Zeroth-order Median Clipping for Non-Smooth Convex Optimization Problems with Heavy-tailed Symmetric Noise

Preprint · February 2024

DOI: 10.48550/arXiv.2402.02461

8 authors, including:

Nikita Kornilov
Moscow Institute of Physics and Technology

Yuriy Dorn
Moscow Institute of Physics and Technology

Aleksandr Lobanov
Moscow Institute of Physics and Technology

Nikolay Kutuzov
Moscow Institute of Physics and Technology

Zeroth-order Median Clipping for Non-Smooth Convex Optimization Problems with Heavy-tailed Symmetric Noise

Nikita Kornilov [1 2] Yuriy Dorn [3 1 4] Aleksandr Lobanov [1 5] Nikolay Kutuzov [1] Innokentiy Shibaev [1 4]
Eduard Gorbunov [6] Alexander Gasnikov [1 5 2] Alexander Nazin [7 1]

Abstract

In this paper, we consider non-smooth convex optimization with a zeroth-order oracle corrupted by symmetric stochastic noise. Unlike the existing high-probability results requiring the noise to have bounded κ-th moment with $\kappa \in (1,2]$, our results allow even heavier noise with any $\kappa > 0$, e.g., the noise distribution can have unbounded 1-st moment. Moreover, our results match the best-known ones for the case of the bounded variance. To achieve this, we use the mini-batched median estimate of the sampled gradient differences, apply gradient clipping to the result, and plug in the final estimate into the accelerated method. We apply this technique to the stochastic multi-armed bandit problem with heavy-tailed distribution of rewards and achieve $O(\sqrt{dT})$ regret by incorporating the additional assumption of noise symmetry.

1. Introduction

During the recent few years, stochastic optimization problems with heavy-tailed noise received a lot of attention and were actively studied by many researchers. In particular, heavy-tailed noise was observed in various problems, such as the training of large language models (Brown et al., 2020; Zhang et al., 2020), generative adversarial networks (Goodfellow et al., 2014; Gorbunov et al., 2022a), finance (Rachev, 2003), and blockchain (Wang et al., 2019). To solve these problems efficiently many algorithms and techniques were proposed. One of the most popular techniques for handling heavy-tailed noise in theory and practice is gradient clipping (Pascanu et al., 2013), see (Gorbunov et al., 2020; Cutkosky & Mehta, 2021; Sadiev et al., 2023; Nguyen et al., 2023; Puchkin et al., 2023) for the recent advances.

[1]MIPT, Moscow, Russia [2]SkolTech, Moscow, Russia [3]MSU AI Institute [4]IITP RAS [5]ISP RAS [6]MBZUAI [7]ICS RAS. Correspondence to: Nikita Kornilov <kornilov.nm@phystech.edu>.

However, most of the mentioned works focus on the gradient-based (first-order) methods. For some problems, such as the bandit optimization problem (Flaxman et al., 2004; Bartlett et al., 2008; Bubeck & Cesa-Bianchi, 2012), only losses are available, and thus, zeroth-order algorithms are required. The only existing works that handle heavy-tailed noise in the case of the zeroth-order optimization are (Kornilov et al., 2023a;b). In particular, Kornilov et al. (2023a;b) obtain high-probability convergence guarantees for non-smooth convex stochastic optimization problems with the noise having bounded κ-th moment for $\kappa \in (1,2]$. These guarantees degenerate when κ tends to 1, and the convergence is not guaranteed for $\kappa = 1$. In the case of first-order methods, this issue is addressed by Puchkin et al. (2023), who consider symmetric (and close to symmetric) heavy-tailed distributions and achieve better complexity guarantees than previous ones. However, the question of the possibility of the extension the results from (Puchkin et al., 2023) to the case of the zeroth-order optimization remains open. In this paper, we address this question.

1.1. Problem setup

We consider a non-smooth convex optimization problem on set $Q \subseteq \mathbb{R}^d$
$$\min_{x \in Q} f(x), \quad (1)$$
where $f : \mathbb{R}^d \to \mathbb{R}$ is M_2-Lipschitz continuous function w.r.t. the Euclidean norm. The optimization is performed only by accessing function evaluations rather than subgradients, i.e., for any of points $x, y \in Q$, a stochastic oracle returns the pair of the scalar values $f(x,\xi), f(y,\xi)$ with the same realization of the stochastic variable ξ, such that
$$f(x,\xi) - f(y,\xi) = f(x) - f(y) + \phi(\xi|x,y),$$
where $\phi(\xi|x,y)$ is the symmetric stochastic noise, whose distribution depends on points x, y. Distribution of $\phi(\xi|x,y)$ is induced by random variable ξ.

Moreover, ϕ is supposed to have heavy tails, i.e., there exists $\kappa > 0$ such that probability density function $p_\phi(u|x,y) \sim \frac{a(x,y)}{1+|u|^{1+\kappa}}$ $(u \in \mathbb{R})$. Previous results on zeroth-order stochastic optimization were based on the assumption of the existence of the finite expectation of the noise ϕ ($\kappa \in (1, 2]$). The assumption on symmetric noise allows us to consider a wider range of problems (e.g., without finite expectation) and achieve a convergence rate as if the noise had a bounded second moment with the correspondingly small number of oracle calls. Although these results cannot be applied in the case of asymmetric noise and the question of their optimality remains open, they dramatically improve existing bounds for the larger family of stochastic noise.

In our work, we consider two types of noise $\phi(\xi|x, y)$ based on the dependency on points x, y. If the distribution of ϕ does not depend on points x, y, then the oracle is called the "one point oracle", since for every point x, we have independent value $f(x, \xi)$. If the distribution of ϕ becomes more concentrated around zero as the distance $\|x - y\|_2$ becomes smaller, then the oracle is called "Lipschitz oracle". This property allows obtaining for two close points x, y better estimation of $f(x) - f(y)$ and, as a consequence, better convergence rate of the proposed algorithms.

We apply the median operator to the batch $\phi(\xi^i|x, y)_{i=1}^{2m+1}$ of size proportional to $\frac{1}{\kappa}$ in order to lighten tails of symmetric ϕ distribution and obtain unbiased oracle for $f(x) - f(y)$ with bounded second moment. Then we plug in this estimate to the existing zeroth-order algorithm. For the most common noise distributions with $\kappa \in [1, 2]$, the number of samples for the median operator allowing computation in parallel is not greater than 10.

1.2. Contributions

For unconstrained convex optimization we propose modified version of ZO-clipped-SSTM (Kornilov et al., 2023b) called ZO-clipped-med-SSTM. This algorithm is accelerated and allow parallel batching. For optimization on convex compact we propose modified version of ZO-clipped-SMD (Kornilov et al., 2023a) called ZO-clipped-med-SMD. This algorithm is not accelerated and considered without batching. Finally, for strongly convex functions we propose restarted versions of the above-mentioned algorithms.

In the Table 1 we provide number of successive iterations required to achieve accuracy ε on function value with probability at least $1 - \beta$ for above-mentioned algorithms.

For each iteration ZO-clip-med-SSTM requires performing in parallel way $\frac{b}{\kappa}$ oracle calls, while ZO-clip-SSTM requires only b calls.

For the constrained optimization the same bounds without acceleration and batching hold true.

For stochastic multi-armed bandit (MAB) problem with heavy tails we propose Clipped-INF-med-SMD algorithm. In Theorem 3 we get $O(\sqrt{Td})$ bound on regret which coincides with lower bound for stochastic MAB with bounded variance on rewards and in general is better then lower bound for heavy-tailed MAB setting. Moreover, this bound holds even in cases, when reward distribution expected value is not defined. To the best of our knowledge, this is the first positive result for the setting.

1.3. Paper organization.

In Section 2, we introduce notations, main assumptions, and a smoothing technique for gradient estimation. Next, in Sections 3, we introduce two novel gradient-free algorithms: ZO-clipped-SSTM and R-ZO-clipped-SSTM for non-smooth stochastic optimization under heavy-tailed noise. Additionally, we present convergence analysis for each algorithm. In Section 4, we introduce the Clipped-INF-med-SMD algorithm for stochastic multi-armed bandits with heavy tails and provide convergence analysis. In Section A.4, we present the results of computational experiments.

2. Preliminaries

Notations. For vector $x \in \mathbb{R}^d$ and $p \in [1, 2]$ we define p-norm as $\|x\|_p \overset{\text{def}}{=} \left(\sum\limits_{i=1}^{d} |x_i|^p\right)^{\frac{1}{p}}$ and its dual norm $\|x\|_q$, where $\frac{1}{p} + \frac{1}{q} = 1$. In the case $q = \infty$, we define $\|x\|_\infty = \max\limits_{i=1,\dots,d} |x_i|$.

Median operator $\text{Median}(\{a_i\}_{i=1}^{2m+1})$ applied to the elements sequence of the odd size $2m + 1, m \in \mathbb{N}$ returns m-th order statistics. We also use short notation for $\max$ operator, i.e. $a \vee b \overset{\text{def}}{=} \max(a, b)$.

We define the Euclidean unit ball $B_2^d \overset{\text{def}}{=} \{x \in \mathbb{R}^d : \|x\|_2 \leq 1\}$, the Euclidean unit sphere $S_2^d \overset{\text{def}}{=} \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$, and the probability simplex $\Delta_+^d \overset{\text{def}}{=} \{x \in \mathbb{R}_+^d : \sum_{i=1}^d x_i = 1\}$. Denote by $[x, y]$ an interval between two fixed endpoints x and y.

Notation $\widetilde{\mathcal{O}}$ hides logarithm factors.

Assumptions. First, we make a convexity and Lipschitz continuity assumptions on optimized function f.

Assumption 1 (Strong convexity) *There exists $\mu \geq 0$ such that function $f : Q \to \mathbb{R}$ is μ-strongly convex on convex set $Q \subseteq \mathbb{R}^d$, i.e.*

$$\begin{aligned} f(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda f(x_1) + (1-\lambda) f(x_2) \\ &- \frac{1}{2}\mu\lambda(1-\lambda)\|x_1 - x_2\|_2^2, \end{aligned}$$

for all $x_1, x_2 \in Q, \lambda \in [0, 1]$. If $\mu = 0$ we can say "convex function" instead of "0-strongly convex function".

Table 1: Number of successive iterations to achieve a function accuracy ε; unconstrained optimization via Lipschitz oracle with bounded κ-th moment. Constants b, M_2' denote the batch size and the Lipschitz constant of the oracle $f(x,\xi)$, respectively.

As 1	ZO-clipped-SSTM (Kornilov et al., 2023b) $\kappa > 1$	ZO-clipped-med-SSTM (this work) $\kappa > 0$, symmetric noise
$\mu = 0$	$\widetilde{\mathcal{O}}\left(\max\left\{\frac{d^{\frac{1}{4}}M_2'}{\varepsilon}, \frac{1}{b}\left(\frac{\sqrt{d}M_2'}{\varepsilon}\right)^{\frac{\kappa}{\kappa-1}}\right\}\right)$	$\widetilde{\mathcal{O}}\left(\max\left\{\frac{d^{\frac{1}{4}}M_2'}{\varepsilon}, \frac{1}{b}\left(\frac{\sqrt{d}M_2'}{\varepsilon}\right)^{2}\right\}\right)$
$\mu > 0$	$\widetilde{\mathcal{O}}\left(\max\left\{\frac{d^{\frac{1}{4}}M_2'}{\varepsilon}, \frac{1}{b}\left(\frac{d(M_2')^2}{\mu\varepsilon}\right)^{\frac{\kappa}{2(\kappa-1)}}\right\}\right)$	$\widetilde{\mathcal{O}}\left(\max\left\{\frac{d^{\frac{1}{4}}M_2'}{\varepsilon}, \frac{1}{b}\frac{d(M_2')^2}{\mu\varepsilon}\right\}\right)$

For a constant $\tau > 0$, let us define an expansion of set Q namely $Q_\tau = Q + \tau B_2^d$, where $+$ stands for Minkowski addition. Further in the paper, we will consider for any point $x \in Q$ its neighborhood $x + \tau B_2^d$, therefore next assumption must hold on larger set Q_τ.

Assumption 2 (Lipschitz continuity) *Function $f : Q \to \mathbb{R}$ is M_2-Lipschitz continuous w.r.t. the Euclidean norm on Q_τ, i.e., for all $x_1, x_2 \in Q_\tau$*

$$|f(x_1) - f(x_2)| \le M_2\|x_1 - x_2\|_2.$$

Next, we make important assumption on symmetric distribution of the noise $\phi(\xi|x,y)$ conditioned by points x, y.

Assumption 3 (Symmetric noise distribution) *For any pair of points $x, y \in Q$ noise $\phi(\xi|x,y)$ has symmetric conditional probability density $p(u|x,y)$, i.e. $p(u|x,y) = p(-u|x,y), \forall u \in \mathbb{R}$.*

We assume that there exist $\kappa > 0$, $\gamma > 0$, and scale function $B(x,y) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, such that $\forall u \in \mathbb{R}$ holds

$$p(u|x,y) \le \frac{\gamma^\kappa \cdot |B(x,y)|^\kappa}{|B(x,y)|^{1+\kappa} + |u|^{1+\kappa}}. \tag{2}$$

We consider two possible oracles

- ***One point oracle**: $\phi(\xi|x,y)$ distribution doesn't depend on points x, y, i.e.*

$$\gamma \cdot B(x,y) \equiv \Delta. \tag{3}$$

- ***Lipschitz oracle**:*

$$|\gamma \cdot B(x,y)| \le \Delta \cdot \|x - y\|_2, \tag{4}$$

  where $\Delta > 0$ is the Lipschitz constant.

For example, in case random variable ξ has Cauchy distribution, then one can use

- One point oracle: $f(x,\xi) = f(x) + \xi_x, f(y,\xi) = f(y) + \xi_y$ with independent ξ_x, ξ_y.
- Lipschitz oracle: $f(x,\boldsymbol{\xi}) = f(x) + \langle \boldsymbol{\xi}, x\rangle, f(y,\boldsymbol{\xi}) = f(y) + \langle \boldsymbol{\xi}, y\rangle$, where $\boldsymbol{\xi}$ is d-dimensional random vector with d independent components ξ_i $(i = 1, \dots, d)$. Oracle gives the same realization of $\boldsymbol{\xi}$ for both x and y.

Remark 1 *In works (Dvinskikh et al., 2022; Kornilov et al., 2023b) different assumption on Lipschitz noise is considered. For any realization of ξ function $f(x,\xi)$ is $M_2'(\xi)$-Lipschitz, i.e. $\forall x, y \in Q$*

$$|f(x,\xi) - f(y,\xi)| \le M_2'(\xi)\|x - y\|_2, \tag{5}$$

and $M_2'(\xi)^\kappa$ has bounded κ-th moment ($\kappa > 1$), i.e. $[M_2']^\kappa \overset{def}{=} \mathbb{E}_\xi[M_2'(\xi)^\kappa] < \infty$.

We emphasize that if Assumption 3 holds with κ then one can find $M_2'(\xi, x, y)$ such that (5) holds for any $1 < \kappa' < \kappa$ with $M_2' = O(M_2 + \Delta)$, where constant in $O(\cdot)$ depends only on κ'.

However, there is slight difference between the assumptions. In our work for Lipschitz oracle we make Assumption 3 on variable $\phi(\xi|x,y)$ with fixed x, y, while in (Kornilov et al., 2023b) authors make assumption on ξ itself and condition x, y by it. For the very same reason, we can not generalize common proof techniques from previous works and in our upper bounds (13) and (16) we obtain not $M_2' = M_2 + \Delta$, but $M_2' = M_2 + d\Delta$.

For more details and intuition behind Assumption 3 we refer to Appendix paragraph A.1.

Randomized smoothing. The main scheme that allows us to develop batch-parallel gradient-free methods for non-smooth convex problems is randomized smoothing (Ermoliev, 1976; Gasnikov et al., 2022b; Nemirovskij & Yudin, 1983; Nesterov & Spokoiny, 2017; Spall, 2005) of a non-smooth function $f(x)$. The smooth approximation to a non-smooth function $f(x)$ is defined as

$$\hat{f}_\tau(x) \overset{\text{def}}{=} \mathbb{E}_{\mathbf{u}}[f(x + \tau\mathbf{u})], \tag{6}$$

where $\mathbf{u} \sim U(B_2^d)$ is a random vector uniformly distributed on the Euclidean unit ball B_2^d.

The next lemma gives estimates for the quality of this approximation. In contrast to $f(x)$, function $\hat{f}_\tau(x)$ is smooth and has several useful properties.

Lemma 1 (Theorem 2.1 (Gasnikov et al., 2022b)) *Let there exist a subset $Q \subseteq \mathbb{R}^d$ and $\tau > 0$ such that Assumptions 1 and 2 hold on Q_τ. Then,*

1. *Function $\hat{f}_\tau(x)$ is μ-strongly convex, Lipschitz with constant M_2 on Q, and satisfies*
$$\sup_{x \in Q} |\hat{f}_\tau(x) - f(x)| \le \tau M_2.$$
2. *Function $\hat{f}_\tau(x)$ is differentiable on Q with the following gradient*
$$\nabla \hat{f}_\tau(x) = \mathbb{E}_{\mathbf{e}}\left[\frac{d}{\tau} f(x + \tau \mathbf{e})\mathbf{e}\right],$$
where $\mathbf{e} \sim U(S_2^d)$ is a random vector uniformly distributed on unit Euclidean sphere.
3. *Function $\hat{f}_\tau(x)$ is L-smooth with $L = \sqrt{d}M_2/\tau$ on Q.*

Our algorithms will aim at minimizing the smooth approximation $\hat{f}_\tau(x)$ with the fixed τ during optimization. Given Lemma 1, the output of the algorithm will also be a good approximate minimizer of $f(x)$ when τ is sufficiently small.

Gradient estimation. Our algorithms will be based on randomized gradient estimate of the function $\hat{f}_\tau(x)$, which will then be used in a first order algorithm. Following (Shamir, 2017), the gradient of $\hat{f}_\tau(x)$ can be estimated by the following vector:

$$g(x, \mathbf{e}, \xi) = \frac{d}{2\tau}(f(x + \tau \mathbf{e}, \xi) - f(x - \tau \mathbf{e}, \xi))\mathbf{e}, \quad (7)$$

where $\tau > 0$ and $\mathbf{e} \sim U(S_2^d)$ is a random vector uniformly distributed on the Euclidean unit sphere S_2^d. Moreover, $\mathbf{e}, \xi$ are independent from each other conditionally on x.

Noise $\phi(\xi | x + \tau \mathbf{e}, x - \tau \mathbf{e})$ might have unbounded expectation, therefore in order to obtain unbiased estimate of $\nabla \hat{f}_\tau(x)$ we lighten tails of ϕ distribution. For this purpose we use component-wise median operator on the batch $\{g(x, \mathbf{e}, \xi^i)\}_{i=1}^{2m+1}$ of $2m+1$ samples with independent ξ^i and the same $x, \mathbf{e}$, i.e.

$$\text{Med}^m(x, \mathbf{e}, \{\xi\}) \overset{\text{def}}{=} \text{Median}(\{g(x, \mathbf{e}, \xi^i)\}_{i=1}^{2m+1}). \quad (8)$$

For large enough m this trick allows $\text{Med}^m(x, \mathbf{e}, \{\xi\})$ to have finite expectation and variance conditionally on x w.r.t. all emerged variables $\xi, \mathbf{e}$.

In addition, we define batched version of the median estimate with smaller second moment in Euclidean norm at point x.

$$\text{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\}) \overset{\text{def}}{=} \frac{1}{b}\sum_{j=1}^{b} \text{Med}^m(x, \mathbf{e}^j, \{\xi\}^j), \quad (9)$$

where $\mathbf{e}^j$ are sampled independently from $U(S_2^d)$.

Lemma 2 *Let function f satisfies Assumptions 1, 2 and symmetric noise ϕ satisfies Assumption 3 with $\kappa > 0$. If median size $m > \frac{2}{\kappa}$ with norm $q \in [2, +\infty]$, then median estimate is unbiased, i.e.*

$$\mathbb{E}_{\mathbf{e},\xi}[\textit{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\}) | x] = \nabla \hat{f}_\tau(x),$$

and has bounded second moment, i.e

$$\mathbb{E}_{\mathbf{e},\xi}[\|\textit{Med}^m(x, \mathbf{e}, \{\xi\}) - \nabla \hat{f}_\tau(x)\|_q^2 | x] \le \sigma^2 a_q^2, \quad (10)$$

$$\mathbb{E}_{\mathbf{e},\xi}[\|\textit{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\}) - \nabla \hat{f}_\tau(x)\|_2^2 | x] \le \frac{\sigma^2}{b},$$

where $a_q = d^{\frac{1}{q} - \frac{1}{2}} \min\{\sqrt{32 \ln d - 8}, \sqrt{2q - 1}\}$ and for

- ***One point oracle**:*
$$\sigma^2 = 8dM_2^2 + 2\left(\frac{d\Delta}{\tau}\right)^2 (2m+1)\left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}},$$
- ***Lipschitz oracle**:*
$$\sigma^2 = 8dM_2^2 + (16m + 8)d^2\Delta^2\left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}}.$$

Proof of Lemma 2 can be found in Appendix paragraph A.2.

3. Zeroth-order algorithms for non-smooth optimization problems

In each algorithm below we use clipping technique which clips tails of gradient estimate distribution and ensures the algorithm convergence. Let $\lambda > 0$ be clipping constant, parameter $q \in [2, +\infty]$ for $q-$norm and $g \in \mathbb{R}^d$, then clipping operator `clip` is defined as

$$\texttt{clip}_q(g, \lambda) = \begin{cases} \frac{g}{\|g\|_q} \min\left(\|g\|_q, \lambda\right), & g \neq 0, \\ 0, & g = 0. \end{cases} \quad (11)$$

Fixed during minimization smoothing parameter τ can be chosen arbitrary. However, its optimal value can be calculated directly based on desired accuracy ε.

Then we feed the oracle vector $\text{BatchMed}^{m,b}(x,\{\mathbf{e}\},\{\xi\})$ that satisfies inequalities from Lemma 2 into various clipped first-order methods which minimize $L = \sqrt{d}M_2/\tau$ smooth function $\hat{f}_\tau(x)$ on set Q.

3.1. ZO-clipped-med-SSTM for unconstrained problems

Let us suppose that the function $f : \mathbb{R}^d \to \mathbb{R}$ is convex, i.e., Assumption 1 is satisfied with $\mu = 0$.

We use first-order Clipped Stochastic Similar Triangles Method (clipped-SSTM) from (Gorbunov et al., 2020) and it's zeroth order version ZO-clipped-SSTM from (Kornilov et al., 2023b) on the whole space $Q = \mathbb{R}^d$ and Euclidean norm $q = 2$.

Algorithm 1 ZO-clipped-med-SSTM

Input: Starting point $x^0 \in \mathbb{R}^d$, number of iterations K, median size m, batch size b, stepsize $a > 0$, smoothing parameter τ, clipping levels $\{\lambda_k\}_{k=0}^{K-1}$.

1: Set $L = \sqrt{d}M_2/\tau$, $\quad A_0 = \alpha_0 = 0, \; y^0 = z^0 = x^0$.
2: **for** $k = 0, \dots, K-1$ **do**
3: Set $\alpha_{k+1} = (k+2)/2aL$, $A_{k+1} = A_k + \alpha_{k+1}$.
4: $x^{k+1} = \frac{A_k y^k + \alpha_{k+1} z^k}{A_{k+1}}$.
5: Sample independently sequences $\{\mathbf{e}\} \sim U(S_2^d)$ and $\{\xi\}$.
6: $g_{med}^{k+1} = \text{BatchMed}^{m,b}(x^{k+1},\{\mathbf{e}\},\{\xi\})$ from (9).
7: $\tilde{g}_{med}^{k+1} = \texttt{clip}_2\left(g_{med}^{k+1}, \lambda_{k+1}\right)$.
8: $z^{k+1} = z^k - \alpha_{k+1}\tilde{g}_{med}^{k+1}$.
9: $y^{k+1} = \frac{A_k y^k + \alpha_{k+1} z^{k+1}}{A_{k+1}}$.
10: **end for**

Output: y^K

Theorem 1 (Convergence of ZO-clipped-med-SSTM) *Let for the function $f : \mathbb{R}^d \to \mathbb{R}$ Assumptions 1, 2 hold with $\mu = 0$ on $Q = \mathbb{R}^d$ and for symmetric noise Assumption 3 holds with $\kappa > 0$.*

We use notation $\|x^0 - x^*\|_2^2 \le R^2$, where x^0 is a starting point and x^* is an optimal solution to (1).

Suppose we run ZO-clipped-med-SSTM for K iterations with smoothing parameter τ, batchsize b, probability $1-\beta$ and further parameters $m = \frac{2}{\kappa} + 1, A = \ln{}^{4K}/_{\beta} \ge 1$, $a = \Theta(\min\{A^2, \sigma K^2\sqrt{A}\tau/\sqrt{bd}M_2R\}), \lambda_k = \Theta(R/(\alpha_{k+1}A))$. Then, with probability at least $1-\beta$, holds true

$$f(y^k) - f(x^*) = 2M_2\tau + \widetilde{\mathcal{O}}\left(\max\left\{\frac{\sqrt{d}M_2R^2}{\tau K^2}, \frac{\sigma R}{\sqrt{bK}}\right\}\right),$$

where σ comes from Lemma 2.

Moreover, with probability at least $1-\beta$ the iterates of ZO-clipped-med-SSTM remain in the Euclidean ball with center x^* and radius $2R$, i.e., $\{x^k\}_{k=0}^{K+1}, \{y^k\}_{k=0}^{K}, \{z^k\}_{k=0}^{K} \subseteq \{x \in \mathbb{R}^d : \|x - x^*\|_2 \le 2R\}$.

Proof of Theorem 1 can be found in Appendix paragraph A.3.

Corollary 1 *In order to achieve accuracy ε, i.e. $f(y^k) - f(x^*) \le \varepsilon$ via ZO-clipped-med-SSTM with probability at least $1-\beta$, the smoothing parameter τ must be chosen as $\tau = \frac{\varepsilon}{4M_2}$ and number of iterations K must be equal*

- ***One point oracle**:*

$$\widetilde{\mathcal{O}}\left(\frac{d^{\frac{1}{4}}M_2R}{\varepsilon} \vee \frac{(\sqrt{d}M_2R)^2}{b \cdot \varepsilon^2}\left(1 \vee \left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}}\frac{d\Delta^2}{\varepsilon^2}\right)\right), \tag{12}$$

- ***Lipschitz oracle**:*

$$\widetilde{\mathcal{O}}\left(\max\left\{\frac{d^{\frac{1}{4}}M_2R}{\varepsilon}, \frac{d(M_2^2 + d\Delta^2/\kappa^{\frac{2}{\kappa}})R^2}{b \cdot \varepsilon^2}\right\}\right). \tag{13}$$

For each iteration one requires $(2m+1)\cdot b = \Theta(\frac{b}{\kappa})$ number of oracle calls.

The first term in bound (13) is optimal in ε for the deterministic case for non-smooth problem (see (Bubeck et al., 2019)) and the second term in bound (13) is optimal in ε for zeroth-order problems with finite variance (see (Nemirovskij & Yudin, 1983)).

In case of one-point oracle, while noise ϕ is "small", i.e.,

$$\Delta \le \left(\frac{\kappa}{4}\right)^{\frac{1}{\kappa}}\frac{\varepsilon}{\sqrt{d}} \tag{14}$$

optimal convergence rate is preserved. This bound is optimal in terms ε upper bound under which convergence is optimal (Lobanov, 2023; Pasechnyuk et al., 2023; Risteski & Li, 2016).

Numerical experiments comparing ZO-clipped-med-SSTM and ZO-clipped-SSTM are located in Appendix paragraph A.4.

Remark 2 *In this section, as in the whole paper, we consider the case where the objective function is non-smooth. However, the estimates presented in Corollary 1 can be improved by introducing a new assumption, namely the assumption that the objective function $f(\cdot)$ is L-smooth with $L > 0$: $\|\nabla f(y) - \nabla f(x)\|_2 \le L\|y - x\|_2$, $\forall x, y \in Q$. Using this assumption we obtain the following value of the smoothing parameter $\tau = \sqrt{\varepsilon/L}$ (see Gasnikov et al., 2022a, the end of Section 4.1). Thus, assuming smoothness and convexity of the function and assuming symmetric noise (Assumption 3), we obtain the following estimates for the iteration complexity: $\widetilde{\mathcal{O}}\left(\max\left\{\sqrt{\frac{LR^2}{\varepsilon}}, \frac{(\sqrt{d}R)^2}{b\cdot\varepsilon^2}\left(M_2^2 \vee \left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}}\frac{dL\Delta^2}{\varepsilon}\right)\right\}\right)$ and $\widetilde{\mathcal{O}}\left(\max\left\{\sqrt{\frac{LR^2}{\varepsilon}}, \frac{d(M_2^2+d\Delta^2/\kappa^{\frac{2}{\kappa}})R^2}{b\cdot\varepsilon^2}\right\}\right)$ for one point oracle and Lipschitz oracle, respectively.*

Remark 3 *The results of Theorem 1 can be extended to the case when the function satisfies the Polyak–Lojasiewicz condition: let a function $f(x)$ is differentiable and there exists constant $\mu > 0$ s.t. $\forall x \in Q$ the following inequality holds $\|\nabla f(x)\|_2^2 \ge 2\mu(f(x) - f(x^*))$. The, assuming smoothness (see Remark 2) and Polyak–Lojasiewicz condition for the function and assuming symmetric noise (Assumption 3), we obtain the following estimates for the iteration complexity: $\widetilde{\mathcal{O}}\left(\max\left\{\frac{L}{\mu}, \frac{dL}{b\mu^2\varepsilon}\left(M_2^2 \vee \left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}}\frac{dL\Delta^2}{\varepsilon}\right)\right\}\right)$ and $\widetilde{\mathcal{O}}\left(\max\left\{\frac{L}{\mu}, \frac{dL(M_2^2+d\Delta^2/\kappa^{\frac{2}{\kappa}})}{b\mu^2\varepsilon}\right\}\right)$ for one point oracle and Lipschitz oracle, respectively.*

For $\mu-$strongly-convex functions (under $\mu > 0$) with Lipschitz oracle or one-point oracle with small noise (14) we apply restarted version of ZO-Clipped-med-SSTM called R-ZO-Clipped-med-SSTM. More details and convergence results are located in Appendix paragraph A.5 and Corollary 3.

3.2. ZO-clipped-med-SMD for constrained problems

To solve convex optimization problems on convex compact $Q \subset \mathbb{R}^d$ we use oracle Med^b in the zeroth-order algorithm ZO-clipped-SMD from (Kornilov et al., 2023a) based on Mirror Gradient Descent.

Let us introduce for this section 1-strongly convex w.r.t. p-norm differentiable prox-function Ψ_p. We denote its Fenchel conjugate and its Bregman divergence respectively as

$$\Psi_p^*(y) = \sup_{x\in\mathbb{R}^d}\{\langle x, y\rangle - \Psi_p(x)\},$$
$$V_{\Psi_p}(y, x) = \Psi_p(y) - \Psi_p(x) - \langle\nabla\Psi_p(x), y - x\rangle.$$

Theorem 2 (Convergence of ZO-clipped-med-SMD) *Let for the function $f(\cdot)$ Assumptions 1, 2 hold with $\mu = 0$ on convex compact Q and for symmetric noise Assumption 3 holds with $\kappa > 0$.*

Denote x^* as an optimal solution to (1).

Suppose we run ZO-clipped-med-SMD for K iterations with smoothing parameter τ, $q-$norm with $q \in [2, +\infty]$, prox-function Ψ_p, probability $1 - \beta$ and further parameters $m = \frac{2}{\kappa} + 1$, $\lambda = \sigma a_q\sqrt{K}$, $\nu = \frac{D_{\Psi_p}}{\lambda}$, where diameter squared $D_{\Psi_p}^2 \overset{def}{=} 2\sup_{x,y\in Q} V_{\Psi_p}(x, y)$. We guarantee that with probability at least $1 - \beta$

$$f(y^k) - f(x^*) = 2M_2\tau + \widetilde{\mathcal{O}}\left(\frac{\sigma a_q D_{\Psi_p}}{\sqrt{K}}\right),$$

where σ comes from Lemma 2.

Proof of Theorem 2 can be found in Appendix paragraph A.3.

Algorithm 2 ZO-clipped-med-SMD

Input: Number of iterations K, median size m, stepsize ν, prox-function Ψ_p, smoothing parameter τ, clipping level λ.

1: $x_0 = \arg\min_{x\in Q}\Psi_p(x)$.
2: **for** $k = 0, 1, \ldots, K - 1$ **do**
3: Sample $\mathbf{e}$ from $U(S_2^d)$ and sequence $\{\xi\}$.
4: $g_{med}^{k+1} = \text{Med}^m(x^{k+1}, \mathbf{e}, \{\xi\})$ from (9).
5: $\tilde{g}_{med}^{k+1} = \texttt{clip}_q\left(g_{med}^{k+1}, \lambda\right)$.
6: $y_{k+1} = \nabla(\Psi_p^*)(\nabla\Psi_p(x_k) - \nu\tilde{g}_{med}^{k+1})$.
7: $x_{k+1} = \arg\min_{x\in Q} V_{\Psi_p}(x, y_{k+1})$.
8: **end for**

Output: $\frac{1}{T}\sum_{k=0}^{T-1} x_k$

Corollary 2 *In order to achieve accuracy ε, i.e. $f(y^k) - f(x^*) \le \varepsilon$ via ZO-clipped-med-SMD with probability at least $1 - \beta$ smoothing parameter τ must be chosen as $\tau = \frac{\varepsilon}{4M_2}$ and number of iterations K must be equal for*

- ***One point oracle**:*

$$\widetilde{\mathcal{O}}\left(\frac{(\sqrt{d}M_2 a_q D_{\Psi_p})^2}{\varepsilon^2}\left(1 \vee \left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}}\frac{d\Delta^2}{\varepsilon^2}\right)\right), \quad (15)$$

- ***Lipchitz oracle**:*

$$\widetilde{\mathcal{O}}\left(\frac{d(M_2^2 + d\Delta^2/\kappa^{\frac{2}{\kappa}})a_q^2 D_{\Psi_p}^2}{\varepsilon^2}\right). \quad (16)$$

For each iteration one requires $(2m + 1)$ *number of oracle calls.*

Bound (16) is optimal in terms of ε for stochastic non-smooth optimization on convex compact with finite variance according to (Vural et al., 2022). In case of one-point oracle, optimal convergence rate is preserved under the same upper bound (14).

Next, we discuss some standard sets Q and prox-functions Ψ_p taken from (Ben-Tal & Nemirovski, 2001). We can choose prox-functions to reduce $a_q D_{\Psi_p}$ and get better convergence constants. The two main setups are given by

1. Ball setup, $p = 2, q = 2$:

$$\Psi_p(x) = \frac{1}{2}\|x\|_2^2,$$

2. Entropy setup, $p = 1, q = \infty$:

$$\Psi_p(x) = (1+\gamma)\sum_{i=1}^{d}(x_i + \gamma/d)\log(x_i + \gamma/d).$$

We consider unit balls $B_{p'}^d$ and standard simplex $\triangle_+^d$ as Q. For $Q = \triangle_+^d$ or B_1^d, the Entropy setup is preferable. Meanwhile, for $Q = B_2^d$ or B_∞^d, the Ball setup is better.

For strongly-convex functions with Lipschitz oracle or one-point oracle with small noise (14) we apply restarted version of ZO-Clipped-SMD called R-ZO-Clipped-SMD. More details and convergence results are located in Appendix paragraph A.5 and Corollary 4.

4. Application to the multi-armed bandit problem with heavy tails

In this section, we present the Clipped-INF-med-SMD algorithm for multi-armed bandit problem with heavy-tailed rewards.

The stochastic multi-armed bandit problem (MAB) with d arms and horizon T can be viewed as an online optimization problem with regret compared to a fixed competitor strategy u defined as

$$\mathbb{E}[\mathcal{R}_T(u)] = \mathbb{E}\left[\sum_{t=1}^{T} l(x_t) - \sum_{t=1}^{T} l(u)\right],$$

with linear loss function $l(x_t) = \langle \mu + \xi_t, x_t\rangle$, with noise ξ_t, $\mathbb{E}[\xi_t] = 0$ and unknown fixed vector of expected rewards $\mu \in \mathbb{R}^d$. Heavy noise assumption usually require that there is $\kappa \in (1, 2]$, such that $\mathbb{E}[\|\mu + \xi_t\|^\kappa] \le \sigma^\kappa$. Decision variable $x_t \in \Delta_+^d$ can be viewed as player's mixed strategy (probability distribution over arms), that he use to sample arms with the aim to maximize expected reward. Loss function is the expected reward (with minus) conditioned by x_t, but the player observe only sampled reward for the chosen arm, i.e. the (sub)gradient $g(x) \in \partial l(x)$ is not observed in the MAB setting, and one must use inexact oracle instead.

Bandits with heavy tails were first introduced in (Bubeck et al., 2013) along with lower bounds on regret $\Omega\left(Md^{\frac{\kappa-1}{\kappa}}T^{\frac{1}{\kappa}}\right)$ and nearly optimal algorithmic scheme Robust UCB. Recently few optimal algorithms were proposed (Lee et al., 2020; Zimmert & Seldin, 2019; Huang et al., 2022; Dorn et al., 2024) with online mirror descent (OMD) as the main ingredient. Currently, these types of algorithms are referred to as best-of-two-worlds algorithms. In this work we consider $HTINF$ and APE as the main benchmark.

In this section we assume that the noise ξ satisfy Assumption 3 for some positive κ and follow the same path and construct Clipped-INF-med-SMD based on online mirror descent, but we show that in case of symmetric noise we can improve regret upper bounds and make it $O(\sqrt{Td})$ which corresponds to lower bound for stochastic MAB with bounded variance of rewards.

In our algorithm we use an importance-weighted estimator:

$$\bar{g}_{t,i} = \begin{cases} \frac{g_{t,i}}{x_{t,i}} & \text{if } i = A_t \\ 0 & \text{otherwise} \end{cases},$$

where A_t is the index of the chosen (at round t) arm.

This estimator is unbiased, i.e. $\mathbb{E}_{x_t}[\bar{g}_t] = g_t$. The main drawback of this estimator is that in the case of small $x_{t,i}$ the value of $\bar{g}_{t,i}$ can be arbitrarily large. In the case when the distribution of the noise $g_t - \mu$ has heavy tails (i.e. $\|g_t - \mu\|_\infty$ can be large with high probability), this drawback can be amplified.

Theorem 3 *Let there exists* $\kappa > 0$ *such that conditional probability density function for each reward satisfies Assumption 3, and* $R^2 \ge \|\mu\|_2^2$. *Then for period* T *the sequence* $\{x_t\}_{t=1}^T$ *generated by* Clipped-INF-med-SMD *with parameters* $K = \frac{T}{2m+1}, \nu = \frac{\sqrt{2(2m+1)}d^{1/4}\sqrt{18\sigma^2+R^2}}{\sqrt{T}}, \lambda = \sigma\sqrt{\frac{T}{2m+1}}$, $m = \frac{2}{\kappa} + 1$, *prox-function* $\Psi_p = \psi(x) = 2(1 - \sum_{i=1}^n x_i^{\frac{1}{2}})$, $p = 2, q = 2$ *satisfies for any fixed* $u \in \Delta_+^d$*:*

$$\frac{1}{T}\mathbb{E}\left[\mathcal{R}_T(u)\right] \le 2\sqrt{2}(2m+1)^{1/2}T^{-1/2}\gamma, \qquad (17)$$

with $\gamma = 2\sigma + d^{1/4}\sqrt{18\sigma^2 + R^2}$, *and high probability bounds from Theorem 2 hold.*

Proof is located in Appendix paragraph A.6.

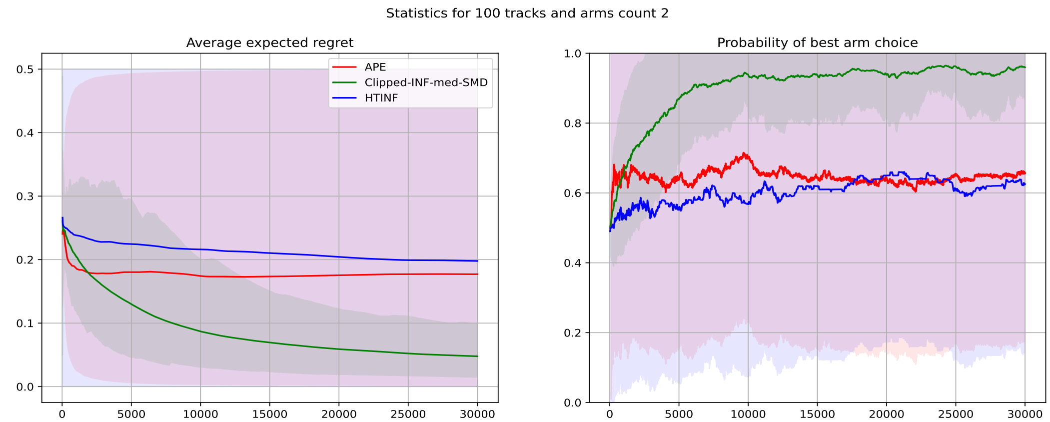

Figure 1: Average expected regret and probability of optimal arm picking mean (aggregated with average filter) for 100 experiments and 30000 samples with 0.95 and 0.05 percentiles for regret and ± std bounds for probabilities

Algorithm 3 Clipped-INF-med-SMD

Input: Time period T, median size m, number of iterations $K = \left\lceil \frac{T-1}{2m+1} \right\rceil$, stepsize ν, prox-function Ψ_p, clipping level λ.

$x_0 = \arg\min_{x \in \Delta_+^d} \Psi_p(x)$.

for $k = 0, 1, \ldots, K$ **do**

Sample $\mathbf{e}$ from $U(S_2^d)$.

Draw A_t for $2m+1$ times ($t = (2m+1)\cdot k + 1, \ldots, (2m+1)\cdot(k+1)$) with $P(A_t = i) = x_{k,i}$, $i = 1, \ldots, d$ and observe rewards g_{t,A_t}.

For each observation construct estimation $\hat{g}_{t,i} = \begin{cases} \frac{g_{t,i}}{x_{k,i}} & \text{if } i = A_t \\ 0 & \text{otherwise} \end{cases}$, $i = 1, \ldots, d$.

$g_{med}^{k+1} = \text{Median}(\{\hat{g}_t\}_{t=(2m+1)\cdot k+1}^{(2m+1)\cdot(k+1)})$.

$\tilde{g}_{med}^{k+1} = \texttt{clip}_q\left(g_{med}^{k+1}, \lambda\right)$.

$y_{k+1} = \nabla(\Psi_p^*)(\nabla\Psi_p(x_k) - \nu\tilde{g}_{med}^{k+1})$.

$x_{k+1} = \arg\min_{x \in \Delta_+^d} V_{\Psi_p}(x, y_{k+1})$.

end for

Output: $\frac{1}{T}\sum_{k=0}^{T-1} x_k$

Note that σ^2 depends linearly on d. Same holds true for $R^2 \simeq \|\mu\|_2^2 \le d \cdot \|\mu\|_\infty^2$. Thus $\frac{1}{T}\mathbb{E}\left[\mathcal{R}_T(u)\right] \le O\left(\sqrt{\frac{d}{T}}\right)$ which corresponds to lower bound for stochastic MAB with bounded variance on reward and in general is better then lower bound for MAB with heavy tails.

For non-linear loss function $l(x)$ we apply ZO-clipped-med-SMD with one-point oracle. We refer to Remark 7 in Appendix paragraph A.6 for more details.

5. Numerical Experiments

We conducted experiments to demonstrate the superior performance of our Clipped-INF-med-SMD algorithm in specific stochastic Multi-Armed Bandit (MAB) scenarios with heavy tails when compared to HTINF and APE. To showcase this, we focus on an experiment involving only two available arms ($d = 2$). Each arm i generates random losses $g_{t,i} \sim \xi_t + \beta_i$. Here parameters $\beta_0 = 3, \beta_1 = 3.5$ are fixed and independent random variables ξ_t have the same $\text{pdf}_{\xi_t}(x) = \frac{1}{3\cdot\left(1+\left(\frac{x}{3}\right)^2\right)\cdot\pi}$.

In this experimental setup, individual experiments are subject to significant random deviations. To enhance the informativeness of the results, we conduct 100 individual experiments and analyze aggregated statistics.

By design, we possess knowledge of the conditional probability of selecting the optimal arm for all algorithms, which remains stochastic due to the nature of the experiment's history.

To mitigate the high dispersion in probabilities, we apply an average filter with a window size of 30 to reduce noise in the plot. APE and HTINF can't handle cases when noise expectation is unbounded, so we modeled this case with a low value of $\alpha = 0.01$, where $1 + \alpha$ is the moment that exists in the problem statement for APE and HTINF.

As can be seen from the graphs, HTINF and APE do not have convergence in probability, while Clipped-INF-med-SMD does, which confirms the efficiency of the proposed method.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

References

Bartlett, P., Dani, V., Hayes, T., Kakade, S., Rakhlin, A., and Tewari, A. High-probability regret bounds for bandit online linear optimization. In *Proceedings of the 21st Annual Conference on Learning Theory-COLT 2008*, pp. 335–342. Omnipress, 2008.

Ben-Tal, A. and Nemirovski, A. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, Philadelphia, 2001.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33: 1877–1901, 2020.

Bubeck, S. and Cesa-Bianchi, N. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012. ISSN 1935-8237. doi: 10.1561/2200000024. URL http://dx.doi.org/10.1561/2200000024.

Bubeck, S., Cesa-Bianchi, N., and Lugosi, G. Bandits with heavy tail. *IEEE Transactions on Information Theory*, 59 (11):7711–7717, 2013.

Bubeck, S., Jiang, Q., Lee, Y.-T., Li, Y., and Sidford, A. Complexity of highly parallel non-smooth convex optimization. *Advances in neural information processing systems*, 32, 2019.

Cutkosky, A. and Mehta, H. High-probability bounds for non-convex stochastic optimization with heavy tails. *Advances in Neural Information Processing Systems*, 34, 2021.

Dorn, Y., Kornilov, N., Kutuzov, N., Nazin, A., Gorbunov, E., and Gasnikov, A. Implicitly normalized forecaster with clipping for linear and non-linear heavy-tailed multi-armed bandits. *Comput Manag Sci*, 21(19), 2024.

Dvinskikh, D., Tominin, V., Tominin, I., and Gasnikov, A. Noisy zeroth-order optimization for non-smooth saddle point problems. In *Mathematical Optimization Theory and Operations Research: 21st International Conference, MOTOR 2022, Petrozavodsk, Russia, July 2–6, 2022, Proceedings*, pp. 18–33. Springer, 2022.

Ermoliev, Y. Stochastic programming methods, 1976.

Flaxman, A. D., Kalai, A. T., and McMahan, H. B. Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*, 2004.

Gasnikov, A., Dvinskikh, D., Dvurechensky, P., Gorbunov, E., Beznosikov, A., and Lobanov, A. Randomized gradient-free methods in convex optimization. *arXiv preprint arXiv:2211.13566*, 2022a.

Gasnikov, A., Novitskii, A., Novitskii, V., Abdukhakimov, F., Kamzolov, D., Beznosikov, A., Takac, M., Dvurechensky, P., and Gu, B. The power of first-order smooth optimization for black-box non-smooth problems. In *International Conference on Machine Learning*, pp. 7241–7265. PMLR, 2022b.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

Gorbunov, E., Danilova, M., and Gasnikov, A. Stochastic optimization with heavy-tailed noise via accelerated gradient clipping. *Advances in Neural Information Processing Systems*, 33:15042–15053, 2020.

Gorbunov, E., Danilova, M., Dobre, D., Dvurechenskii, P., Gasnikov, A., and Gidel, G. Clipped stochastic methods for variational inequalities with heavy-tailed noise. *Advances in Neural Information Processing Systems*, 35: 31319–31332, 2022a.

Gorbunov, E., Dvurechensky, P., and Gasnikov, A. An accelerated method for derivative-free smooth stochastic convex optimization. *SIAM Journal on Optimization*, 32 (2):1210–1238, 2022b.

Huang, J., Dai, Y., and Huang, L. Adaptive best-of-both-worlds algorithm for heavy-tailed multi-armed bandits. In *International Conference on Machine Learning*, pp. 9173–9200. PMLR, 2022.

Kornilov, N., Gasnikov, A., Dvurechensky, P., and Dvinskikh, D. Gradient-free methods for non-smooth convex stochastic optimization with heavy-tailed noise on convex compact. *Computational Management Science*, 20(1):37, 2023a.

Kornilov, N., Shamir, O., Lobanov, A., Dvinskikh, D., Gasnikov, A., Shibaev, I., Gorbunov, E., and Horváth, S. Accelerated zeroth-order method for non-smooth stochastic convex optimization problem with infinite variance. *arXiv preprint arXiv:2310.18763*, 2023b.

Lee, K., Yang, H., Lim, S., and Oh, S. Optimal algorithms for stochastic multi-armed bandits with heavy tailed rewards. *Advances in Neural Information Processing Systems*, 33:8452–8462, 2020.

Lobanov, A. Stochastic adversarial noise in the "black box" optimization problem. In *Optimization and Applications*, pp. 60–71, Cham, 2023. Springer Nature Switzerland. ISBN 978-3-031-47859-8.

Nemirovskij, A. S. and Yudin, D. B. *Problem complexity and method efficiency in optimization*. J. Wiley @ Sons, 1983.

Nesterov, Y. and Spokoiny, V. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17:527–566, 2017.

Nguyen, T. D., Ene, A., and Nguyen, H. L. Improved convergence in high probability of clipped gradient methods with heavy tails. *arXiv preprint arXiv:2304.01119*, 2023.

Pascanu, R., Mikolov, T., and Bengio, Y. On the difficulty of training recurrent neural networks. In *International conference on machine learning*, pp. 1310–1318, 2013.

Pasechnyuk, D. A., Lobanov, A., and Gasnikov, A. Upper bounds on maximum admissible noise in zeroth-order optimisation. *arXiv preprint arXiv:2306.16371*, 2023.

Puchkin, N., Gorbunov, E., Kutuzov, N., and Gasnikov, A. Breaking the heavy-tailed noise barrier in stochastic optimization problems. *arXiv preprint arXiv:2311.04161*, 2023.

Rachev, S. T. *Handbook of heavy tailed distributions in finance: Handbooks in finance, Book 1*. Elsevier, 2003.

Risteski, A. and Li, Y. Algorithms and matching lower bounds for approximately-convex optimization. *Advances in Neural Information Processing Systems*, 29, 2016.

Sadiev, A., Danilova, M., Gorbunov, E., Horváth, S., Gidel, G., Dvurechensky, P., Gasnikov, A., and Richtárik, P. High-probability bounds for stochastic optimization and variational inequalities: the case of unbounded variance. *arXiv preprint arXiv:2302.00999*, 2023.

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

Spall, J. C. *Introduction to stochastic search and optimization: estimation, simulation, and control*. John Wiley & Sons, 2005.

Vural, N. M., Yu, L., Balasubramanian, K., Volgushev, S., and Erdogdu, M. A. Mirror descent strikes again: Optimal stochastic convex optimization under infinite noise variance. In *Conference on Learning Theory*, pp. 65–102. PMLR, 2022.

Wang, P., Xu, H., Jin, X., and Wang, T. Flash: efficient dynamic routing for offchain networks. In *Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies*, pp. 370–381, 2019.

Zhang, J., Karimireddy, S. P., Veit, A., Kim, S., Reddi, S., Kumar, S., and Sra, S. Why are adaptive methods good for attention models? *Advances in Neural Information Processing Systems*, 33:15383–15393, 2020.

Zimmert, J. and Seldin, Y. An optimal algorithm for stochastic and adversarial bandits. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 467–475. PMLR, 2019.

A. Appendix

A.1. Remarks about the assumption on the noise.

Remark 4 (Standard oracles examples) *To build noise $\phi(\xi|x,y)$ satisfying Assumption 3 with $\kappa > 0$ we will use independent random variables $\{\xi_k\}$ with symmetric probability density functions $p_{\xi_k}(u)$*

$$p_{\xi_k}(u) \leq \frac{|\gamma_k \Delta_k|^\kappa}{|\Delta_k|^{1+\kappa} + |u|^{1+\kappa}}, \quad \Delta_k, \gamma_k > 0,$$

such that for any real numbers $\{a_k\}_{k=1}^n$ and sum $\sum\limits_{k=1}^n a_k \xi_k$ it holds

$$p_{\sum\limits_{k=1}^n a_k\xi_k}(u) \leq \frac{\left(\sum\limits_{k=1}^n |\gamma_k a_k \Delta_k|\right)^\kappa}{\left(\sum\limits_{k=1}^n |a_k\Delta_k|\right)^{1+\kappa} + |u|^{1+\kappa}}. \tag{18}$$

Moreover, using Cauchy-Schwarz inequality we bound

$$\sum_{k=1}^n |\gamma_k a_k \Delta_k| \leq \|(\gamma_1\Delta_1, \dots, \gamma_n\Delta_n)^\top\|_2 \cdot \|(a_1, \dots, a_k)^\top\|_2. \tag{19}$$

For example, ξ_k could have Cauchy distribution with $\kappa = 1$ and $p(u) = \frac{1}{\pi}\frac{\Delta_k}{\Delta_k^2 + u^2}$ parametrized by scale Δ_k. For the independent Cauchy variables with scales $\{\Delta_k\}_{k=1}^n$ and any real numbers $\{a_k\}_{k=1}^n$ sum $\sum\limits_{k=1}^n a_k\xi_k$ is the Cauchy variable with scale $\sum\limits_{k=1}^n |a_k|\Delta_k$. Therefore, inequality (18) for Cauchy variables holds true.

- ***One point oracle:***

  $f(x,\xi) = f(x) + \xi_x$, $f(y,\xi) = f(y) + \xi_y$, $\phi(\xi|x,y) = \xi_x - \xi_y$, where ξ_x, ξ_y are independent samples for each point x and y.

- ***Lipschitz oracle:***

  $f(x,\boldsymbol{\xi}) = f(x) + \langle \boldsymbol{\xi}, x\rangle$, $f(y,\boldsymbol{\xi}) = f(y) + \langle \boldsymbol{\xi}, y\rangle$, $\phi(\boldsymbol{\xi}|x,y) = \langle \boldsymbol{\xi}, x-y\rangle$, where $\boldsymbol{\xi}$ is d-dimensional random vector with independent components ξ_k. Oracle gives the same realization of $\boldsymbol{\xi}$ for both x and y.

  In that case, we have γ and $B(x,y)$ from Assumption 3 equal to

$$\begin{aligned} a_k &= x_k - y_k, \\ \gamma &= \max_{k=1,\dots,d} \gamma_k, \\ B(x,y) &= \sum_{k=1}^d |\Delta_k (x-y)_k| \overset{(19)}{\leq} \|(\Delta_1, \dots, \Delta_d)^\top\|_2 ||x-y||_2. \end{aligned}$$

Remark 5 (Relation between two definitions of M_2') *Let noise $\phi(\xi|x,y)$ satisfies Assumption 3 with Lipschitz oracle and $\kappa > 1$, then it holds*

$$\begin{aligned} |f(x,\xi) - f(y,\xi)| &= |f(x) - f(y) + \phi(\xi|x,y)| \\ &\leq |f(x) - f(y)| + |\phi(\xi|x,y)| \\ &\overset{\text{As } 2}{\leq} M_2\|x-y\|_2 \\ &+ \frac{|\phi(\xi|x,y)|}{\|x-y\|_2}\|x-y\|_2. \end{aligned}$$

Let us denote $M_2'(\xi,x,y) \overset{def}{=} M_2 + \frac{|\phi(\xi|x,y)|}{\|x-y\|_2}$ *and show that for any* $1<\kappa'<\kappa$ *random variable* $M_2'(\xi,x,y)$ *has bounded* κ'*-th moment which doesn't depend on* x,y*. We notice that*

$$\begin{aligned}\mathbb{E}_\xi[|\phi(\xi|x,y)|^{\kappa'}] &= \int\limits_{-\infty}^{+\infty} |u|^{\kappa'} p(u|x,y)du \\ &\le \int\limits_{-\infty}^{+\infty} \frac{|u|^{\kappa'}\gamma^\kappa|B(x,y)|^\kappa}{|B(x,y)|^{1+\kappa}+|u|^{1+\kappa}}du.\end{aligned}$$

After substitution $t = {}^u/_{|B(x,y)|}$ *we get*

$$\begin{aligned}\mathbb{E}_\xi[|\phi(\xi|x,y)|^{\kappa'}] &\le \frac{\gamma^\kappa|B(x,y)|^\kappa}{|B(x,y)|^{\kappa-\kappa'}}\int\limits_{0}^{+\infty}\frac{|t|^{\kappa'}}{1+|t|^{1+\kappa}}dt \\ &\overset{(4)}{\le} \gamma^{\kappa-\kappa'}\Delta^{\kappa'}\|x-y\|_2^{\kappa'}\int\limits_{0}^{+\infty}\frac{|t|^{\kappa'}}{1+|t|^{1+\kappa}}dt.\end{aligned}$$

Integral $I(\kappa') = \int\limits_0^{+\infty}\frac{\gamma^{\kappa-\kappa'}|t|^{\kappa'}dt}{1+|t|^{1+\kappa}}$ *converges since* $\kappa'<\kappa$ *but its value tends to* ∞ *as* $\kappa'\to\kappa-0$*. Finally, we have*

$$\begin{aligned}&\mathbb{E}_\xi[M_2'(\xi,x,y)^{\kappa'}] \\ =\;& \mathbb{E}_\xi\left[\left|M_2+\frac{|\phi(\xi|x,y)|}{\|x-y\|_2}\right|^{\kappa'}\right] \\ \overset{\text{Jensen inq, } \kappa'>1}{\le}\;& 2^{\kappa'-1}\left[M_2^{\kappa'}+\frac{\mathbb{E}_\xi\left[|\phi(\xi|x,y)|^{\kappa'}\right]}{\|x-y\|_2^{\kappa'}}\right] \\ \le\;& 2^{\kappa'-1}\left[M_2^{\kappa'}+I(\kappa')\Delta^{\kappa'}\right].\end{aligned}$$

Therefore, $M_2' = (\mathbb{E}_\xi[M_2'(\xi,x,y)^{\kappa'}])^{\frac{1}{\kappa'}} = O(M_2+\Delta)$*, where constant in* $O(\cdot)$ *depends only on* κ'*.*

Remark 6 (Role of the scale function $B(x,y)$**)** *In inequality (2) due to normalization property of probability density we must ensure that*

$$\int\limits_{-\infty}^{+\infty}\frac{\gamma^\kappa|B(x,y)|^\kappa}{|B(x,y)|^{1+\kappa}+|u|^{1+\kappa}}du \ge \int\limits_{-\infty}^{+\infty}p(u|x,y)du = 1.$$

One can make substitution $t = u/|B(x,y)|$ *and ensure that for* $\kappa\le 2$

$$\int\limits_{-\infty}^{+\infty}\frac{\gamma^\kappa|B(x,y)|^\kappa du}{|B(x,y)|^{1+\kappa}+|u|^{1+\kappa}} = \gamma^\kappa\int\limits_{-\infty}^{+\infty}\frac{dt}{1+|t|^{1+\kappa}} \overset{\kappa=1}{\ge} \gamma^\kappa\pi.$$

Hence, γ *is sufficient to satisfy*

$$\gamma \ge \left(\frac{1}{\pi}\right)^{\frac{1}{\kappa}}.$$

As scale value $|B(x,y)|$ *decreases, quantiles of* $p(u|x,y)$ *gets closer to zero. Therefore,* $|B(x,y)|$ *can be considered as analog of variance of distribution* $p(u|x,y)$*.*

A.2. Proof of Lemma 2.

First, we notice from our construction of the oracle

$$f(x,\xi) - f(y,\xi) = f(x) - f(y) + \phi(\xi|x,y), \quad \forall x,y \in Q,$$

that

$$\begin{aligned} g(x,\mathbf{e},\xi) &= \frac{d}{2\tau}(f(x+\tau\mathbf{e},\xi) - f(x-\tau\mathbf{e},\xi)) \\ &= \frac{d}{2\tau}[f(x+\tau\mathbf{e}) - f(x-\tau\mathbf{e})]\mathbf{e} + \frac{d}{2\tau}\phi(\xi|x+\tau\mathbf{e}, x-\tau\mathbf{e})\mathbf{e} \end{aligned}$$

and for $\text{Med}(x,\mathbf{e},\{\xi\})$ we have

$$\begin{aligned} & \text{Med}(x,\mathbf{e},\{\xi\}) = \text{Median}\left(\left\{g(x,\mathbf{e},\xi^i)\right\}_{i=1}^{2m+1}\right) \\ = \; & \text{Median}\left(\left\{\frac{d}{2\tau}[f(x+\tau\mathbf{e}) - f(x-\tau\mathbf{e})]\mathbf{e} + \frac{d}{2\tau}\phi(\xi^i|x+\tau\mathbf{e}, x-\tau\mathbf{e})\mathbf{e}\right\}_{i=1}^{2m+1}\right) \\ = \; & \frac{d}{2\tau}[f(x+\tau\mathbf{e}) - f(x-\tau\mathbf{e})]\mathbf{e} \qquad (20) \\ + \; & \frac{d}{2\tau}\text{Median}\left(\left\{\phi(\xi^i|x+\tau\mathbf{e}, x-\tau\mathbf{e})\right\}_{i=1}^{2m+1}\right)\mathbf{e}. \qquad (21) \end{aligned}$$

Finite second moment:

Further, we analyze two terms: gradient estimation term (20) and the noise term (21).

Following work (Kornilov et al., 2023a) [Lemma 2.3.] we have an upper bound for the second moment of (20)

$$\mathbb{E}_{\mathbf{e}}\left[\left\|\frac{d}{2\tau}[f(x+\tau\mathbf{e}) - f(x-\tau\mathbf{e})]\mathbf{e}\right\|_q^2\right] \leq d a_q^2 M_2^2, \tag{22}$$

where $a_q = d^{\frac{1}{q}-\frac{1}{2}} \min\{\sqrt{32\ln d - 8}, \sqrt{2q-1}\}$ is a special coefficient, such that,

$$\mathbb{E}_{\mathbf{e}}[\|\mathbf{e}\|_q^2] \leq a_q^2. \tag{23}$$

See Lemma 2.1 from (Gorbunov et al., 2022b) and Lemma 8.4 from (Kornilov et al., 2023a) for more details.

Next, we deal with noise term (21). For symmetric variable $\phi(\xi|x,y)$ for all $x,y \in Q$ under Assumption 3 it holds

$$p(u) \leq \frac{\gamma^\kappa |B(x,y)|^\kappa}{|B(x,y)|^{1+\kappa} + |u|^{1+\kappa}}.$$

Further, we prove that for large enough m noise term has finite variance. For this purpose we denote $Y \stackrel{\text{def}}{=} \text{Median}\left(\left\{\phi(\xi^i|x,y)\right\}_{i=1}^{2m+1}\right)$ and cumulative distribution function of Y

$$P(t) \stackrel{\text{def}}{=} \int\limits_{-\infty}^{t} p(u)du.$$

Median of $2m+1$ i.i.d. variables distributed according to $p(u)$ is $(m+1)$-th order statistic, which has probability density function

$$(2m+1)\binom{2m}{m} P(t)^m (1-P(t))^m p(t).$$

The second moment $\mathbb{E}[Y^2]$ can be calculated via

$$\begin{aligned}
\mathbb{E}[Y^2] &= \int\limits_{-\infty}^{+\infty} (2m+1)\binom{2m}{m} t^2 P(t)^m (1-P(t))^m p(t)dt \\
&\le (2m+1)\binom{2m}{m} \sup_t \{t^2 P(t)^m (1-P(t))^m\} \int\limits_{-\infty}^{+\infty} p(t)dt \\
&\le (2m+1)\binom{2m}{m} \sup_t \{t^2 P(t)^m (1-P(t))^m\}.
\end{aligned}$$

For any $t < 0$ we have

$$\begin{aligned}
P(t) &= \int\limits_{-\infty}^{t} p(u)du \le \int\limits_{-\infty}^{t} \frac{|\gamma B(x,y)|^\kappa}{|B(x,y)|^{1+\kappa} + |u|^{1+\kappa}} \\
&\le \int\limits_{-\infty}^{t} \frac{|\gamma B(x,y)|^\kappa}{|u|^{1+\kappa}} \le \frac{|\gamma B(x,y)|^\kappa}{\kappa} \cdot \frac{1}{|t|^\kappa}.
\end{aligned}$$

Similarly, one can prove that for any $t > 0$

$$1 - P(t) = \int\limits_{t}^{\infty} p(u)du \le \frac{|\gamma B(x,y)|^\kappa}{\kappa} \cdot \frac{1}{t^\kappa}.$$

Since for any number $a \in [0,1]$ holds $a(1-a) \le \frac{1}{4}$ we have for any $t \in \mathbb{R}$

$$P(t)(1-P(t)) \le \min\left\{\frac{1}{4}, \frac{|\gamma B(x,y)|^\kappa}{\kappa} \cdot \frac{1}{|t|^\kappa}\right\}$$

along with

$$t^2 P(t)^m (1-P(t))^m \le \min\left\{\frac{t^2}{4^m}, \left(\frac{|\gamma B(x,y)|^\kappa}{\kappa}\right)^m \cdot \frac{1}{|t|^{m\kappa-2}}\right\}. \tag{24}$$

If $m\kappa > 2$ first term of (24) increasing and the second one decreasing with the growth of $|t|$, then the maximum of the minimum (24) is achieved when

$$\begin{aligned}
\frac{t^2}{4^m} &= \left(\frac{|\gamma B(x,y)|^\kappa}{\kappa}\right)^m \cdot \frac{1}{|t|^{m\kappa-2}}, \\
|t| &= |\gamma B(x,y)| \left(\frac{4}{\kappa}\right)^{\frac{1}{\kappa}}.
\end{aligned}$$

Therefore, we get for any $t \in \mathbb{R}$

$$t^2 P(t)^m (1-P(t))^m \le \frac{|\gamma B(x,y)|^2}{4^m} \left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}},$$

and, as a consequence

$$\mathbb{E}[Y^2] \le (2m+1)\binom{2m}{m} \frac{|\gamma B(x,y)|^2}{4^m} \left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}}.$$

It only remains to note

$$\binom{2m}{m} = \frac{(2m)!}{m! \cdot m!} = \prod_{j=1}^{m} \frac{2j}{j} \cdot \prod_{j=1}^{m} \frac{2j-1}{j} \leq 4^m.$$

Since Y has the finite second moment, it has finite math expectation

$$\mathbb{E}[Y] = \int_{-\infty}^{+\infty} (2m+1) \binom{2m}{m} tP(t)^m (1-P(t))^m p(t) dt.$$

For any $t \in \mathbb{R}$ due to symmetry of $p(t)$ we have $P(t) = (1 - P(-t))$ and $p(t) = p(-t)$ and, as a consequence,

$$\mathbb{E}[Y] = \int_{-\infty}^{+\infty} (2m+1) \binom{2m}{m} tP(t)^m (1-P(t))^m p(t) dt = 0.$$

Finally, we have an upper bound for (21)

$$\begin{aligned}
& \mathbb{E}_{\mathbf{e},\xi} \left|\left| \frac{d}{2\tau} \text{Med}\big(\{\phi(\xi^i | x + \tau\mathbf{e}, x - \tau\mathbf{e})\}\big)\mathbf{e} \right|\right|_q^2 \\
= & \left(\frac{d}{2\tau}\right)^2 \mathbb{E}_{\mathbf{e}}[\mathbb{E}_\xi[Y^2|\mathbf{e}] \cdot \|\mathbf{e}\|_q^2] \\
\leq & \left(\frac{d}{2\tau}\right)^2 (2m+1) \left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}} \cdot \mathbb{E}_{\mathbf{e}}[|\gamma B(x+\tau\mathbf{e}, x-\tau\mathbf{e})|^2 \|\mathbf{e}\|_q^2].
\end{aligned} \tag{25}$$

In case of **one-point oracle** from Assumption 3 and (3) we simplify

$$\mathbb{E}_{\mathbf{e}}[|\gamma B(x+\tau\mathbf{e}, x-\tau\mathbf{e})| \|\mathbf{e}\|_q^2] \leq \Delta^2 \mathbb{E}_{\mathbf{e}}[\|\mathbf{e}\|_q^2] \overset{(23)}{\leq} \Delta^2 a_q^2. \tag{26}$$

In case of **Lipschitz oracle** we use (4) and get

$$\begin{aligned}
\mathbb{E}_{\mathbf{e}}[|\gamma B(x+\tau\mathbf{e}, x-\tau\mathbf{e})| \|\mathbf{e}\|_q^2] &\leq 4\Delta^2\tau^2 \mathbb{E}_{\mathbf{e}}[\|\mathbf{e}\|_2^2 \|\mathbf{e}\|_q^2] \\
&\overset{(23)}{\leq} 4\Delta^2\tau^2 a_q^2.
\end{aligned} \tag{27}$$

Combining upper bounds (22) and (26) or (27) we obtain total bound

$$\mathbb{E}_{\mathbf{e},\xi}[\|\text{Med}(x, \mathbf{e}, \{\xi\})\|_q^2] \quad \leq \quad 2 \cdot (22) + 2 \cdot (26)/(27).$$

For the batched gradient estimation $\text{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\})$ and $q = 2$ we use Lemma 4 from (Kornilov et al., 2023b), that states

$$\mathbb{E}_{\mathbf{e},\xi}[\|\text{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\})\|_2^2] \leq \frac{1}{b} \cdot \mathbb{E}_{\mathbf{e},\xi}[\|\text{Med}(x, \mathbf{e}, \{\xi\})\|_2^2].$$

For the bound of the centered second moment we use Jensen inequality for any random vector X

$$\mathbb{E}[||X - \mathbb{E}[X]||_q^2] \leq 2\mathbb{E}[||X||_q^2] + 2||\mathbb{E}[X]||_q^2 \leq 4\mathbb{E}[||X||_q^2].$$

Unbiasedness:

According to Lemma 1 term (20) is unbiased estimation of the gradient $\nabla \hat{f}^\tau(x)$. Indeed, the distribution of $\mathbf{e}$ is symmetrical and we can derive

$$\mathbb{E}_{\mathbf{e}}\left[\frac{d}{2\tau}[f(x+\tau\mathbf{e}) - f(x-\tau\mathbf{e})]\mathbf{e}\right] = \mathbb{E}_{\mathbf{e}}\left[\frac{d}{\tau}[f(x+\tau\mathbf{e})]\right] = \nabla \hat{f}^\tau(x).$$

Since Y has the finite second moment, it has finite math expectation

$$\mathbb{E}[Y] = \int\limits_{-\infty}^{+\infty} (2m+1)\binom{2m}{m} tP(t)^m(1-P(t))^m p(t)dt.$$

For any $t \in \mathbb{R}$ due to symmetry of $p(t)$ we have $P(t) = (1 - P(-t))$ and $p(t) = p(-t)$ and, as a consequence,

$$\mathbb{E}[Y] = \int\limits_{-\infty}^{+\infty} (2m+1)\binom{2m}{m} tP(t)^m(1-P(t))^m p(t)dt = 0.$$

Hence, we obtained, that $\mathbb{E}_{\mathbf{e},\xi}[\text{Med}(x, \mathbf{e}, \{\xi\})] = \nabla \hat{f}_\tau(x)$ along with $\mathbb{E}_{\mathbf{e},\xi}[\text{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\})] = \nabla \hat{f}_\tau(x)$ as the batching is the mean of random vectors with the same math expectation.

A.3. Proof of Convergence Theorems 1 and 2.

We might consider $\text{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\})$ to be the oracle for the gradient of $\hat{f}_\tau(x)$ that satisfies Assumption 4.

Assumption 4 *Let $G(x, \mathbf{e}, \xi)$ be the oracle for the gradient of function $\hat{f}_\tau(x)$, such that for any point $x \in Q$ it is unbiased, i.e.*

$$\mathbb{E}_{\mathbf{e},\xi}[G(x, \mathbf{e}, \xi)] = \nabla \hat{f}_\tau(x),$$

and has bounded second moment, i.e.

$$\mathbb{E}_{\mathbf{e},\xi}[\|G(x, \mathbf{e}, \xi) - \nabla \hat{f}_\tau(x)\|_q^2] \le \Sigma_q^2, \tag{28}$$

where Σ_q might depend on τ.

Thus, in order to prove convergence of ZO-clipped-med-SSTM and ZO-clipped-med-SMD we use general convergence theorems with oracle satisfying Assumption 4 for ZO-clipped-SSTM (Theorem 1 from (Kornilov et al., 2023b) with $\alpha = 2$) and ZO-clipped-SMD (Theorem 4.3 from (Kornilov et al., 2023a) with $\kappa = 1$).

Next, we take $\text{BatchMed}^{m,b}(x, \{\mathbf{e}\}, \{\xi\})$ $\text{Med}^m(x, \mathbf{e}, \{\xi\})$ as the necessary oracles and substitute Σ_q from (28) with σ or σa_q from Lemma 2. But each call of median operator requires $(2m+1)$ calls of zeroth-order oracle $\delta(x, \xi)$.

Theorem 4 (Convergence of ZO-clipped-SSTM) *Let for the function f Assumptions 1, 2 hold with $\mu = 0$ on $Q = \mathbb{R}^d$ and for oracle Assumption 4 holds with Σ_2.*

Let be $\|x^0 - x^*\|_2^2 \le R^2$, where x^0 is a starting point and x^* is an optimal solution to (1).

Suppose ZO-clipped-SSTM is run for K iterations with smoothing parameter τ and batch size b, probability $1-\beta$ and further parameters $A = \ln {}^{4K}/_{\beta} \ge 1$, $a = \Theta(\min\{A^2, {}^{\Sigma_2 K^2\sqrt{A}\tau}/_{\sqrt{db}M_2R}\})$, $\lambda_k = \Theta({}^{R}/_{(\alpha_{k+1}A)})$. We guarantee that with probability at least $1-\beta$

$$f(y^k) - f(x^*) = 2M_2\tau + \widetilde{\mathcal{O}}\left(\max\left\{\frac{\sqrt{d}M_2R^2}{\tau K^2}, \frac{\Sigma_2 R}{\sqrt{bK}}\right\}\right).$$

Moreover, with probability at least $1-\beta$ the iterates of ZO-clipped-SSTM remain in the Euclidean ball with center x^* and radius $2R$, i.e., $\{x^k\}_{k=0}^{K+1}, \{y^k\}_{k=0}^{K}, \{z^k\}_{k=0}^{K} \subseteq \{x \in \mathbb{R}^d : \|x - x^*\|_2 \le 2R\}$.

Theorem 5 (Convergence of ZO-clipped-SMD) *Let for the function f Assumptions 1, 2 hold with $\mu = 0$ on convex compact Q and for oracle Assumption 4 holds with Σ_q.*

We use notation x^* as an optimal solution to (1).

Suppose we run ZO-clipped-SMD for K iterations with smoothing parameter τ, norm $q \in [2, +\infty]$, prox-function Ψ_p, probability $1-\beta$ and further parameters $\lambda = \Sigma_q \sqrt{K}$, $\nu = \frac{D_{\Psi_p}}{\lambda}$, where squared diameter $D^2_{\Psi_p} \overset{def}{=} 2 \sup\limits_{x,y \in Q} V_{\Psi_p}(x,y)$. We guarantee that with probability at least $1-\beta$

$$f(y^k) - f(x^*) = 2M_2\tau + \widetilde{\mathcal{O}}\left(\frac{\Sigma_q D_{\Psi_p}}{\sqrt{K}}\right).$$

A.4. Additional Numerical Experiments

Following (Kornilov et al., 2023b) we conducted experiments on the following problem

$$\min_{x \in \mathbb{R}^d} \|Ax - b\|_2 + \langle \xi, x \rangle,$$

where ξ is a random vector with independent components sampled from the symmetric Levy α-stable distribution with $\alpha = {}^3/_2$, $A \in \mathbb{R}^{l \times d}$, $b \in \mathbb{R}^l$ (we used $d = 16$ and $l = 200$). For ZO-clipped-med-SSTM we used median size $m = 3$. Figure 2 presents the comparison of convergences averaged over 9 launches with different noise, and we see ZO-clipped-med-SSTM outperforming ZO-clipped-SSTM.

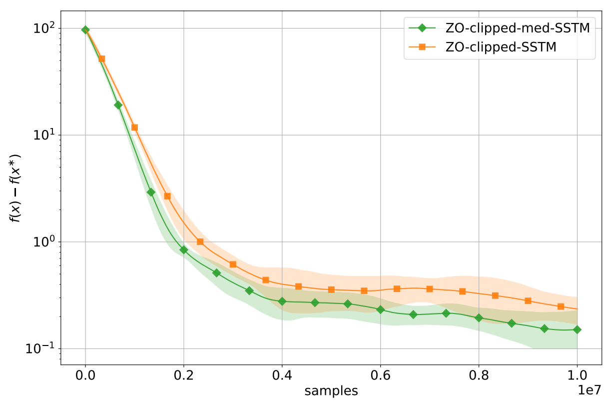

Figure 2: Convergence of ZO-clipped-SSTM and ZO-clipped-med-SSTM in terms of a gap function w.r.t. the number of consumed samples from the dataset

A.5. Restarted algorithms R-ZO-clipped-SSTM and R-ZO-clipped-SMD.

The restart technique is to run in cycle algorithm $\mathcal{A}$ taking the output point from the previous run as the initial point for the current one.

Algorithm 4 R-ZO-clipped-$\mathcal{A}$

Input: Starting point x^0, number of restarts N_r, number of iterations $\{K_t\}_{t=1}^{N_r}$, algorithm $\mathcal{A}$, parameters $\{P_t\}_{t=1}^{N_r}$.
1: $\hat{x}^0 = x^0$.
2: **for** $t = 1, \ldots, N_r$ **do**
3: Run algorithms $\mathcal{A}$ with parameters P_t and starting point $\hat{x}^{t-1}$. Set output point as $\hat{x}^t$.
4: **end for**
Output: $\hat{x}^{N_r}$

Strong convexity of function f with minimum x^* implies upper bound for the distance between point x and solution as

$$\frac{\mu}{2}\|x - x^*\|_2^2 \leq f(x) - f(x^*).$$

Considering upper bounds from Corollary 1, 2 for $f(x) - f(x^*)$ one can construct a relation between $\|x_0 - x^*\|_2$ and $\|x - x^*\|_2$ after K iterations. Based on this relation one can calculate iteration after which it is more efficient to start a new run rather than continue with slow convergence rate current one.

We apply the general Convergence Theorem 2 from (Kornilov et al., 2023b) for R-ZO-clipped-SSTM and Theorem 5.2 from (Kornilov et al., 2023a) for R-ZO-clipped-SMD with oracle satisfying Assumption 4. However, oracle couldn't depend on, τ which means that we consider either Lipschitz oracle or one-point oracle with small noise, i.e.,

$$\Delta \leq \left(\frac{\kappa}{4}\right)^{\frac{1}{\kappa}} \frac{\varepsilon}{\sqrt{d}}. \tag{29}$$

In the Convergence Theorems minimal necessary value of $\tau = \frac{\varepsilon}{4M_2}$, hence

$$\begin{aligned} \sigma^2 &= 8dM_2^2 + 2\left(\frac{d\Delta}{\tau}\right)^2 (2m+1)\left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}} \\ &\leq 32(2m+1)\cdot dM_2^2. \end{aligned}$$

Theorem 6 (Convergence of R-ZO-clipped-SSTM) *Let for the function $f(x)$ Assumptions 1, 2 hold with $\mu > 0$ on $Q = \mathbb{R}^d$ and for oracle Assumption 4 holds with Σ_2.*

Let be $\|x^0 - x^*\|^2 \leq R^2$, where x^0 is a starting point and x^* is the optimal solution to (1).

Let ε be desired accuracy, $1-\beta$ be desired probability and $N_r = \lceil \log_2(\mu R^2/2\varepsilon) \rceil$ be the number of restarts. Suppose at each stage $t = 1, ..., N_r$ ZO-clipped-SSTM is run with batch size b_t, $\tau_t = \varepsilon_t/4M_2$, $L_t = M_2\sqrt{d}/\tau_t$, $K_t = \widetilde{\Theta}(\max\{\sqrt{L_t R_{t-1}^2/\varepsilon_t}, (\Sigma_2 R_{t-1}/\varepsilon_t)^2/b_t\})$, $a_t = \widetilde{\Theta}(\max\{1, \Sigma_2 K_t^{\frac{3}{2}}/\sqrt{b_t} L_t R_t\})$ and $\lambda_k^t = \widetilde{\Theta}(R/\alpha_{k+1}^t)$, where $R_{t-1} = 2^{-\frac{(t-1)}{2}}R$, $\varepsilon_t = \mu R_{t-1}^2/4$, $\ln 4N_rK_t/\beta \geq 1$, $\beta \in (0,1]$. Then to guarantee $f(\hat{x}^{N_r}) - f(x^*) \leq \varepsilon$ with probability at least $1-\beta$, R-ZO-clipped-SSTM *requires*

$$\widetilde{\mathcal{O}}\left(\max\left\{\sqrt{\frac{M_2^2\sqrt{d}}{\mu\varepsilon}}, \frac{\Sigma_2^2}{\mu\varepsilon}\right\}\right) \tag{30}$$

total number of oracle calls.

Corollary 3 *Let Assumption 3 holds with $\kappa > 0$ for Lipschitz oracle.*

In order to achieve accuracy ε, i.e. $f(\hat{x}^{N_r}) - f(x^*) \leq \varepsilon$ via R-ZO-clipped-med-SSTM *with probability at least $1-\beta$ median size must be $m = \frac{2}{\kappa} + 1$, other parameters must be set according to Theorem 6 ($\Sigma_2 = \sigma$ from Lemma 2). In that case* R-ZO-clipped-med-SSTM *requires for*

- ***One point oracle under (29):***

$$\widetilde{\mathcal{O}}\left((2m+1)\cdot\max\left\{\sqrt{\frac{M_2^2\sqrt{d}}{\mu\varepsilon}}, \frac{dM_2^2}{\kappa\mu\varepsilon}\right\}\right), \tag{31}$$

- ***Lipschitz oracle:***

$$\widetilde{\mathcal{O}}\left((2m+1)\cdot\max\left\{\sqrt{\frac{M_2^2\sqrt{d}}{\mu\varepsilon}}, \frac{d(M_2^2 + d\Delta^2/\kappa^{\frac{2}{\kappa}})}{\mu\varepsilon}\right\}\right) \tag{32}$$

total number of oracle calls.

Similar to the convex case, the first term in bounds (32), (31) is optimal in ε for the deterministic case for non-smooth strongly convex problems (see (Bubeck et al., 2019)) and the second term in is optimal in ε for zeroth-order problems with finite variance (see (Nemirovskij & Yudin, 1983)).

Theorem 7 *Let for the function f Assumptions 1, 2 hold with $\mu > 0$ and for oracle Assumption 4 holds with Σ_2 on convex compact Q.*

We set the prox-function Ψ_p and norm $p \in [1, 2]$. Denote $R_0^2 \overset{def}{=} \sup_{x,y\in Q} 2V_{\Psi_p}(x, y)$ for diameter of set Q and $R_t = R_0/2^t$. Let ε be desired accuracy and $N = \widetilde{O}\left(\frac{1}{2}\log_2\left(\frac{\mu R_0^2}{2\varepsilon}\right)\right)$ be the number of restarts. Suppose at each stage $t = 1, \ldots, N_r$ ZO-clipped-SMD *is run with $K_t = \widetilde{O}\left(\left[\frac{\Sigma_q}{\mu R_t}\right]^2\right)$, $\tau_t = \frac{\Sigma_q R_t}{M_2\sqrt{K_t}}$, $\lambda_t = \sqrt{K_t}\Sigma_q$ and $\nu_t = \frac{R_t}{\lambda_t}$.*

Then to guarantee $f(\hat{x}^{N_r}) - f(x^*) \le \varepsilon$ with probability at least $1-\beta$, R-ZO-clipped-SMD *requires*

$$\widetilde{O}\left(\frac{\Sigma_q^2}{\mu\varepsilon}\right)$$

total number of oracle calls.

Corollary 4 *Let Assumption 3 holds with $\kappa > 0$ for Lipschitz oracle.*

In order to achieve accuracy ε, i.e. $f(\hat{x}^{N_r}) - f(x^*) \le \varepsilon$ via R-ZO-clipped-med-SMD *with probability at least $1-\beta$ median size must be $m = \frac{2}{\kappa} + 1$, other parameters must be set according to Theorem 7 ($\Sigma_q = \sigma a_q$ from Lemma 2). In that case* R-ZO-clipped-med-SMD *requires for*

- ***One point oracle under (29):***

$$\widetilde{\mathcal{O}}\left((2m+1)\cdot\frac{dM_2^2 a_q^2}{\kappa\mu\varepsilon}\right), \tag{33}$$

- ***Lipschitz oracle:***

$$\widetilde{\mathcal{O}}\left((2m+1)\cdot\frac{d(M_2^2 + d\Delta^2/\kappa^{\frac{2}{\kappa}})a_q^2}{\mu\varepsilon}\right) \tag{34}$$

total number of oracle calls, where $a_q = d^{\frac{1}{q}-\frac{1}{2}}\min\{\sqrt{32\ln d - 8}, \sqrt{2q-1}\}$.

A.6. Proof of Theorem 3

Lemma 3 *Let $f(x)$ be a linear function and random vector $\mathbf{u}$ satisfies $\mathbb{E}_{\mathbf{u}}[u] = 0$, then $\nabla f(x) = \nabla \hat{f}_\tau(x)$.*

Proof:

$$\begin{aligned}\nabla \hat{f}_\tau(x) &= \nabla\mathbb{E}_{\mathbf{u}}[f(x+\tau\mathbf{u})] = \nabla\mathbb{E}_{\mathbf{u}}[\langle\mu, x+\tau\mathbf{u}\rangle] \\ &= \nabla\langle\mu, x+\tau\mathbb{E}_{\mathbf{u}}[u]\rangle = \nabla\langle\mu, x\rangle = \nabla f(x).\end{aligned}$$

Lemma 4 *Suppose that* Clipped-INF-med-SMD *with $1/2$-Tsallis entropy*

$$\psi(x) = 2\left(1 - \sum_{i=1}^d x_i^{1/2}\right), \quad x \in \Delta_+^d$$

as prox-function generates the sequences $\{x_k\}_{k=0}^{K}$ *and* $\{\tilde{g}^k_{med}\}_{k=0}^{K}$*, then for any* $u \in \Delta^d_+$ *holds:*

$$\sum_{k=0}^{K}\sum_{s=1}^{2m+1}\langle \tilde{g}^k_{med}, x_k - u\rangle$$
$$\leq (2m+1)\left[2\frac{d^{1/2} - \sum_{i=1}^{d} u_i^{1/2}}{\nu} + \nu\sum_{k=0}^{K}\sum_{i=1}^{d}(\langle \tilde{g}^k_{med})_i^2 \cdot x_{k,i}^{3/2}\right].$$

Proof:

By definition the *Bregman divergence* $V_\psi(x,y)$ is:

$$V_\psi(x,y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y\rangle$$
$$= 2\left(1 - \sum_{i=1}^{d} x_i^{1/2}\right) - 2\left(1 - \sum_{i=1}^{d} y_i^{1/2}\right) + \sum_{i=1}^{d} y_i^{-1/2}(x_i - y_i)$$
$$= -2\sum_{i=1}^{d} x_i^{1/2} + 2\sum_{i=1}^{d} y_i^{1/2} + \sum_{i=1}^{d} y_i^{-1/2}(x_i - y_i).$$

Note that the algorithm can be considered as an online mirror descent (OMD) with batching and the Tsallis entropy used as prox:

$$x_{k+1} = \arg\min_{x\in\Delta^d_+}\left[\nu x^{\mathtt{T}}\tilde{g}^k_{med} + V_\psi(x, x_k)\right].$$

Thus standard inequation for OMD holds:

$$\langle \tilde{g}^k_{med}, x_k - u\rangle \leq \frac{1}{\nu}\left[V_\psi(u,x_k) - V_\psi(u,x_{k+1}) - V_\psi(x_{k+1},x_k)\right] + \langle \tilde{g}^k_{med}, x_k - x_{k+1}\rangle. \tag{35}$$

From Tailor theorem we have

$$V_\psi(z,x_k) = \frac{1}{2}(z - x_k)^T\nabla^2\psi(y_k)(z - x_k) = \frac{1}{2}\|z - x_k\|^2_{\nabla^2\psi(y_k)}$$

for some point $y_k \in [z, x_k]$.

Hence we have

$$\langle \tilde{g}^k_{med}, x_k - x_{k+1}\rangle - \frac{1}{\nu}V_\psi(x_{k+1},x_k)$$
$$\leq \max_{z\in R^d_+}\left[\langle \tilde{g}^k_{med}, x_k - z\rangle - \frac{1}{\nu}V_\psi(z,x_k)\right]$$
$$= \left[\langle \tilde{g}^k_{med}, x_k - z^*_k\rangle - \frac{1}{\nu}V_\psi(z^*_k,x_k)\right]$$
$$\leq \frac{\nu}{2}\|\tilde{g}^k_{med}\|^2_{(\nabla^2\psi(y_k))^{-1}} + \frac{1}{2}\|z^* - x_k\|^2_{\nabla^2\psi(y_k)} - \frac{1}{\nu}V_\psi(z^*,x_k)$$
$$= \frac{\nu}{2}\|\tilde{g}^k_{med}\|^2_{(\nabla^2\psi(y_k))^{-1}},$$

where $z^* = \arg\max_{z\in\mathbb{R}^d_+}\left[\langle \tilde{g}^k_{med}, x_k - z\rangle - \frac{1}{\nu}V_\psi(z,x_k)\right]$.

Proceeding with (35), we get:

$$\langle \tilde{g}^k_{med}, x_k - u\rangle \leq \frac{1}{\nu}\left[V_\psi(u,x_k) - V_\psi(u,x_{k+1})\right] + \frac{\nu}{2}\|\tilde{g}^k_{med}\|^2_{(\nabla^2\psi(y_k))^{-1}}.$$

Sum over k gives

$$\begin{aligned}
&\sum_{k=0}^{K}\langle \tilde{g}_{med}^k, x_k - u\rangle \\
&\leq \frac{V_\psi(x_0,u)}{\nu} + \frac{\nu}{2}\sum_{k=0}^{K}(\tilde{g}_{med}^k)^T\left(\nabla^2\psi(y_k)\right)^{-1}\tilde{g}_{med}^k \\
&= 2\frac{d^{1/2}-\sum_{i=1}^d u_i^{1/2}}{\nu} + \nu\sum_{k=0}^{K}\sum_{i=1}^{d}(\tilde{g}_{med}^k)_i^2 y_{k,i}^{3/2},
\end{aligned} \tag{36}$$

where $y_k \in [x_k, z_k^*]$ and $z_k^* = \arg\max_{z\in R_+^d}\left[\langle \tilde{g}_{med}^k, x_k - z\rangle - \frac{1}{\nu}V_\psi(z,x_k)\right]$.

From the first-order optimality condition for z_k^* we obtain

$$-\nu(\tilde{g}_{med}^k)_i + (x_{k,i})^{1/2} = (z_{k,i}^*)^{1/2}$$

and thus we get $z_{k,i}^* \leq x_{k,i}$.

Thus (36) becomes

$$\sum_{k=0}^{K}\langle \tilde{g}_{med}^k, x_k - u\rangle \leq 2\frac{d^{1/2}-\sum_{i=1}^d u_i^{1/2}}{\nu} + \nu\sum_{k=0}^{K}\sum_{i=1}^{d}(\tilde{g}_{med}^k)_i^2 \cdot x_{k,i}^{3/2}$$

and concludes the proof.

Lemma 5 ***[Lemma 5.1 from (Sadiev et al., 2023)]*** *Let X be a random vector in $\mathbb{R}^d$ and $\bar{X} = clip(X,\lambda) = X \cdot \min\left\{1, \frac{\lambda}{\|X\|_2}\right\}$, then*

$$\|\bar{X} - \mathbb{E}[\bar{X}]\|_2 \leq 2\lambda. \tag{37}$$

Moreover, if for some $\sigma \geq 0$

$$\mathbb{E}[X] = x \in \mathbb{R}^n, \quad \mathbb{E}[\|X - x\|_2^2] \leq \sigma^2$$

and $\|x\|_2 \leq \frac{\lambda}{2}$, then

$$\left\|\mathbb{E}[\bar{X}] - x\right\|_2 \leq \frac{4\sigma^2}{\lambda}, \tag{38}$$

$$\mathbb{E}\left[\left\|\bar{X} - x\right\|_2^2\right] \leq 18\sigma^2, \tag{39}$$

$$\mathbb{E}\left[\left\|\bar{X} - \mathbb{E}[\bar{X}]\right\|_2^2\right] \leq 18\sigma^2. \tag{40}$$

Lemma 6 *Suppose that* Clipped-INF-med-SMD *with $1/2$-Tsallis entropy as prox-function generates the sequences $\{x_k\}_{k=0}^K$ and $\{\tilde{g}_{med}^k\}_{k=0}^K$, and for each arm i random reward $g_{t,i}$ at any step t has bounded expectation $\mathbb{E}[g_{t,i}] \leq \frac{\lambda}{2}$ and the noise $g_{t,i} - \mu_i$ has symmetric distribution, then for any $u \in \Delta_+^d$ holds:*

$$\mathbb{E}_{x_k,\mathbf{e}_{[k]},\xi_{[k]}}\left[\sum_{i=1}^{d}(\tilde{g}_{med}^k)_i^2 \cdot x_{k,i}^{3/2}\right] \leq 18\sigma^2 + \|\mu\|_2^2. \tag{41}$$

Proof:

$$\mathbb{E}_{x_k,\mathbf{e}_{[k]},\xi_{[k]}}\left[\sum_{i=1}^{d}(\tilde{g}_{med}^k)_i^2 \cdot x_{k,i}^{3/2}\right] \leq \mathbb{E}[\|\tilde{g}_{med}^k\|_2^2] \leq \mathbb{E}[\|\tilde{g}_{med}^k - \mu\|_2^2 + \|\mu\|_2^2] \overbrace{\leq}^{\text{Lemma 6 + Lemma 2}} 18\sigma^2 + \|\mu\|_2^2.$$

Proof of Theorem 3:

Firstly, for any $x, y \in \triangle_+^d$ we have

$$\|x - y\|_2 \leq \sqrt{2}. \tag{42}$$

Next we obtain

$$\begin{aligned}
\mathbb{E}\left[\mathcal{R}_T(u)\right] &= \mathbb{E}\left[\sum_{t=1}^{T} l(x_t) - \sum_{t=1}^{T} l(u)\right] \leq \mathbb{E}\left[\sum_{t=1}^{T}\langle \nabla l(x_t), x_t - u\rangle\right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T}\langle \mu - g_{med}^{k(t)}, x_{k(t)} - u\rangle\right] + \mathbb{E}\left[\sum_{t=1}^{T}\langle g_{med}^{k(t)} - \tilde{g}_{med}^{k(t)}, x_{k(t)} - u\rangle\right] + \mathbb{E}\left[\sum_{t=1}^{T}\langle \tilde{g}_{med}^{k(t)}, x_{k(t)} - u\rangle\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\langle g_{med}^{k(t)} - \tilde{g}_{med}^{k(t)}, x_{k(t)} - u\rangle\right] + \mathbb{E}\left[\sum_{t=1}^{T}\langle \tilde{g}_{med}^{k(t)}, x_{k(t)} - u\rangle\right] \\
&\leq \left[\sum_{t=1}^{T} \|\mathbb{E}[g_{med}^{k(t)}] - \mathbb{E}[\tilde{g}_{med}^{k(t)}]\|_2 \cdot \|x_{k(t)} - u\|_2\rangle\right] + \mathbb{E}\left[\sum_{t=1}^{T}\langle \tilde{g}_{med}^{k(t)}, x_{k(t)} - u\rangle\right] \\
&\overbrace{\leq}^{\text{Lemma 5, (42)}} \frac{4\sigma^2 T}{\lambda}\cdot\sqrt{2} + (2m+1)\mathbb{E}\left[\sum_{k=0}^{K}\langle \tilde{g}_{med}^{k}, x_k - u\rangle\right] \\
&\overbrace{\leq}^{\text{Lemma 4}} \frac{4\sqrt{2}\sigma^2 T}{\lambda} + 2(2m+1)\frac{d^{1/2} - \sum_{i=1}^{d} u_i^{1/2}}{\nu} + \nu(2m+1)\mathbb{E}\left[\sum_{k=0}^{K}\sum_{i=1}^{d}(\tilde{g}_{med}^{k})_i^2 \cdot x_{k,i}^{3/2}\right] \\
&\overbrace{\leq}^{\text{Lemma 6}} \frac{4\sqrt{2}\sigma^2 T}{\lambda} + 2(2m+1)\frac{\sqrt{d}}{\nu} + \nu T(18\sigma^2 + \|\mu\|_2^2) \\
&\leq 4\sqrt{2}\sigma(2m+1)^{1/2}T^{1/2} + 2\sqrt{2}d^{1/4}(2m+1)^{1/2}T^{1/2}\sqrt{18\sigma^2 + \|\mu\|_2^2} \\
&= 2\sqrt{2}(2m+1)^{1/2}T^{1/2}\left(2\sigma + d^{1/4}\sqrt{18\sigma^2 + \|\mu\|_2^2}\right).
\end{aligned}$$

Remark 7 *Following work (Dorn et al., 2024) let us consider M_2-Lipschitz non-linear loss function $l(x)$ defined on complex compact set $Q \subset \mathbb{R}^d$, which absolute value is bounded by constant Δ_l, i.e.,*

$$|l(x)| \leq \Delta_l, \quad \forall x \in Q.$$

Values of $l(x)$ are available via one point oracle $l(x, \xi) = l(x) + \xi$, where ξ is symmetric noise satisfying Assumption 3 with $\kappa > 0$.

In that case, we build one-point analog of gradient estimation (7), namely,

$$g(x, \mathbf{e}, \xi) = \frac{d}{\tau} l(x + \tau\mathbf{e}, \xi)\mathbf{e}.$$

Then it is guaranteed that Med^m with $m = \frac{2}{\kappa} + 1$ samples has bounded second moment, i.e.,

$$\mathbb{E}_{\mathbf{e},\xi}[\|\mathit{Med}^m(x, \mathbf{e}, \{\xi\})\|_q^2 | x] \leq \sigma_l^2 a_q^2,$$

where $\sigma_l^2 = 2\left(\frac{d\Delta_l}{\tau}\right)^2 + (4m+2)\left(\frac{4}{\kappa}\right)^{\frac{2}{\kappa}}\left(\frac{d\Delta}{\tau}\right)^2$.

Therefore, we apply the algorithm ZO-clipped-med-SMD *to solve MAB problem with σ_l instead of σ. In order to achieve average regret accuracy $\frac{1}{T}\mathbb{E}\left[\mathcal{R}_T(u)\right] \leq \varepsilon$ for the sequence of points $z_t = x_t + \tau\mathbf{e}_t$ generated by* ZO-clipped-med-SMD *at each oracle call one requires horizon*

$$T = O\left(\frac{1}{\kappa}\frac{d^2 M_2^2(\Delta_l + \Delta/\kappa^{\frac{1}{\kappa}})^2 a_q^2 D_{\Psi_p}^2}{\varepsilon^4}\right).$$