a grid search across a small range of values. We did not undertake specific tuning efforts for optimality. Across all methods, including both baselines and LEARN, a consistent step size of $\alpha = \frac{1}{\sqrt{T}}$ was employed. We refer to Appendix J.5 for additional experiments involving different step sizes for baselines.

Results The experiments were executed across 30 independent runs, and the outcomes are presented in Figure 4. Referring back to Theorem 3.3, we anticipate the clean regret to vary in the order of $\mathcal{O}(\sqrt{T} + k)$. Since $V_T = 0$, this implies sublinear regret for LEARN until $k = T^{\frac{2}{3}}$ and constant regret when $k = \frac{T}{4}$. The same behavior can be observed from the plots in Figure 4. The empirical observations derived from the experiment are outlined below:

- LEARN adopts a cautious updating strategy in the uncorrupted regime: This is illustrated by the observed gap in clean regret in Figure 4a. In this context, the clean regret for other baselines flattens at a significantly lower value compared to LEARN. This behavior stems from LEARN's inherent cautious update mechanism, specifically designed to anticipate the presence of corrupted rounds. Consequently, LEARN experiences a substantial accumulation of regret initially, navigating a careful trajectory until it converges to the optimal action, leading to the eventual flattening of the regret curve.
- LEARN exhibits robustness to outliers: With an increasing value of k, the performance of the baselines markedly deteriorates, while LEARN remains relatively unaffected until Figure 4d. This is evident when inspecting the maximum regret values across various curves. For LEARN, the maximum regret hovers around 900 until Figure 4d, whereas other baselines experience a steep surge in regret in the presence of outliers.
- LEARN is good at capturing the underlying ground truth: In cases where the data adheres to a specific generating process, the presence of a flat curve indicates convergence to the optimal action (θ^*). Remarkably, LEARN maintains a flat region in Figures 4a to 4c even in the presence of outliers, showcasing its ability to learn and adapt to the true underlying dynamics. In contrast, other methods falter in achieving this sustained flatness. Upon closer examination of the rate of growth, it becomes apparent that OGD is particularly susceptible to the disruptive influence of outliers.
- A constant proportion of outliers proves detrimental for all methods: As anticipated by Theorem 3.3, LEARN begins to accumulate increasing regret when $k = \frac{T}{4}$. This phenomenon is evident in Figure 4d. In fact, all the baseline methods exhibit poor performance in this regime.

Visualizing the Decision Boundary As the data (side information) adheres to a noiseless generative model, the online SVM problem can be thought of as a learning task where the objective is to learn the classification parameter θ^*. The decision boundaries learned by different methods with varying numbers of outliers are illustrated in Figure 5 for a single run. The decision boundary for θ^* (plotted as $\langle \theta^*, x \rangle = 0$) is also depicted. To enhance clarity, only 500 rounds out of the total $T = 10^4$ rounds are randomly selected for plotting.

In scenarios with uncorrupted data, all methods accurately learn θ^*. However, as the number of corrupted rounds increases, the decision boundaries for other baselines gradually deviate from the optimal boundary. In contrast, LEARN exhibits robustness against outliers, maintaining a decision boundary close to the optimal configuration even in the presence of corrupted data.

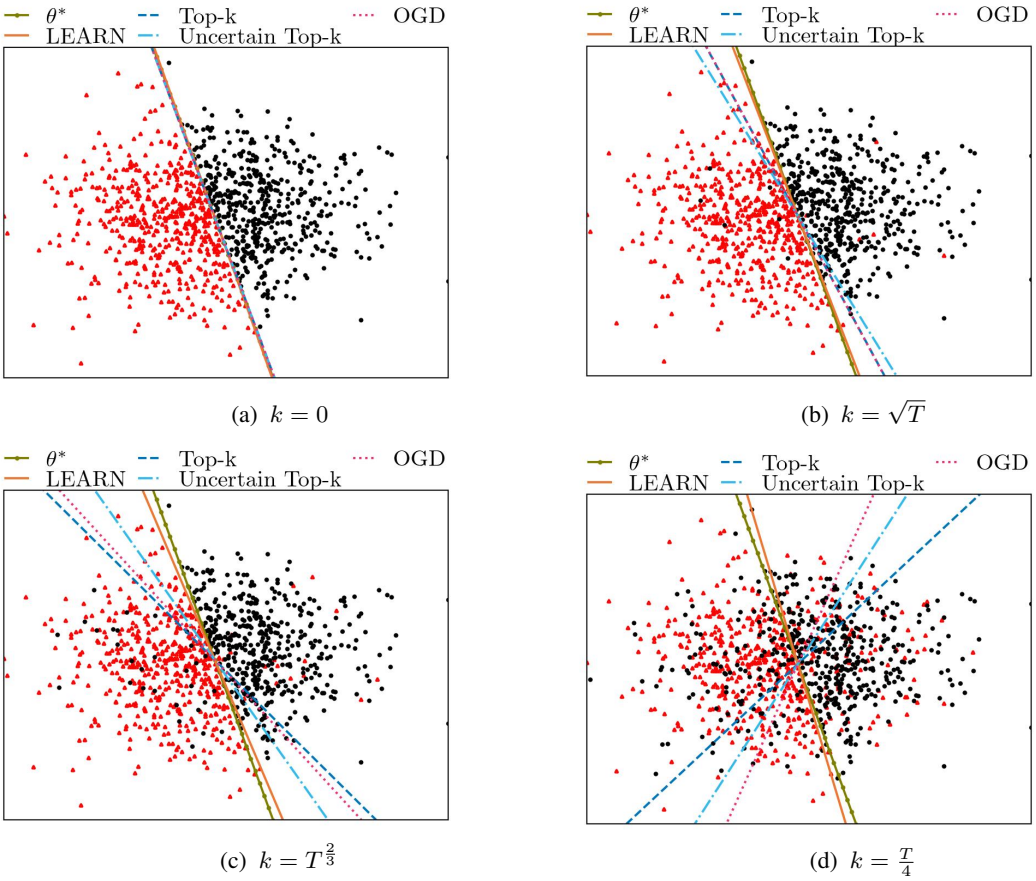
J.3 ONLINE LINEAR REGRESSION

In our subsequent experimental setup, we undertook a comparative analysis, juxtaposing the performance of LEARN against other baseline methods for a regression problem. Specifically, we explored the online ridge regression setting. The details of this experimental configuration are elaborated below:

Uncorrupted Data Generation The data for uncorrupted rounds ($t \in \mathcal{S}$) adhered to a specific data generation process:

$$y_t = \langle \theta^*, x_t \rangle + e_t$$

In this context, the entries of $\theta^* \in \mathbb{R}^{100}$ were uniformly drawn at random from the interval $[-1, 1]$, followed by normalization to a unit norm. Note that $s_t = (x_t, y_t)$ is again the side information. The entries of x_t for uncorrupted rounds were independently drawn from a normal distribution $\mathcal{N}(0, 1)$. Additionally, a small additive noise e_t was independently chosen from $\mathcal{N}(0, 10^{-6})$.



(a) $k = 0$ (b) $k = \sqrt{T}$

(c) $k = T^{\frac{2}{3}}$ (d) $k = \frac{T}{4}$

Figure 5: Visualization of Decision Boundary in Presence of Varying Number of Outliers (k).

Corrupted Data Generation We conducted a total of 10^5 rounds, providing the adversary with the flexibility to corrupt any k out of the T rounds, with the choice made uniformly at random. The experiments covered scenarios where k assumed values from the set $\left\{0, \sqrt{T}, T^{\frac{2}{3}}, \frac{T}{4}\right\}$. In the context of corrupted rounds, y_t was chosen uniformly at random from the interval $[0, 1]$.

Strongly Convex Loss Function We employed the following fixed loss function f_t for each round:

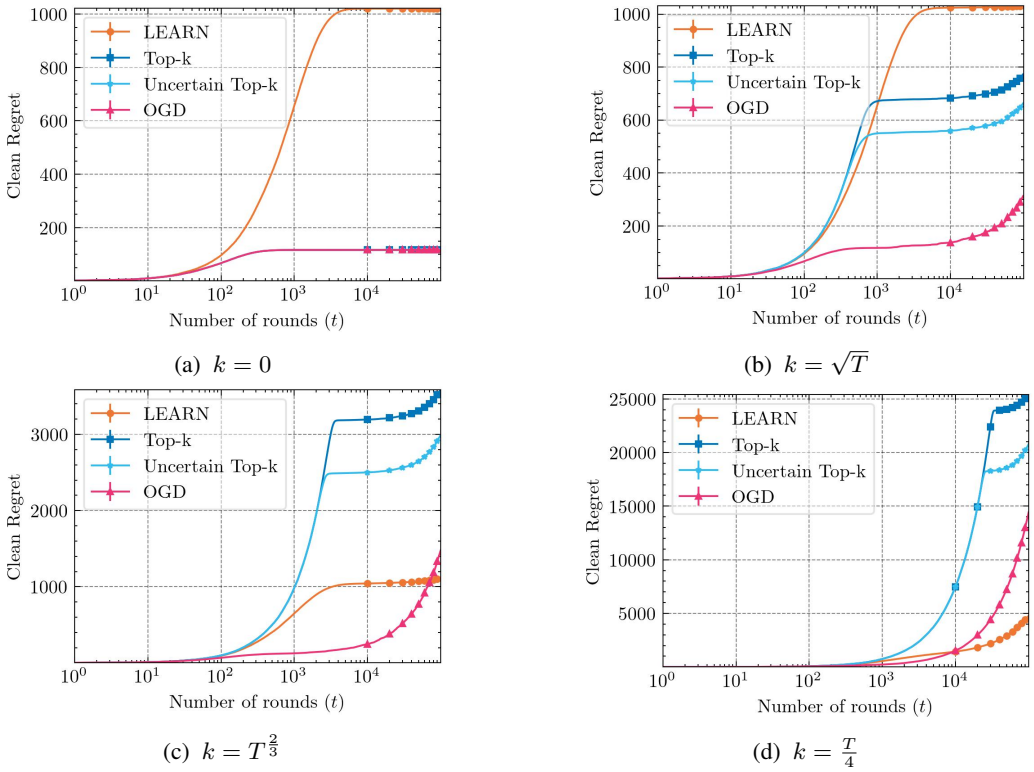
$$f_t\big((x_t, y_t), \theta\big) = \frac{\lambda}{2}\|\theta\|^2 + (y_t - \langle x_t\,, \theta\rangle)^2\,. \tag{25}$$

The parameter λ was set to 10^{-4}. To enforce an unbounded domain for θ, we considered it to be an element of $\mathbb{R}^{100}$. In each round, the comparator θ_t^* was selected by minimizing the function $f_t((x_t, y_t), \theta)$ as defined in Equation (25). Although it may not precisely match the ground truth θ^* due to the low influence of additive noise, our experimental results indicate its proximity to θ^*.

Choice of Parameters In the context of LEARN, the parameter a was set to 10, and b was held at 10. Once again, this choice resulted from a grid search across a limited range of values, with no dedicated tuning efforts for optimality. Consistently, across all methods, including both baselines and LEARN, a uniform step size of $\alpha = \frac{1}{\sqrt{T}}$ was utilized.

Results The experiments were executed across 30 independent runs, and the outcomes are presented in Figure 6. The results show the similar behavior as in the case of online SVM. We recall them here in the context of ridge regression:

- Similar to Section J.2, LEARN again employs a cautious update strategy, evident in the gap in clean regret in Figure 6a. Unlike other baselines, LEARN accumulates initial regret due to its cautious update mechanism designed for anticipating corrupted rounds, leading to slow flattening of the regret curve.



(a) $k = 0$ (b) $k = \sqrt{T}$

(c) $k = T^{\frac{2}{3}}$ (d) $k = \frac{T}{4}$

Figure 6: Clean Dynamic Regret for Online Linear Regression in Presence of Varying Number of Outliers (k).

- In the presence of outliers (until Figure 6d), LEARN remains robust, maintaining a maximum regret around 1000, while other baselines experience a sharp surge in regret.
- In our experiments, due to low noise levels, LEARN captures the underlying ground truth, sustaining a flat region in Figures 6a to 6c even with outliers. This contrasts with other methods struggling to achieve sustained flatness. OGD, in particular, proves susceptible to outlier influence.
- A constant proportion of outliers adversely affects all methods, as expected from Theorem 3.3. LEARN starts accumulating increasing regret when $k = \frac{T}{4}$, depicted in Figure 6d, with all baseline methods exhibiting poor performance in this regime.

J.4 DYNAMIC REGRET WITH DIMINISHING VARIATIONS IN OPTIMAL POINTS

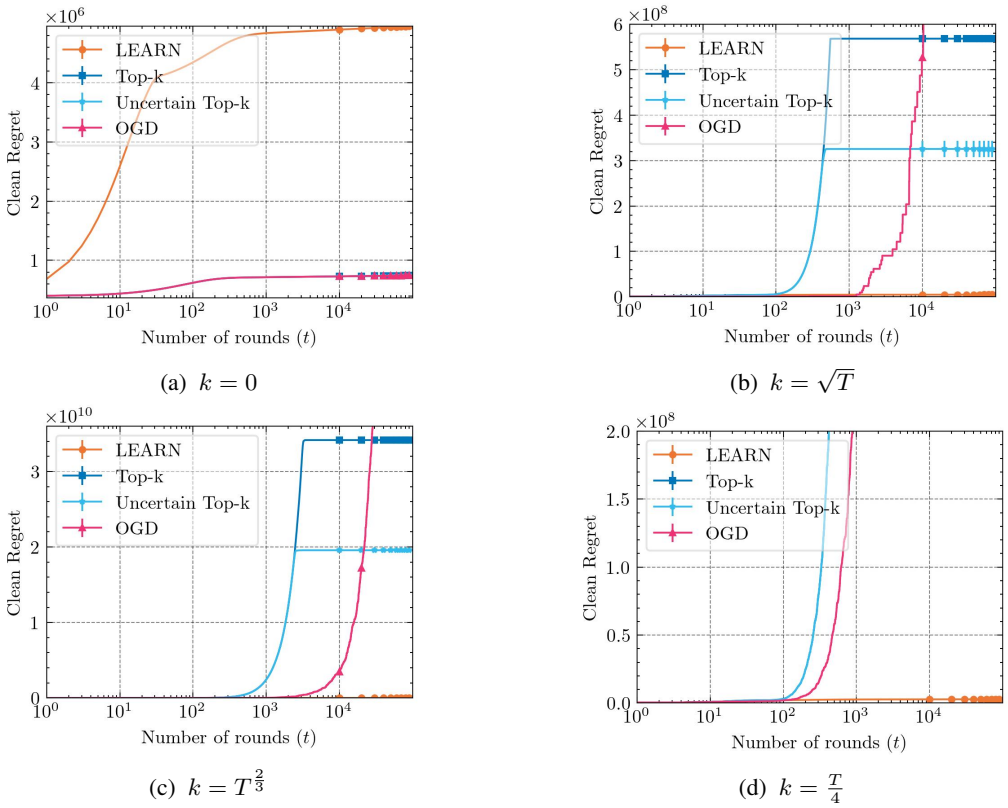
Next, we examine the impact of a dynamically changing environment on the performance of LEARN. Adopting the setup from Mokhtari et al. [2016], we extend it by allowing for an unbounded domain size, as opposed to the original circular domain. For comparison, we also evaluate the performance of baseline methods. The details of the experimental configuration are provided below:

Data Generation and the loss function We define the decision variable as $\theta = \begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix} \in \mathbb{R}^2$ and consider a quadratic loss function $f_t((a_t, b_t, c_t), \theta)$ at time t , expressed as:

$$f_t\left((a_t, b_t, c_t), \theta\right) = \rho(\theta_1 - a_t)^2 + (\theta_2 - b_t)^2 + c_t \ .$$

The variables (a_t, b_t, c_t) act as side information. Clearly, $\theta_t^* = \begin{bmatrix}a_t \\ b_t\end{bmatrix}$. Using the setting from Mokhtari et al. [2016], we fix $\rho = 100$ and $b_t = 100$ and $c_t = 5$ for all rounds. For uncorrupted rounds, a_t satisfies the recursive formula

$$a_{t+1} = a_t + 5\sqrt{\frac{1}{t}}, \qquad t \in [T] \ ,$$



(a) $k = 0$ (b) $k = \sqrt{T}$

(c) $k = T^{\frac{2}{3}}$ (d) $k = \frac{T}{4}$

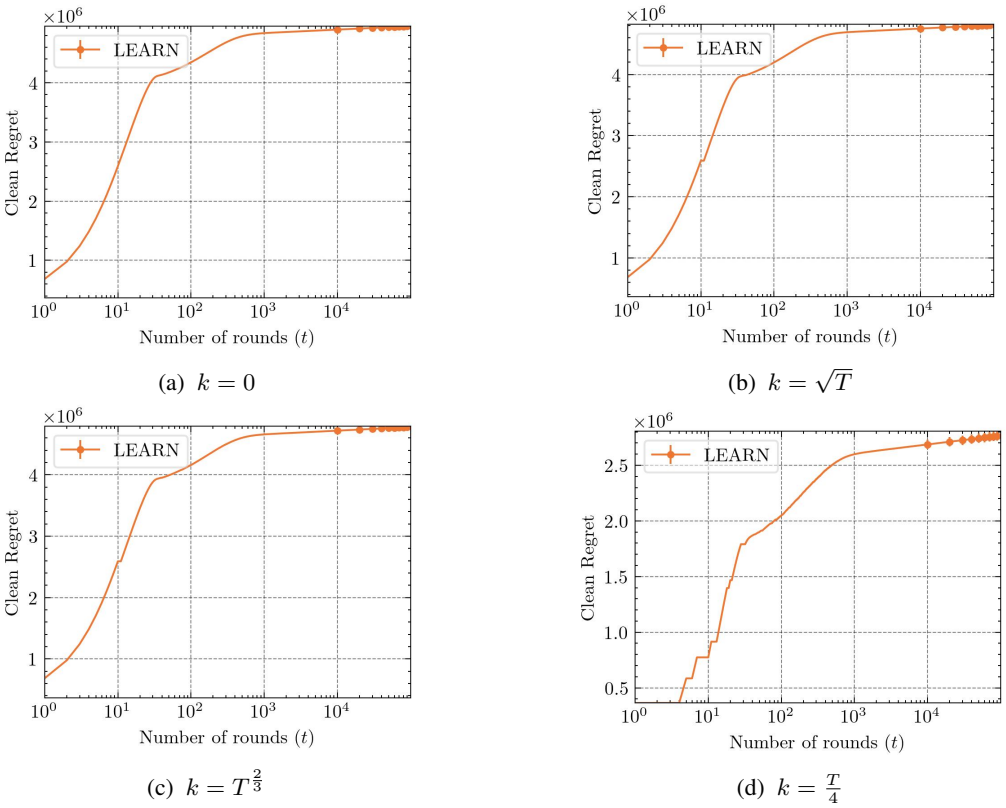
Figure 7: Clean Dynamic Regret for a Sequence of Quadratic Loss Functions in the Presence of Varying Number of Outliers (k).

with a_1 fixed at -60. In corrupted rounds, the adversary selects a_t uniformly at random from $[-50, 50]$. The experiments were run for $T = 10^5$ rounds, allowing the adversary to corrupt any k rounds chosen uniformly at random. The scenarios explored include k values from the set $\left\{0, \sqrt{T}, T^{\frac{2}{3}}, \frac{T}{4}\right\}$.

Choice of Parameters In the context of LEARN, the parameter a was set to 10^4, and b was held at 1. Once again, this choice resulted from a grid search across a limited range of values, with no dedicated tuning efforts for optimality. Consistently, across all methods, including both baselines and LEARN, a uniform step size of $\alpha = \frac{1}{\sqrt{T}}$ was utilized. Mokhtari et al. [2016] suggest a constant step size $\frac{1}{\gamma}$ for baselines where $\frac{1}{\gamma} \leq \frac{1}{L}$ with L being the uniform upper bound on the Lipschitz constants for f_ts. We note that $L = 2\rho = 200$ for our case and choosing $\gamma = \sqrt{T}$ satisfies the prescribed property for baselines.

Results The experiments were conducted over 30 independent runs, with the results shown in Figure 7 and compared against baseline methods. Additionally, Figure 8 highlights the performance of LEARN across varying outlier counts.

In our example, we observe $V_T = \mathcal{O}(\sqrt{T})$ leading to a dynamic environment. In such a dynamic environment without corrupted rounds, the baselines perform well. In this scenario, LEARN initially progresses slowly, incurring higher regret, but eventually stabilizes with a flatter regret curve (Figure 7a). However, as the number of outliers increases, the regret curve for OGD deteriorates significantly and diverges (Figures 7b, 7c). The top-k filter algorithm demonstrates some robustness to increasing outliers, albeit at the cost of higher cumulative regret. However, it also fails to maintain stability when the proportion of outliers becomes constant (Figure 7d). In contrast, LEARN outperforms the baselines as the number of outliers increases. Notably, Figure 8 confirms that LEARN retains its shape of the regret curve and delivers robust performance even in the presence of a constant proportion of outliers.



(a) $k = 0$ (b) $k = \sqrt{T}$

(c) $k = T^{\frac{2}{3}}$ (d) $k = \frac{T}{4}$

Figure 8: Clean Dynamic Regret for only LEARN algorithm with a Sequence of Quadratic Loss Functions in the Presence of Varying Number of Outliers (k). The plots show that even in a dynamic environment LEARN's performance is robust with respect to the number of outliers.

J.5 EFFECT OF STEP SIZE

Working with m-strongly convex functions, one might consider using a step size of $\alpha = \frac{1}{mt}$ for baseline methods. In a static setting with a bounded domain, this choice can yield an upper bound of $\mathcal{O}\left(\frac{\log T}{m}\right)$. The logarithmic upper bounds from the static case do not extend to the dynamic setting (see Theorem B.1). Additionally, the step size $\alpha = \frac{1}{mt}$, borrowed from static settings, lacks theoretical justification for application in dynamic environments. Besides, several other factors make this step size unsuitable for our experiments. First, our experiments are conducted in an unbounded domain, where the baselines lack theoretical guarantees. Moreover, the value of m in our experiments is very small ($m = \lambda = 10^{-4}$), leading to an excessively large step size. As a result, the wishful regret bound of $\mathcal{O}(\frac{\log T}{m})$ becomes impractically large.

We demonstrate this by conducting additional numerical experiments[2] using a modified step size for the Online SVM. For comparison, we also include results using the step size $\alpha_t = \frac{1}{\sqrt{T}}$. The experimental setup remains consistent with Section J.2, and results are averaged over 30 independent runs. The tables presented below show the clean regret in various rounds (first column) across different values of k (second row).

Performance of LEARN For comparison, we report the performance of LEARN algorithm in the new set of experiments in Table 2.

Performance of OGD Table 3 shows the performance of OGD under two different step sizes. We see that the clean dynamic regret of OGD is significantly large when using the step size $\alpha_t = \frac{1}{mt}$.

Performance of Top-k filter algorithm Table 4 shows the performance of Top-k filter algorithm under two different step sizes. We see that after $k + 1$ initial rounds (until Top-k filter finishes creating the initial filter list) the clean dynamic regret

[2] on newly generated data

Table 2: Average clean dynamic regret for LEARN in the presence of a varying number of outliers k

	$\alpha_t = \frac{1}{\sqrt{T}}$			
$t \downarrow \quad k \rightarrow$	0	$\sqrt{T}$	$T^{\frac{2}{3}}$	$0.25T$
1	0.995	0.995	0.995	0.995
10	7.635	7.635	7.635	6.851
100	30.455	30.455	31.518	34.430
1000	123.584	128.780	150.671	277.845
10000	510.256	647.695	1173.455	2595.375

Table 3: Average clean dynamic regret for OGD with different step sizes in the presence of a varying number of outliers k

	$\alpha_t = \frac{1}{\sqrt{T}}$				$\alpha_t = \frac{1}{mt}$			
t, k	0	$\sqrt{T}$	$T^{\frac{2}{3}}$	$0.25T$	0	$\sqrt{T}$	$T^{\frac{2}{3}}$	$0.25T$
1	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
10	4.67	4.67	4.67	6.17	2274221.69	2274221.69	2274221.69	2298131.77
100	16.10	16.10	20.90	40.23	2524798.94	2524798.94	2559633.76	2780628.95
1000	68.78	86.41	156.97	449.25	2600361.98	2625348.08	2711570.17	3256749.28
10000	275.65	672.57	1516.33	4221.43	2632377.85	2685959.45	2857902.15	3685032.15

of Top-k becomes significantly large when using the step size $\alpha_t = \frac{1}{mt}$.

Table 4: Average clean dynamic regret for Top-k filter algorithm with different step sizes in the presence of a varying number of outliers k

	$\alpha_t = \frac{1}{\sqrt{T}}$				$\alpha_t = \frac{1}{mt}$			
t, k	0	$\sqrt{T}$	$T^{\frac{2}{3}}$	$0.25T$	0	$\sqrt{T}$	$T^{\frac{2}{3}}$	$0.25T$
1	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
10	4.67	9.91	9.91	7.92	2274221.69	9.91	9.91	7.92
100	16.10	99.04	96.06	79.20	2524798.94	99.04	96.06	79.20
1000	68.78	189.83	535.21	725.91	2600361.98	56646.91	22964.39	725.91
10000	275.65	772.24	1811.41	4503.58	2632377.85	111748.09	141835.60	166198.45

Based on these numerical evaluations, we found that using $\alpha = \frac{1}{\sqrt{T}}$ produced a significantly smaller regret for the baselines. As a result, we adopted this more favorable step size in our experimental framework.

K REMOVING THE DEPENDENCE ON TIME HORIZON

We can remove the dependence on the knowledge of the time horizon by applying the doubling trick and solving the online optimization problem in several epochs. We provide an informal discussion on the doubling procedure:

1. Let $T_0 = 0$ and start with an estimate of time horizon $T_1 = 2$
2. When $t > T_i$, increase $T_{i+1} = 2T_i$
3. In regret analysis, bound the regret across different epochs M_i where M_i contains the rounds $\{T_i + 1, \ldots, T_{i+1}\}$. Clearly, $|M_0| = 2$ and $|M_i| = 2^i, \forall i \in \{1, \ldots, p\}$.

Observe that after T rounds, p would at max be $\mathcal{O}(\log T)$. We define $V_{M_i} = \sum_{t \in M_i} \|\theta_t^* - \theta_{t+1}^*\|$ and k_{M_i} to be the number of corrupted rounds in epoch i. Consequently, based on our existing results in Theorem 3.3, the order of regret bound for any epoch i contains five kind of terms: $\mathcal{O}(\sqrt{V_{M_i}|M_i|}) + \mathcal{O}(k_{M_i}) + \mathcal{O}(V_{M_i}) + \mathcal{O}(\sqrt{|M_i| \log |M_i|}) + \mathcal{O}(1)$. It is easy to see

that after summing it across all the epochs:

$$\sum_{i=0}^{p} \mathcal{O}(\sqrt{V_{M_i}|M_i|}) = \mathcal{O}(\sqrt{V_T T})$$

$$\sum_{i=0}^{p} \mathcal{O}(k_{M_i}) = \mathcal{O}(k)$$

$$\sum_{i=0}^{p} \mathcal{O}(V_{M_i}) = \mathcal{O}(V_T)$$

$$\sum_{i=0}^{p} \mathcal{O}(\sqrt{|M_i| \log |M_i|}) = \tilde{\mathcal{O}}(\sqrt{T \log T})$$

$$\sum_{i=0}^{p} \mathcal{O}(1) = \tilde{\mathcal{O}}(1)$$

Overall, we incur the same order of regret with an additional factor of $\mathcal{O}(\log T)$.

L HANDLING ARBITRARILY LARGE CORRUPTED GRADIENTS IN BOUNDED DOMAIN

The quadratically bounded gradient assumption in Equation (1) places additional restrictions on the norm of the corrupted gradients in a bounded domain. We may avoid this by adding an appropriate filtering mechanism to our algorithm. For instance, van Erven et al. [2021] address arbitrarily large corrupted gradients by applying a filtering mechanism before passing them to a standard online learning algorithm (ALG). Specifically, they construct a Top-k filter that leverages knowledge of the number of corrupted rounds k to ensure that the norm of the gradients passed to ALG is at most $2G$, where G is the bound on the norm of gradients from uncorrupted rounds. Similarly, if our algorithm has access to k, we can incorporate a similar filtering strategy to exclude large corrupted gradients. The introduction of Equation (1) lets us avoid any oracle knowledge of k.

Even when k is unknown, a filtering algorithm can be designed if the values of G, L, and m are provided (even when D is unknown or infinite). Notably, a naive approach of discarding gradients with norms exceeding G is insufficient in our setting, as uncorrupted gradients can also exhibit large norms while adhering to the growth rule specified in Equation (2). Instead, we propose computing $\eta_t \|\nabla f_t(s_t, \theta_t)\|$ for each round t and filtering out rounds where this quantity exceeds $G + \max\left(\frac{mL}{2ab}, \frac{4a^2L}{m^2b}\right)$. This approach ensures that valid uncorrupted gradients satisfying the growth condition are not inadvertently excluded.