THREE-POINT METHOD WITH ZERO-ORDER ORACLE THAT INEXACTLY AND RANDOMLY COMPARES FUNCTION VALUES

Anonymous authors
Paper under double-blind review

ABSTRACT

In this paper, we consider zero-order optimization setting, in which it is not possible to return the function values; instead, the oracle is only capable of comparing these values. In order to address this formulation, one can utilize the well-established stochastic three-point method, which is able to select the minimum from the three points under consideration at each iteration. Furthermore, in this setting, we assume that the oracle produces inaccurate and random results. In particular, we consider strategies in which the probability of selecting the correct value is either constant (determined by a coin flip) or dependent on the difference between the values of the function at the current point and the minimum of the three points selected at this iteration. In a further strategy, we consider the possibility that the difference value may be subject to noise, whether random or deterministic. These settings aim to obtain a more approximate description of the real-world problems that arise, for instance, in human feedback systems. We select parameters in the stochastic three-point method for all considered strategies in the different cases and evaluated the convergence rates for strongly convex, convex and non-convex optimization problems. The obtained results are verified on practical examples.

Keywords: zero-order optimization · derivative-free optimization · stochastic optimization · stochastic three-point method · human feedback

1 INTRODUCTION

One of the most significant challenges in optimization is the availability of oracles, which are information sources that provide important data such as function values and gradients (Gasnikov et al., 2023). In many real-world applications, obtaining gradients is a major challenge. This is especially true for the following cases: ”black-box” functions (Liao et al., 2024), (Lobanov, 2024), for which we do not have access to their internal structure or how they are computed, and consequently can not obtain gradients for them; or complex simulators, where the computational cost makes gradient acquisition extremely expensive and impractical. Such problems frequently arise in the context of real-world applications, including, for example, classical machine learning (Chen et al., 2023), deep learning (Zhang et al., 2020), reinforcement learning (Yani et al., 2021), natural language processing (Munos et al., 2023) or emerging signal processing (Liu et al., 2020).

In instances where the gradient of a function is unavailable, gradient-free optimization methods may be employed as an alternative solution. The field of zero-order methods (Nesterov & Spokoiny, 2017) represents a significant area of research that has been developing for a considerable period of time, resulting in a substantial body of literature (Alarie et al., 2021), (Nocedal & Wright, 2006), (Conn et al., 2009). Nowadays, zero-order methods, which approximate the gradient through finite differences, occupy a distinctive position within the optimization community. Additionally, there are other well-known approaches, such as Monte Carlo methods (Mohamed et al., 2020), dichotomy methods, and the ellipsoid method (Nemirovski). Gradient-free methods perform well on low-dimensional problems; however, they converge more slowly and require more careful tuning (Sudharshan & Jeyakumar, 2021).

The development of the theory of gradient-free optimization has a long history, but modern realities require the investigation of novel formulations that are also zero-order (Borghi et al., 2024), (Hutchinson & Alizadeh, 2024), (Sahinoglu & Shahrampour, 2024). To illustrate, there are issues that require input from human beings, such as the ranking data that is displayed in search engines or social media feeds. Such tasks frequently arise in the context of reinforcement learning (Ouyang et al., 2022), (Bai et al., 2022), (Tang et al., 2023). Settings described above include a more complex model in which the oracle can no longer return function values, but is instead capable of comparing them with each other and output the point at which the value is smaller. Such examples are evident in the work of (Lobanov et al., 2024). However, a significant limitation of these results is that it is more theoretical than practical. This approach necessitates the sampling of a substantial number of points along a single direction, which yields favourable theoretical convergence.

A more practical approach would be to change the direction at each iteration, which may yield faster results for some problems. That is the reason in our paper we consider one of the classical direct search methods: the stochastic three-point method (STP) (Bergou et al., 2020) in a more complex model of oracle described above.

1.1 Contributions

Different concepts of the oracle: In this study, we examine a zero-order oracle that is susceptible to errors in the comparison of function values. A number of settings of this inexactness is considered, including the following:

• The oracle's output is randomly generated with a constant probability, and in the event of an incorrect response, the current value is returned. In the context of human feedback, this corresponds to the situation that the user sometimes sees no difference between some options to choose from and leaves the current option unchanged.

• The previous setting but the probability is equal to a sigmoid depending on the difference of function values between the next x^{k+1} and the current x^k value of the argument. This setting reflects the sensitivity of the system, whereby a greater difference will result in a greater perception of dissimilarity by the user.

• And more complex, in addition to the probability with sigmoid dependence, the oracle responses are subject to noise, comprising both random and deterministic components, which are added to the difference.

The impact of our concepts of the oracle on the original algorithmic complexities: Similarly to the paper of (Bergou et al., 2020), the convergence of the algorithm is evaluated for three cases of the L-smooth target function, namely non-convex, convex and strongly convex. We analyse how convergence differs between different concepts of the oracle and the original deterministic algorithm from (Bergou et al., 2020). In light of the increased complexity of the productions, a modified analysis is required for each of the aforementioned items. We also present the parameters that should be selected for the methods to obtain the convergence estimates described above.

Experiments: We conduct experiments on convex and strongly convex target functions, with parameters selected to align with these classes. The experiments illustrate the impact of the problem formulation on the convergence of the method, particularly evident in the formulations with the sigmoid function.

2 Problem statement

2.1 Notation

By $\mathbb{E}[\cdot]$ is denoted the mathematical expectation and by $\mathbb{E}[\cdot|A]$ – the conditional mathematical expectation (subject to condition A). We denote by $\mathcal{D}$ a probability distribution over $\mathbb{R}^n$. The Euclidean norm is denoted by $\|\cdot\|_2$.

For the next notation, it is essential to make the following assumption as in the paper of (Bergou et al., 2020), which can be interpreted as a definition of a norm dependent on $\mathcal{D}$, denoted by $\|\cdot\|_{\mathcal{D}}$.

Assumption 1. The probability distribution $\mathcal{D}$ on $\mathbb{R}^n$ is characterised by the following properties:

1. $0 < \gamma_{\mathcal{D}} := \mathbb{E}_{s\sim\mathcal{D}}\, \|s\|_2^2 < \infty$.
2. $\exists$ a constant $\mu_{\mathcal{D}} > 0$ and norm $\|\cdot\|_{\mathcal{D}}$ on $\mathbb{R}^n$ such $\forall g \in \mathbb{R}^n$,

$$\mathbb{E}_{s\sim\mathcal{D}}\, |\langle g, s\rangle| \geq \mu_{\mathcal{D}} \|g\|_{\mathcal{D}}. \tag{1}$$

2.2 ASSUMPTIONS

In this subsection, we introduce several assumptions that we use in the paper.

Assumption 2. The objective function f is said to have a Lipschitz continuous gradient with a constant $L > 0$ (it is L-smooth), i.e. the following conditions are met:

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2, \qquad \forall x, y \in \mathbb{R}^n$$

Additionally, the function f is bounded from below by $f^* = f(x^*) \in \mathbb{R}$, $x^* \in \mathbb{R}^n$ and $f(x) \geq f(x^*)$ for all $x \in \mathbb{R}^n$.

This is one of the most fundamental assumptions for analyzing a function.

In order to proceed, it is necessary to consider one of the gradient-free methods that can be employed in conjunction with the oracle that we use: the Stochastic Three Points (STP) method (Bergou et al., 2020). At iteration k of `STP`, a random direction s_k is generated by sampling from the selected distribution $\mathcal{D}$. The subsequent iteration is as follows:

$$x^{k+1} = \arg\min\left\{f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k), f(x^k)\right\}, \tag{2}$$

where $\alpha_k > 0$ is an appropriately chosen stepsize. We make the same assumption regarding the function f and the probability law $\mathcal{D}$ as in the paper of (Bergou et al., 2020).

Assumption 3. The probability distribution $\mathcal{D}$ on $\mathbb{R}^n$ is such that it generates points that lie on the unit hypersphere. In other words, every point s sampled from $\mathcal{D}$ has a Euclidean norm of 1. Given that all norms in $\mathbb{R}^n$ are equivalent, there exists a positive constant $C_{\mathcal{D}}$ such that for all $x \in \mathbb{R}^n$: $\|x\|_2 \leq C_{\mathcal{D}}\|x\|_{\mathcal{D}}$.

Note also several options for the distribution law $\mathcal{D}$, which are discussed in the STP paper (see Lemma 3.4 from the paper of (Bergou et al., 2020)).

Lemma 4. Let $g \in \mathbb{R}^n$. If the distribution law $\mathcal{D}$ is

1. the uniform distribution on the unit sphere in $\mathbb{R}^n$, then
$$\gamma_{\mathcal{D}} = 1, \quad \mathbb{E}_{s\sim\mathcal{D}}|\langle g, s\rangle| \sim \tfrac{1}{\sqrt{2\pi n}}\|g\|_2, \quad \|\cdot\|_{\mathcal{D}} = \|\cdot\|_2 \quad \text{and} \quad \mu_{\mathcal{D}} \sim \tfrac{1}{\sqrt{2\pi n}}.$$

2. the uniform distribution on $\{e_1, \ldots, e_n\}$, then
$$\gamma_{\mathcal{D}} = 1, \quad \mathbb{E}_{s\sim\mathcal{D}}\,|\langle g, s\rangle| = \tfrac{1}{n}\|g\|_1, \quad \|\cdot\|_{\mathcal{D}} = \|\cdot\|_1 \quad \text{and} \quad \mu_{\mathcal{D}} = \tfrac{1}{n}.$$

The next assumptions are used in only a few sections. At the beginning of these sections there are references to the assumptions used.

Assumption 5. The objective function f is said to be convex, i.e. the following condition is met:

$$f(y) \geq f(x) + \langle\nabla f(x), y - x\rangle, \qquad \forall x, y \in \mathbb{R}^n.$$

Moreover, the function f has a minimizer x^* and a bounded level set at x_0, i.e. the following condition is satisfied:

$$R_0 := \max\{\|x - x^*\|_{\mathcal{D}}^* \ : \ f(x) \leq f(x_0)\} < +\infty, \qquad \forall x \in \mathbb{R}^n,$$

where $\|\xi\|_{\mathcal{D}}^* := \max\{\langle\xi, x\rangle \mid \|x\|_{\mathcal{D}} \leq 1\}$ defines the dual norm to $\|\cdot\|_{\mathcal{D}}$.

Assumption 6. The objective function f is said to be λ-strongly convex with respect to the norm $\|\cdot\|_{\mathcal{D}}$, i.e. the following condition is fulfilled:

$$f(y) \geq f(x) + \langle\nabla f(x), y - x\rangle + \tfrac{\lambda}{2}\|x - y\|_{\mathcal{D}}^2, \qquad \forall x, y \in \mathbb{R}^n.$$

2.3 THE DESCRIPTION OF THE MODEL UNDER CONSIDERATION

In this paper, we consider the problem of minimizing a given smooth objective function $f : \mathbb{R}^n \to \mathbb{R}$ over the real vector space $\mathbb{R}^n$

$$\min_{x\in\mathbb{R}^n} f(x). \tag{3}$$

As previously stated, it is assumed that we do not have access to the derivatives of f and only have access to a function evaluation oracle: it is only capable of comparing points with each other, but not of counting the difference between the values of the function at those points. Formally, the oracle can only return a minimum of several values, including an error, which are described further in several ways.

3 THREE MODIFICATIONS OF THE ORACLE FOR THE STP METHOD

As mentioned above, in the context of real-world tasks, such as recommender systems, issues related to metrics arise. Metrics such as click-through rate (CTR) or precision are discrete and may not change with minor alterations to the model parameters. Consequently, even when utilizing smooth functions to model recommendations, gradient-free optimization methods may become "stuck" due to the discrete nature of the metrics. However, the STP method does not take into account the "stuck" problem described above. This is the issue that we seek to address in the proposed modification of the oracle for the STP method.

As mentioned above, the oracle gives an inexact solution, i.e. the equation (2) is obtained only with some probability:

$$x^{k+1} = \begin{cases} \arg\min\left\{f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k), f(x^k)\right\}, & \text{with probability } p_k, \\ x^k, & \text{with probability } 1 - p_k \end{cases}, \tag{4}$$

where the probability p_k depends on the model considered.

Strategies for obtaining this probability can be different, and we consider them in the next three subsections.

3.1 THE INITIAL MODIFICATION OF THE ORACLE FOR STP (WITH A CONSTANT PROBABILITY)

In our initial modification of the oracle for STP, we assume that with a constant probability p, a method step is taken, and with probability $1 - p$, the value of x^k remains unchanged. Consequently, the iteration is represented in the equation (4), where $p \in [0, 1]$ is a constant.

The following theorem gives the expressions that we use later on for convergence estimates.

Lemma 7. If Assumptions 1 and 2 hold, then for all $k \geq 0$,

$$\mathbb{E}\left[f(x^{k+1}) \mid x^k\right] \leq f(x^k) - p\mu_{\mathcal{D}}\alpha_k \|\nabla f(x^k)\|_{\mathcal{D}} + \tfrac{Lp}{2}\alpha_k^2, \tag{5}$$

and

$$\theta_{k+1} \leq \theta_k - p\mu_{\mathcal{D}}\alpha_k g_k + \tfrac{Lp}{2}\alpha_k^2, \tag{6}$$

where $\theta_k = \mathbb{E}[f(x^k)]$ and $g_k = \mathbb{E}[\|\nabla f(x^k)\|_{\mathcal{D}}]$.

See the proof in Appendix A.1.

As can be observed, the results obtained are comparable to those obtained for the STP method (see Lemma 3.5 from (Bergou et al., 2020)). The only difference is that the final two summands are multiplied by the constant p.

3.2 THE PRINCIPAL MODIFICATION OF THE ORACLE FOR STP AND ITS SPECIAL CASE (WITH A SIGMOID PROBABILITY DISTRIBUTION INCORPORATING NOISE)

In this section, we set the main challenge of our study. In real-world problems, it is not uncommon for the desired result to be obtained with some probability, and for the influence of noise on the obtained result to be significant. Therefore, in our theoretical formulation, we attempt to take into account the aforementioned aspects of practical challenges.

As in the previous modification, we assume that with a probability p_k, a method step is taken, and with probability $1 - p_k$, the method remains in place (see the equation (4)). The difference is that here p_k is a sigmoid function with noise:

$$p_k = \left(1 + \exp\left(-f(x^k) + \min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} + \Delta_k\right)\right)^{-1}.$$

In order to clarify the nature of the noise Δ_k, it is necessary to make the following assumption.

Assumption 8. Δ represents random noise, which can be broken down into the following sum:

$$\Delta = \xi + \delta,$$

where $|\delta| < \infty$, $\mathbb{E}[\xi] = 0$, and $\mathbb{E}\left[\xi^2\right] = \sigma^2$.

We also assume that the noise Δ can be equal to 0, this special case is considered further. The step there is as in the equation (4), but unlike the principal modification of the oracle, this one does not take noise into account. Subsequently, we evaluate the convergence of this modification. Before the lemma, we need an assumption for the proof (it is not needed for the special case under consideration).

Assumption 9. ξ is a sub-Gaussian random variable, i.e. $\mathbb{E}[e^{c\xi^2}] < \infty$ for some constant $c > 0$. Using that $\mathbb{E}[\xi] = 0$, this is equivalent to the statement that

$$\mathbb{E}[e^{q\xi}] \leq e^{\sigma^2 q^2/2},$$

where $q \in \mathbb{R}$, $\mathbb{E}[\xi^2] \leq \sigma^2$.

The following lemma can now be formulated:

Lemma 10. If Assumptions 1, 2, 8 and 9 hold, then for all $k \geq 0$,

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \leq f(x^k) - \tfrac{1}{1+e^{\sigma^2/2}e^{\delta}} \cdot \mu_{\mathcal{D}}\alpha_k\|\nabla f(x^k)\|_{\mathcal{D}} + \tfrac{L}{2}\alpha_k^2, \tag{7}$$

and

$$\theta_{k+1} \leq \theta_k - \tfrac{1}{1+e^{\sigma^2/2}e^{\delta}} \cdot \mu_{\mathcal{D}}\alpha_k g_k + \tfrac{L}{2}\alpha_k^2, \tag{8}$$

where $\theta_k = \mathbb{E}[f(x^k)]$ and $g_k = \mathbb{E}[\|\nabla f(x^k)\|_{\mathcal{D}}]$.

We also present a special case of Lemma 10, where the noise Δ is equal to 0.

Lemma 11. If Assumptions 1, 2 and 8 hold and $\Delta = 0$, then for all $k \geq 0$,

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \leq f(x^k) - \tfrac{1}{2}\mu_{\mathcal{D}}\alpha_k\|\nabla f(x^k)\|_{\mathcal{D}} + \tfrac{L}{4}\alpha_k^2, \tag{9}$$

and

$$\theta_{k+1} \leq \theta_k - \tfrac{1}{2}\mu_{\mathcal{D}}\alpha_k g_k + \tfrac{L}{4}\alpha_k^2, \tag{10}$$

where $\theta_k = \mathbb{E}[f(x^k)]$ and $g_k = \mathbb{E}[\|\nabla f(x^k)\|_{\mathcal{D}}]$.

The proof for both Lemmas 10 and 11 is in Appendix A.1.

The result differs from that of the analogous result for STP (see Lemma 3.5 from (Bergou et al., 2020)) only in that the negative summand is multiplied by the constant $\frac{1}{1+e^{\sigma^2/2}e^{\delta}} \leq 1$, and for special case last two summands are multiplied by $\frac{1}{2}$.

4 The Selection of Parameters and Estimation of Convergence

This section presents the convergence estimates that can be obtained for the proposed methods by selecting the parameters in question. Three classes of target functions are considered in this study: non-convex, convex, and strongly convex. The parameters selected are either fixed or dependent on the iteration number.

The proofs proposed next depend on the equations (5), (6), (7), and (8) obtained for each method. In order to avoid the repetition of similar proofs, we prove the inequalities in general form.

The general form of the equations (5) and (7) is as follows:

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \leq f(x^k) - C_k^1 \cdot \mu_{\mathcal{D}}\alpha_k\|\nabla f(x^k)\|_{\mathcal{D}} + C_k^2 \cdot \tfrac{L}{2}\alpha_k^2, \tag{11}$$

while for (6) and (8), the form is:

$$\theta_{k+1} \leq \theta_k - C_k^1 \cdot \mu_{\mathcal{D}}\alpha_k g_k + C_k^2 \cdot \tfrac{L}{2}\alpha_k^2, \tag{12}$$

where the constants C_k^1, C_k^2 are assumed to be positive.

Thus the main differences from the key lemma of the original paper are in constants, which have estimates: $0 \leq C_k^1 \leq C_k^2 \leq 1$, in the considered cases.

The following three sections discuss the selection of the stepsize for different classes of functions. We make all the same assumptions as in similar sections for the initial modification of the oracle for STP and prove similar theorems.

4.1 Non-convex

In this section, we present our most general complexity result. We make no assumptions regarding the smoothness or boundedness of the function f (see Assumption 2).

Lemma 12 (Monotonicity). STP produces a monotonic sequence of iterates, i.e., $f(x^{k+1}) \leq f(x^k)$ for all $k \geq 0$. As a consequence,

$$\mathbb{E}[f(x^{k+1}) \mid x^k] \leq f(x^k). \tag{13}$$

The following theorem considers a method in which the stepsize α_k is chosen to decrease with the iteration number. The convergence of this method is evaluated in light of this parameter choice.

Theorem 13 (Decreasing stepsize). Let Assumptions 1, 2, 8, 9 and Lemma 12 hold. Choose $\alpha_k = \frac{\alpha_0}{\sqrt{k+1}}$, where $\alpha_0 > 0$. If

$$K \geq \frac{2\left(\frac{\sqrt{2}(f(x_0)-f^*)}{C_k^1\alpha_0} + \frac{C_k^2 L\alpha_0}{2C_k^1}\right)^2}{\mu_{\mathcal{D}}^2\varepsilon^2}, \tag{14}$$

then $\min\limits_{k=0,1,\ldots,K} \mathbb{E}\left[\|\nabla f(x^k)\|_{\mathcal{D}}\right] \leq \varepsilon.$

See the proof in Appendix A.1.

Although the stepsize α_k obtained in this case is identical to that obtained in the same case for STP, the estimate for the number of iterations K increased slightly.

It should be noted that the complexity of the problem is dependent on the value of α_0.

Corollary 14. The optimal choice, which minimizes the complexity bound, is

$$\alpha^* = 8^{1/4}\sqrt{\frac{f(x_0)-f^*}{C_k^2 L}},$$

in which case the complexity bound from the equation (14) takes the form

$$\frac{4\sqrt{2}C_k^2(f(x_0)-f^*)L}{(C_k^1)^2\mu_{\mathcal{D}}^2\varepsilon^2}. \tag{15}$$

Furthermore, since f is L-smooth, and it is necessary $\nabla f(x^*) = 0$, the optimal stepsize is no greater than

$$\alpha^* \leq \frac{2^{1/4}}{\sqrt{C_k^2}}\|x_0 - x^*\|_2.$$

Moreover, it is possible to select a fixed step size for this particular case. It can be also observed that the estimated values have undergone a slight increase.

Theorem 15 (Fixed stepsize). Let f satisfy Assumption 2, Lemma 12. Choose a fixed stepsize $\alpha_k = \alpha_0$ with $0 < \alpha_0 < \frac{2C_k^1\mu_{\mathcal{D}}\varepsilon}{C_k^2 L}$. If

$$K \geq k(\varepsilon) := \left\lceil \frac{f(x_0)-f^*}{\left(C_k^1\mu_{\mathcal{D}}\varepsilon - \frac{C_k^2 L}{2}\alpha_0\right)\alpha_0} \right\rceil - 1, \tag{16}$$

then $\min\limits_{k=0,1,\ldots,K} \mathbb{E}\left[\|\nabla f(x^k)\|_{\mathcal{D}}\right] \leq \varepsilon$. In particular, if $\alpha = \frac{C_k^1\mu_{\mathcal{D}}\varepsilon}{C_k^2 L}$, then $k(\varepsilon) = \left\lceil \frac{2C_k^2 L(f(x_0)-f^*)}{(C_k^1)^2\mu_{\mathcal{D}}^2\varepsilon^2} \right\rceil - 1$.

See the proof in Appendix A.1.

It should be noted that if the constants C_k^1 and C_k^2 are identical, as is the case with the initial iteration of the STP method, then the step size α_k remains unchanged.

4.2 Convex

In this section, we estimate the complexity of the modification of the oracle for STP in the case of a convex f (see Assumption 5). It should be noted that if the aforementioned assumption is valid, then for any value of x such that $f(x) \leq f(x_0)$, we obtain $f(x) - f(x^*) \leq \langle \nabla f(x), x - x^* \rangle \leq \|\nabla f(x)\|_{\mathcal{D}}\|x - x^*\|_{\mathcal{D}}^* \leq R_0\|\nabla f(x)\|_{\mathcal{D}}$. That is,

$$\|\nabla f(x)\|_{\mathcal{D}} \geq \frac{f(x)-f(x^*)}{R_0}. \tag{17}$$

In the following theorem we present our primary result on complexity. We commence our analysis with the examination of the first modification of the oracle for STP algorithm with constant stepsizes.

Theorem 16 (Constant stepsize). Let Assumptions 1, 2 and 5 be satisfied. Let $0 < \varepsilon < \frac{C_k^2 L R_0^2}{(C_k^1)^2 \mu_{\mathcal{D}}^2}$ and choose constant stepsize $\alpha_k = \alpha = \frac{C_k^1 \varepsilon \mu_{\mathcal{D}}}{C_k^2 L R_0}$. If

$$K \geq \frac{C_k^2 L R_0^2}{(C_k^1)^2 \mu_{\mathcal{D}}^2 \varepsilon} \log\left(\frac{2(f(x_0) - f(x^*))}{\varepsilon}\right), \tag{18}$$

then $\mathbb{E}\left[f(x^K) - f(x^*)\right] \leq \varepsilon$.

See the proof in Appendix A.1.

The following theorem demonstrates the optimal selection of a variable step for the proposed modification of the oracle for STP.

Theorem 17 (Variable stepsize). Let Assumptions 1, 2 and 5 be satisfied.

Let $\alpha_k = \alpha_0 \left(f(x^k) - f(x^*)\right)$, where $0 < \alpha_0 < \frac{2C_k^1 \mu_{\mathcal{D}}}{C_k^2 L R_0}$. Define $a = \frac{C_k^1 \mu_{\mathcal{D}} \alpha_0}{R_0} - \frac{C_k^2 L \alpha_0^2}{2} > 0$. If $k \geq k(\varepsilon) := \frac{1}{a}\left(\frac{1}{\varepsilon} - \frac{1}{r_0}\right)$, then $\mathbb{E}\left[f(x^k) - f(x^*)\right] \leq \varepsilon$.

See the proof in Appendix A.1.

In the previous theorem, the stepsize α_k was dependent on $f(x^*)$, but it is not always possible to obtain the value of $f(x^*)$. Therefore, let us formulate the following theorem.

Theorem 18 (Solution-free stepsize). Let Assumptions 1, 2, 3 and 5 be satisfied. Let $\alpha_k = \frac{C_k^1 |f(x^k + t s_k) - f(x^k)|}{C_k^2 L t}$, where

$$0 < t \leq \frac{\sqrt{2} C_k^1 \mu_{\mathcal{D}} \mathbb{E}\left[f(x^{K-1}) - f^*\right]}{C_k^2 L R_0}.$$

Define $a = \frac{(C_k^1)^2 \mu_{\mathcal{D}}^2}{4 C_k^2 L R_0^2}$. If $K \geq k(\varepsilon) := \frac{1}{a}\left(\frac{1}{\varepsilon} - \frac{1}{r_0}\right)$, then $\mathbb{E}\left[f(x^K) - f(x^*)\right] \leq \varepsilon$.

See the proof in Appendix A.1.

4.3 Strongly Convex

In this subsection, we define x^* as the unique minimizer of f. The function f is λ-strongly convex (see Assumption 6). As in the previous subsections, we find the appropriate stepsize α_k.

Theorem 19. Let Assumptions 1, 2 and 6 be satisfied. Let stepsize $\alpha_k = \frac{C_k^1 \mu_{\mathcal{D}}}{C_k^2 L} \sqrt{2\lambda(f(x^k) - f(x^*))}$. If

$$K \geq \frac{C_k^2 L}{(C_k^1)^2 \lambda \mu_{\mathcal{D}}^2} \log\left(\frac{f(x_0) - f(x^*)}{\varepsilon}\right), \tag{19}$$

then $\mathbb{E}\left[f(x^K) - f(x^*)\right] \leq \varepsilon$.

See the proof in Appendix A.1.

5 The final convergence table

Several tables illustrating the convergence results obtained for the various proposed methods and the parameters chosen are presented below.

For simplicity, we present the values of the constants for each concept of the oracle are presented in tabular form (see Table 5).

It is crucial to acknowledge that despite the introduction of "unproductive" steps, which have led to some estimates becoming less favourable, these steps are computationally cost-effective. This is because we are not required to compute two additional points ($x_k \pm \alpha_k s_k$) and the values of the target function in them. Furthermore, the value at the current point is already known from the previous iteration.

	Fixed stepsize	Variable stepsize
NC	$K \geq \left\lceil 2\frac{f(x_0)-f^*}{\left(2C_k^1\mu_{\mathcal{D}}\varepsilon - C_k^2 L\alpha_0\right)\alpha_0}\right\rceil - 1$	$K \geq \frac{2}{\mu_{\mathcal{D}}^2\varepsilon^2}\left(\frac{\sqrt{2}(f(x_0)-f^*)}{C_k^1\alpha_0} + \frac{C_k^2 L\alpha_0}{2C_k^1}\right)^2$
C	$K \geq \frac{C_k^2 L R_0^2}{(C_k^1)^2\mu_{\mathcal{D}}^2\varepsilon}\log\left(\frac{2(f(x_0)-f(x^*))}{\varepsilon}\right)$	$K \geq \frac{1}{a}\left(\frac{1}{\varepsilon} - \frac{1}{r_0}\right)$, $a = \frac{C_k^1\mu_{\mathcal{D}}\alpha_0}{R_0} - \frac{C_k^2 L\alpha_0^2}{2} > 0$
SC	–	$K \geq \frac{C_k^2 L}{(C_k^1)^2\lambda\mu_{\mathcal{D}}^2}\log\left(\frac{f(x_0)-f(x^*)}{\varepsilon}\right)$

Table 1: Convergence estimates for the general case

Oracle's modification	C_k^1	C_k^2	Lemma
Constant p	p	p	7
Sigmoid p	$\frac{1}{2}$	$\frac{1}{2}$	11
Sigmoid with noise p	$\frac{1}{1+e^{\sigma^2/2}e^{\delta}}$	1	10

Table 2: Values of constants for modifications

5.1 Why sigmoid can be better

Let us consider the function $f(x) = x^2$, where $x \in \mathbb{R}$. We determine the number of iterations required by the modifications for this function to achieve ε accuracy. Let the parameters be as follows: $\varepsilon = 0.5$, $x_0 = 10$, $\alpha_k = \frac{1}{\sqrt{1+k}}$.

Note that s_k in the one-dimensional case is equal to ± 1. At each step, the method selects $\arg\min\{f(x^k), f(x^k \pm \alpha_k s_k)\}$, hence we assume $s_k = 1$.

Further reasoning can be found in the Appendix A.2. Please refer to the Appendix for some explanations and experiments using α_k from Theorem 19.

	$\varepsilon = 0.5$ $x_0 = 10$	$\varepsilon = 0.05$ $x_0 = 10$	$\varepsilon = 0.005$ $x_0 = 10$	$\varepsilon = 0.5$ $x_0 = 50$	$\varepsilon = 0.5$ $x_0 = 100$
$p = 1/2$	99	109	113	2501	10004
Sigmoid	46	59	64	1073	4259
Sigmoid + noise $\in \mathcal{N}(0.1, 0.5)$	56	80	90	1337	5332

Table 3: The number of iterations required to achieve accuracy ε

6 Experiments

In this section, we describe the experiments performed.

6.1 Toy problem

The experiments for the function $f(x) = x^2$ were conducted under the same parameters for which the theoretical results are obtained (see Subsection 5.1). Each method is executed 100 times, and the resulting iteration count was taken as the mean value. The theoretical results are in close alignment with the practical outcomes.

The plots below illustrate the number of iterations required by each method to achieve a specified level of accuracy defined as ε.

The following conclusions may be drawn:

• The convergence with constant probability differs from the original STP for this problem by no more than an order of magnitude.

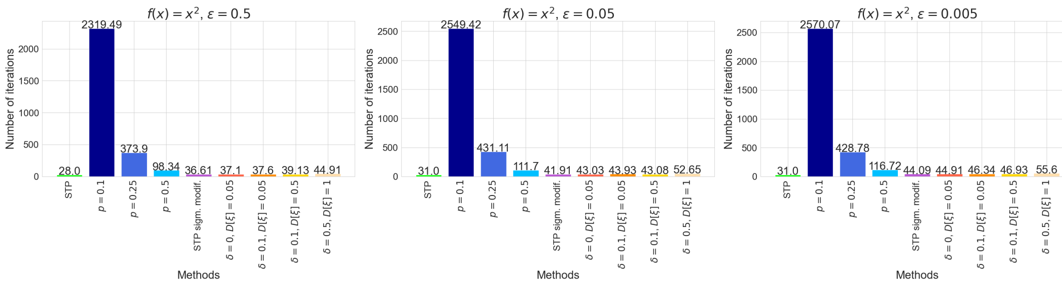

Figure 1: The number of iterations needed to achieve accuracy ε

• The sigmoid function makes convergence with respect to $p = \frac{1}{2}$ for "large steps", which is typical for the initial iterations.

• In this problem, the addition of noise to the sigmoid function resulted in a deterioration of the results. However, even the theoretical estimates of convergence for the sigmoid with noise demonstrate superior outcomes in comparison to those of the constant $p = \frac{1}{2}$.

6.2 More complex experiments

We considered the convex:

$$f(x) = \tfrac{1}{2}(x_1^2 + \sum_{i=1}^{n-1}(x_{i+1} - x_i)^2 + x_n^2) - x_1$$

and strongly convex target functions:

$$f(x) = \tfrac{\mu}{2}(x_1^2 + \sum_{i=1}^{n-1}(x_{i+1} - x_i)^2 - 2x_1 + \|x\|^2).$$

The targets functions and its lower bounds were taken from (Nesterov, 2013). The stepsizes α_k were selected in accordance with Theorem 18 and Theorem 19. The remaining parameters are labelled in the legend or taken from those selected in the STP paper: $t = 10^{-4}$, $L = 1$. All methods were run 100 times, with the resulting values averaged at each iteration.

The STP sigmoid modification represents a modification of the oracle for the STP method with a sigmoidal probability distribution. Its convergence is analogous to that of the initial modification of the oracle for STP with $p = 0.5$.

In the legend, the p value is specified to the initial modification of the oracle for STP method and the variables δ and $\mathcal{D}[\xi]$ are specified to the principal modification.

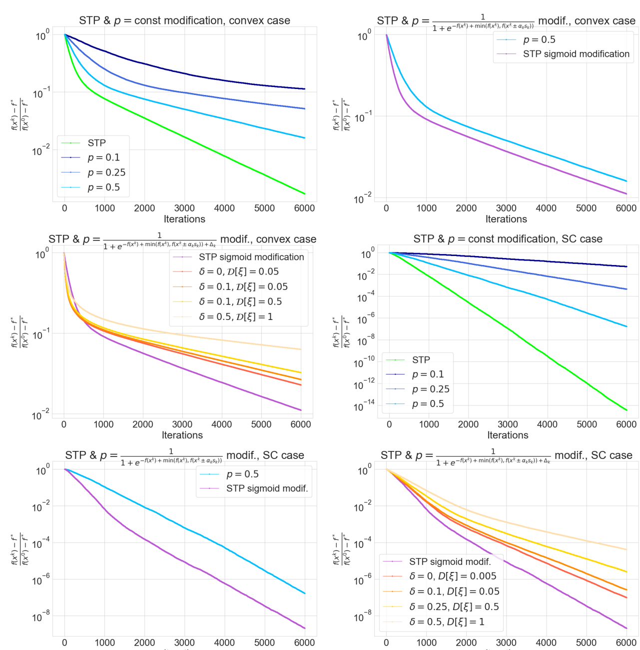

Figure 2: Convergence of STP in the context of different oracle's modifications

For an understanding of the impact of the sigmoid function in experimental contexts, please refer to the Appendix A.3. Please refer to the Appendix A.4 for some additional results on the convergence of different productions.

References

Stéphane Alarie, Charles Audet, Aïmen E Gheribi, Michael Kokkolaras, and Sébastien Le Digabel. Two decades of blackbox optimization applications. *EURO Journal on Computational Optimiza-*

tion, 9:100011, 2021.

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*, 2022.

El Houcine Bergou, Eduard Gorbunov, and Peter Richtárik. Stochastic three points method for unconstrained smooth minimization. *SIAM Journal on Optimization*, 30(4):2726–2749, 2020.

Giacomo Borghi, Hui Huang, and Jinniao Qiu. A particle consensus approach to solving nonconvex-nonconcave min-max problems. *arXiv preprint arXiv:2407.17373*, 2024.

Aochuan Chen, Yimeng Zhang, Jinghan Jia, James Diffenderfer, Jiancheng Liu, Konstantinos Parasyris, Yihua Zhang, Zheng Zhang, Bhavya Kailkhura, and Sijia Liu. Deepzero: Scaling up zeroth-order optimization for deep model training. *arXiv preprint arXiv:2310.02025*, 2023.

A.R. Conn, K. Scheinberg, and L.N. Vicente. *Introduction to Derivative-free Optimization*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2009. ISBN 9780898718768. URL `https://books.google.ru/books?id=7c7X6tlcaHEC`.

Alexander Gasnikov, Darina Dvinskikh, Pavel Dvurechensky, Eduard Gorbunov, Aleksandr Beznosikov, and Alexander Lobanov. Randomized gradient-free methods in convex optimization. In *Encyclopedia of Optimization*, pp. 1–15. Springer, 2023.

Spencer Hutchinson and Mahnoosh Alizadeh. Safe online convex optimization with multi-point feedback. In *6th Annual Learning for Dynamics & Control Conference*, pp. 168–180. PMLR, 2024.

Zhaohe Liao, Jiangtong Li, Li Niu, and Liqing Zhang. Align and aggregate: Compositional reasoning with video alignment and answer aggregation for video question-answering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13395–13404, 2024.

Sijia Liu, Pin-Yu Chen, Bhavya Kailkhura, Gaoyuan Zhang, Alfred O Hero III, and Pramod K Varshney. A primer on zeroth-order optimization in signal processing and machine learning: Principals, recent advances, and applications. *IEEE Signal Processing Magazine*, 37(5):43–54, 2020.

Aleksandr Lobanov. Improved iteration complexity in black-box optimization problems under higher order smoothness function condition. *arXiv preprint arXiv:2407.03507*, 2024.

Aleksandr Lobanov, Alexander Gasnikov, and Andrei Krasnov. The order oracle: a new concept in the black box optimization problems. *arXiv preprint arXiv:2402.09014*, 2024.

Shakir Mohamed, Mihaela Rosca, Michael Figurnov, and Andriy Mnih. Monte carlo gradient estimation in machine learning. *Journal of Machine Learning Research*, 21(132):1–62, 2020.

Rémi Munos, Michal Valko, Daniele Calandriello, Mohammad Gheshlaghi Azar, Mark Rowland, Zhaohan Daniel Guo, Yunhao Tang, Matthieu Geist, Thomas Mesnard, Andrea Michi, et al. Nash learning from human feedback. *arXiv preprint arXiv:2312.00886*, 2023.

A Nemirovski. Efficient methods. *Ekonomika i Mat. Metody*, 15:79.

Yurii Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2013.

Yurii Nesterov and Vladimir Spokoiny. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017.

J. Nocedal and S. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2006. ISBN 9780387227429. URL `https://books.google.ru/books?id=7wDpBwAAQBAJ`.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35: 27730–27744, 2022.

Emre Sahinoglu and Shahin Shahrampour. Online optimization perspective on first-order and zero-order decentralized nonsmooth nonconvex stochastic optimization. *arXiv preprint arXiv:2406.01484*, 2024.

S. Sudharshan and G. Jeyakumar. Gradient-based versus gradient-free algorithms for reinforcement learning. In Aruna Tiwari, Kapil Ahuja, Anupam Yadav, Jagdish Chand Bansal, Kusum Deep, and Atulya K. Nagar (eds.), *Soft Computing for Problem Solving*, pp. 115–124, Singapore, 2021. Springer Singapore. ISBN 978-981-16-2709-5.

Zhiwei Tang, Dmitry Rybin, and Tsung-Hui Chang. Zeroth-order optimization meets human feedback: Provable learning via ranking oracles. *arXiv preprint arXiv:2303.03751*, 2023.

Mohamad Yani, Fernando Ardilla, Azhar Aulia Saputra, and Naoyuki Kubota. Gradient-free deep q-networks reinforcement learning: Benchmark and evaluation. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–5. IEEE, 2021.

Sushen Zhang, Ruijuan Chen, Wenyu Du, Ye Yuan, and Vassilios S Vassiliadis. A hessian-free gradient flow (hfgf) method for the optimisation of deep learning neural networks. *Computers & Chemical Engineering*, 141:107008, 2020.

A Appendix

Claim 20 (Jensen's inequality). Let f be a convex function and X a random variable. Then

$$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X]),$$

if both mathematical expectations exist.

A.1 Proofs

Proof of Lemma 7. The proof is analogous to that of a similar theorem for the STP method; therefore, certain calculations are not detailed here.

Firstly, it should be noted that the L-smoothness of f implies that

$$f(x^k + \alpha_k s_k) \quad \leq f(x^k) + \alpha_k \langle \nabla f(x^k), s_k \rangle + \tfrac{L}{2} \alpha_k^2 \|s_k\|_2^2,$$

and

$$f(x^k - \alpha_k s_k) \leq f(x^k) - \alpha_k \langle \nabla f(x^k), s_k \rangle + \tfrac{L}{2} \alpha_k^2 \|s_k\|_2^2.$$

Then,

$$\begin{aligned}
\mathbb{E}\left[f(x^{k+1})|x^k\right] &= p \min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} + (1-p) f(x^k) \\
&\leq p \min\{f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} + f(x^k) - p f(x^k) \\
&\leq p(f(x^k) - \alpha_k |\langle \nabla f(x^k), s_k \rangle| + \tfrac{L}{2} \alpha_k^2 \|s_k\|_2^2) + f(x^k) - p f(x^k) \\
&= p(-\alpha_k |\langle \nabla f(x^k), s_k \rangle| + \tfrac{L}{2} \alpha_k^2 \|s_k\|_2^2) + f(x^k).
\end{aligned}$$

To conclude the equation (5), we need to take the expectation in the above inequality with respect to $s_k \sim \mathcal{D}$, conditional on x^k, and use (1). By taking the expectation in (5) we get the equation (6). □

Proof of Lemmas 10 and 11. Similarly to the proof of the similar lemma (see Lemma 7):

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] = p \min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} + (1-p) f(x^k),$$

where

$$p = (1 + \exp(-f(x^k) + \min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} + \Delta_k))^{-1}.$$

Substituting, we obtain

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] = \frac{\min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} - f(x^k)}{1 + \exp(\min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} - f(x^k) + \Delta_k)} + f(x^k).$$

Let us denote $t := \min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} - f(x^k)$. Notice, that $t \le 0$ and it tends to 0 as k increases.

Then,

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] = \frac{t}{1 + e^{t+\Delta_k}} + f(x^k).$$

Let us estimate from above $\mathbb{E}\left[\frac{t}{1+e^{t+\Delta_k}}\right]$. Since $t \le 0$,

$$t \le 0 \Rightarrow \frac{t}{1 + e^{t+\Delta_k}} \le \frac{t}{1 + e^{\Delta_k}}.$$

We get

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \le \frac{\min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} - f(x^k)}{1 + e^{\Delta_k}} + f(x^k).$$

It has previously been proven that

$$\min\{f(x^k), f(x^k + \alpha_k s_k), f(x^k - \alpha_k s_k)\} \le f(x^k) - \alpha_k |\langle \nabla f(x^k), s_k \rangle| + \tfrac{L}{2}\alpha_k^2 \|s_k\|_2^2.$$

Hence,

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \le f(x^k) - \tfrac{\alpha_k}{1+e^{\Delta_k}} |\langle \nabla f(x^k), s_k \rangle| + \tfrac{L}{2(1+e^{\Delta_k})}\alpha_k^2 \|s_k\|_2^2.$$

Taking the expectation of the above inequality with respect to $s_k \sim \mathcal{D}$, conditional on x^k, and to apply the equation (1), we obtain

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \le f(x^k) - \mathbb{E}\left[\tfrac{1}{1+e^{\Delta_k}}|x^k\right] \mu_{\mathcal{D}} \alpha_k \|\nabla f(x^k)\|_{\mathcal{D}} + \mathbb{E}\left[\tfrac{L}{2+2e^{\Delta_k}}|x^k\right] \alpha_k^2. \tag{20}$$

By taking the expectation in (20), we obtain

$$\theta_{k+1} \le \theta_k - \mathbb{E}\left[\tfrac{1}{1+e^{\Delta_k}}\right] \mu_{\mathcal{D}} \alpha_k g_k + \mathbb{E}\left[\tfrac{L}{2+2e^{\Delta_k}}\right] \alpha_k^2. \tag{21}$$

Note, that for the special case in which the noise $\Delta = 0$, we obtain

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \le f(x^k) - \frac{\alpha_k |\langle \nabla f(x^k), s_k \rangle|}{2} + \tfrac{L}{4}\alpha_k^2 \|s_k\|_2^2,$$

where one is able to take the expectation and get (10), finalizing the proof for Lemma 11.

We proceed only for Lemma 10 to evaluate the upper and lower limits of $\mathbb{E}\left[\frac{1}{1+e^{\Delta_k}}\right]$.

Using the Jensen's inequality (20), as $\frac{1}{1+y}$ is a convex function for $y > 0$,

$$\mathbb{E}[\tfrac{1}{1+e^{\Delta_k}}] \ge \tfrac{1}{1+\mathbb{E}[e^{\Delta_k}]}. \tag{22}$$

It remains to evaluate the upper limit of $\mathbb{E}[e^{\Delta_k}]$, given that δ is a small random noise. It is not possible to make any statements about the distribution of this noise, that is the reason we evaluate only $\mathbb{E}[e^{\xi}]$.

In accordance with Assumption 9,

$$\mathbb{E}\left[e^{\xi}\right] \leq e^{\sigma^2/2}.$$

The equation (22) can be rewritten in the following form:

$$\mathbb{E}[\tfrac{1}{1+e^{\Delta_k}}] \geq \tfrac{1}{1+\mathbb{E}[e^{\xi+\delta}]} \geq \tfrac{1}{1+e^{\sigma^2/2}e^{\delta}}. \quad (23)$$

The upper limit of $\mathbb{E}\left[\frac{1}{1+e^{\Delta_k}}\right]$ derived from $\frac{1}{1+e^{\Delta_k}} < 1$:

$$\mathbb{E}[\tfrac{1}{1+e^{\Delta_k}}] < 1. \quad (24)$$

Given (23) and (24), it is possible to rewrite the equation (20):

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \leq f(x^k) - \tfrac{1}{1+e^{\sigma^2/2}e^{\delta}} \cdot \mu_{\mathcal{D}}\alpha_k\|\nabla f(x^k)\|_{\mathcal{D}} + \tfrac{L}{2}\alpha_k^2. \quad (25)$$

By taking the expectation in (25), we obtain the equation (8). □

Proof of Theorem 13. As in the analogous theorem for STP, the proof is based on the analysis of the equation (12), rewrote in the following form:

$$\mathbb{E}[\|\nabla f(x^k)\|_{\mathcal{D}}] \leq \tfrac{1}{\mu_{\mathcal{D}}}\left(\tfrac{\theta_k-\theta_{k+1}}{C_k^1\alpha_k} + \tfrac{C_k^2 L}{2C_k^1}\alpha_k\right) = \tfrac{1}{\mu_{\mathcal{D}}}\left(\tfrac{(\theta_k-\theta_{k+1})\sqrt{k+1}}{C_k^1\alpha_0} + \tfrac{C_k^2 L\alpha_0}{2C_k^1\sqrt{k+1}}\right). \quad (26)$$

From (13) and the boundness of f: $f^* \leq \theta_{k+1} \leq \theta_k \leq f(x_0)$ for all k. Choosing $l = \lfloor K/2 \rfloor$, we get

$$\sum_{j=l}^{2l}(\theta_j - \theta_{j+1}) = \theta_l - \theta_{2l+1} \leq f(x_0) - f_* := C \Rightarrow \exists j \in \{l, \ldots, 2l\} : \theta_j - \theta_{j+1} \leq C/(l+1).$$

It is now possible to make an estimate of the expectation of the gradient norm.

$$\begin{aligned}
\mathbb{E}[\|\nabla f(x^j)\|_{\mathcal{D}}] &\overset{(26)}{\leq} \tfrac{1}{\mu_{\mathcal{D}}}\left(\tfrac{(\theta_j-\theta_{j+1})\sqrt{j+1}}{C_k^1\alpha_0} + \tfrac{C_k^2 L\alpha_0}{2C_k^1\sqrt{j+1}}\right) \leq \tfrac{1}{\mu_{\mathcal{D}}}\left(\tfrac{C\sqrt{j+1}}{C_k^1\alpha_0(l+1)} + \tfrac{C_k^2 L\alpha_0}{2C_k^1\sqrt{j+1}}\right) \\
&\leq \tfrac{1}{\mu_{\mathcal{D}}}\left(\tfrac{C\sqrt{2l+1}}{C_k^1\alpha_0(l+1)} + \tfrac{C_k^2 L\alpha_0}{2C_k^1\sqrt{l+1}}\right) \leq \tfrac{1}{\mu_{\mathcal{D}}\sqrt{l+1}}\left(\tfrac{\sqrt{2}C}{C_k^1\alpha_0} + \tfrac{C_k^2 L\alpha_0}{2C_k^1}\right) \\
&\leq \tfrac{1}{\mu_{\mathcal{D}}\sqrt{K/2}}\left(\tfrac{\sqrt{2}C}{C_k^1\alpha_0} + \tfrac{C_k^2 L\alpha_0}{2C_k^1}\right) \overset{(14)}{\leq} \varepsilon.
\end{aligned}$$

□

Proof of Theorem 15. If $g_k \leq \varepsilon$ for some $k \leq k(\varepsilon)$, then the theorem is proven.

If not, assume by contradiction that $g_k > \varepsilon$ for all $k \leq k(\varepsilon)$. We use (12)

$$\theta_{k+1} \leq \theta_k - C_k^1 \cdot \mu_{\mathcal{D}}\alpha g_k + C_k^2 \cdot \tfrac{L}{2}\alpha^2,$$

where $\theta_k = \mathbb{E}\left[f(x_k)\right]$ and $g_k = \mathbb{E}\left[\|\nabla f(x_k)\|_{\mathcal{D}}\right]$. Consequently,

$$f^* \leq \theta_{K+1} < \theta_0 - (K+1)\left(C_k^1\mu_{\mathcal{D}}\alpha_0\varepsilon - \tfrac{C_k^2 L}{2}\alpha_0^2\right) \overset{(16)}{\leq} \theta_0 - (f(x_0) - f^*) = f^*,$$

which is a contradiction. □

Proof of Theorem 16. In order to proceed, we substitute (17) into the equation (12):

$$\theta_{k+1} \leq \theta_k - \tfrac{C_k^1\mu_{\mathcal{D}}\alpha}{R_0}(\theta_k - f(x^*)) + \tfrac{C_k^2 L}{2}\alpha^2. \quad (27)$$

Let $r_k = \theta_k - f(x^*)$ and $c = 1 - \frac{C_k^1 \mu_{\mathcal{D}} \alpha}{R_0} \in (0, 1)$. Subtracting $f(x^*)$ from both sides of the equation (27), we obtain

$$\begin{aligned} r_K &\leq c r_{K-1} + \tfrac{C_k^2 L}{2}\alpha^2 \quad \leq \quad c^K r_0 + \tfrac{C_k^2 L}{2}\alpha^2 \sum_{i=0}^{K-1} c^i \\ &\leq \exp(-C_k^1 \mu_{\mathcal{D}} \alpha K / R_0) r_0 + \tfrac{C_k^2 L \alpha^2}{2(1-c)} \quad = \quad \exp(-C_k^1 \mu_{\mathcal{D}} \alpha K / R_0) r_0 + \tfrac{\varepsilon}{2} \overset{(18)}{\leq} \varepsilon. \end{aligned}$$

□

Proof of Theorem 17. Let us substitute (17) into the equation (11), and then substrate $f(x^*)$ from both sides:

$$\mathbb{E}\left[f(x^{k+1}) \mid x^k\right] - f(x^*) \leq f(x^k) - f(x^*) - C_k^1 \mu_{\mathcal{D}} \alpha_k \tfrac{f(x^k) - f(x^*)}{R_0} + \tfrac{C_k^2 L}{2}\alpha_k^2.$$

Let $r_k = \mathbb{E}\left[f(x^k)\right] - f(x^*)$.

By utilising the selected value of α_k in the preceding equation and subsequently calculating the expectation, the following result is obtained: $r_{k+1} \leq r_k - \left(\frac{C_k^1 \mu_{\mathcal{D}} \alpha_0}{R_0} - \frac{C_k^2 L \alpha_0^2}{2}\right) r_k^2 = r_k - a r_k^2$. Therefore,

$$\tfrac{1}{r_{k+1}} - \tfrac{1}{r_k} = \tfrac{r_k - r_{k+1}}{r_k r_{k+1}} \geq \tfrac{r_k - r_{k+1}}{r_k^2} \geq a \Rightarrow \tfrac{1}{r_k} \geq \tfrac{1}{r_0} + ka \Rightarrow r_k \leq \tfrac{1}{\frac{1}{r_0} + ka}.$$

For $k \geq \frac{1}{a}\left(\frac{1}{\varepsilon} - \frac{1}{r_0}\right)$ we have $r_k \leq \frac{1}{\frac{1}{r_0} + ka} \leq \varepsilon$. □

Proof of Theorem 18. From Lemmas 7 and 10 we have

$$\mathbb{E}\left[f(x^{k+1}) | x^k\right] \leq f(x^k) - C_k^1 \alpha_k |\langle \nabla f(x^k), s_k \rangle| + \tfrac{C_k^2 L \alpha_k^2}{2}. \tag{28}$$

In $\alpha_k^{\text{opt}} = \frac{C_k^1 |\langle \nabla f(x^k), s_k \rangle|}{C_k^2 L}$, that minimizes the right-hand side of (28), we can replace unknown $\nabla f(x^k)$ with the the directional derivative of f, which can be approximated by finite differences of the function values of f.

$$\alpha_k = \tfrac{C_k^1 |f(x^k + t s_k) - f(x^k)|}{C_k^2 L t} = \tfrac{C_k^1 |\langle \nabla f(x^k), s_k \rangle|}{C_k^2 L} + \tfrac{C_k^1 |f(x^k + t s_k) - f(x^k)|}{C_k^2 L t} - \tfrac{C_k^1 |\langle \nabla f(x^k), s_k \rangle|}{C_k^2 L} := \alpha_k^{\text{opt}} + \delta_k.$$

Therefore,

$$\begin{aligned} \mathbb{E}\left[f(x^{k+1}) | x^k\right] &\leq f(x^k) - \tfrac{(C_k^1)^2 |\langle \nabla f(x^k), s_k \rangle|^2}{C_k^2 L} - C_k^1 \delta_k |\langle \nabla f(x^k), x^k \rangle| + \tfrac{(C_k^1)^2 |\langle \nabla f(x^k), x^k \rangle|^2}{2 C_k^2 L} \\ &\quad + C_k^1 \delta_k |\langle \nabla f(x^k), x^k \rangle| + \tfrac{C_k^2 L}{2}(\delta_k)^2 \\ &= f(x^k) - \tfrac{(C_k^1)^2 |\langle \nabla f(x^k), s_k \rangle|^2}{2 C_k^2 L} + \tfrac{C_k^2 L}{2}(\delta_k)^2. \end{aligned}$$

Using L-smoothness of f:

$$\begin{aligned} |\delta_k| &= \tfrac{1}{Lt}\left| |f(x^k + t s_k) - f(x^k)| - |\langle \nabla f(x^k), t s_k \rangle| \right| \\ &\leq \tfrac{1}{Lt}\left| f(x^k + t s_k) - f(x^k) - \langle \nabla f(x^k), t s_k \rangle \right| \leq \tfrac{1}{Lt} \cdot \tfrac{L}{2}\|t s_k\|_2^2 = \tfrac{t}{2}. \end{aligned}$$

Then

$$\mathbb{E}\left[f(x^{k+1}) | x^k\right] \leq f(x^k) - \tfrac{(C_k^1)^2 \left|\langle \nabla f(x^k), s_k \rangle\right|^2}{2 C_k^2 L} + \tfrac{C_k^2 L t^2}{8}.$$

By taking the mathematical expectation for all randomnesses from the previous inequality, we obtain

$$\begin{aligned} \underbrace{\mathbb{E}[f(x^{k+1})] - f^*}_{r_{k+1}} &\overset{(*)}{\leq} \underbrace{\mathbb{E}[f(x^k)] - f^*}_{r_k} - \tfrac{(C_k^1)^2 \mu_{\mathcal{D}}^2}{2 C_k^2 L}\mathbb{E}[\|\nabla f(x^k)\|_{\mathcal{D}}^2] + \tfrac{C_k^2 L t^2}{8} \\ &\overset{(**)}{\leq} r_k - \tfrac{(C_k^1)^2 \mu_{\mathcal{D}}^2}{2 C_k^2 L R_0^2} r_k^2 + \tfrac{C_k^2 L t^2}{8}, \end{aligned}$$

where $(*)$ is due to tower property of mathematical expectation and (1):

$$\begin{aligned}\mathbb{E}[|\langle \nabla f(x^k), s_k\rangle|^2] &= \mathbb{E}\left[\mathbb{E}[|\langle \nabla f(x^k), s_k\rangle|^2 \mid x^k]\right] \geq \mathbb{E}\left[\left(\mathbb{E}[|\langle \nabla f(x^k), s_k\rangle| \mid x^k]\right)^2\right] \\ &\overset{(1)}{\geq} \mu_{\mathcal{D}}^2 \mathbb{E}[\|\nabla f(x^k)\|_{\mathcal{D}}^2];\end{aligned}$$

$(**)$ follows from the Assumption 5: $\mathbb{E}[\|\nabla f(x^k)\|_{\mathcal{D}}^2] \geq \frac{\mathbb{E}\left[\left(f(x^k)-f^*\right)^2\right]}{R_0^2} \geq \frac{\left(\mathbb{E}\left[f(x^k)-f^*\right]\right)^2}{R_0^2} = \frac{r_k^2}{R_0^2}$.

Considering the monotonicity of $\{f(x^k)\}_{k\geq 0}$:

$$\frac{1}{r_{k+1}} - \frac{1}{r_k} \geq \frac{r_{k+1}-r_k}{r_k r_{k+1}} \geq \frac{\frac{(C_k^1)^2\mu_{\mathcal{D}}^2}{2C_k^2 L R_0^2} r_k^2 - \frac{C_k^2 L t^2}{8}}{r_k^2} \geq \frac{(C_k^1)^2\mu_{\mathcal{D}}^2}{2C_k^2 L R_0^2} - \frac{C_k^2 L}{8}\left(\frac{t}{r_k}\right)^2.$$

If $k \leq K-1$ and $0 < t \leq \frac{\sqrt{2}C_k^1 \mu_{\mathcal{D}} r_{K-1}}{C_k^2 L R_0}$, then

$$\frac{1}{r_{k+1}} - \frac{1}{r_k} \geq \frac{(C_k^1)^2\mu_{\mathcal{D}}^2}{4C_k^2 L R_0^2} = a,$$

since $r_k \leq r_{K-1}$. Finally, $\frac{1}{r_k} \geq \frac{1}{r_0} + ka \Rightarrow r_k \leq \frac{1}{\frac{1}{r_0}+ka}$ for all $k \leq K$. Thus, if $K \geq \frac{1}{a}\left(\frac{1}{\varepsilon} - \frac{1}{r_0}\right)$, then $r_K \leq \frac{1}{\frac{1}{r_0}+Ka} \leq \varepsilon$. □

Proof of Theorem 19. By introducing the variable α_k into the equation (11) and then applying the substrate $f(x^*)$ to both sides, we obtain

$$\begin{aligned}\mathbb{E}[f(x^{k+1}) \mid x^k] - f(x^*) &\leq f(x^k) - f(x^*) - \frac{(C_k^1)^2\mu_{\mathcal{D}}^2\sqrt{2\lambda(f(x^k)-f(x^*))}\|\nabla f(x^k)\|_{\mathcal{D}}}{C_k^2 L} \\ &\quad + \frac{(C_k^1)^2\mu_{\mathcal{D}}^2\lambda(f(x^k)-f(x^*))}{C_k^2 L}.\end{aligned}$$

f is strongly convex, then $\|\nabla f(x^k)\|_{\mathcal{D}}^2 \geq 2\lambda(f(x^k)-f(x^*))$ and

$$\begin{aligned}\mathbb{E}[f(x^{k+1})|x^k] - f(x^*) &\leq f(x^k) - f(x^*) - \frac{2(C_k^1)^2\mu_{\mathcal{D}}^2\lambda(f(x^k)-f(x^*))}{C_k^2 L} + \frac{(C_k^1)^2\mu_{\mathcal{D}}^2\lambda(f(x^k)-f(x^*))}{C_k^2 L} \\ &\leq f(x^k) - f(x^*) - \frac{(C_k^1)^2\mu_{\mathcal{D}}^2\lambda(f(x^k)-f(x^*))}{C_k^2 L}.\end{aligned}$$

Let $r_k = \mathbb{E}\left[f(x^k)\right] - f(x^*)$. By taking the expectation of the last inequality we get $r_{k+1} \leq \left(1 - \frac{(C_k^1)^2\mu_{\mathcal{D}}^2\lambda}{C_k^2 L}\right) r_k$, and therefore

$$r_k \leq \left(1 - \frac{(C_k^1)^2\mu_{\mathcal{D}}^2\lambda}{C_k^2 L}\right)^k r_0.$$

Hence if K satisfies (19), we get $r_K \leq \varepsilon$. □

A.2 The discussion on the number of iterations required

Let us calculate the number of iterations for the STP method.

The stepsize α_k is such that $\arg\min\{f(x^k), f(x^k \pm \alpha_k)\}$ is equal to $x^k - \alpha_k$ as long as $x^k \leq \alpha_k$. Let us write this as an inequality:

$$f\left(x_0 - \sum_{k=0}^{K}\alpha_k\right) \leq \varepsilon, \qquad \sum_{k=0}^{K}\alpha_k \leq x_0. \tag{29}$$

The second inequality in (29) is required to verify that the method does not bypass the solution.

Let us estimate $\sum\limits_{k=0}^{K}\alpha_k$ on both sides:

$$2\sqrt{K+2} - 2 = \int_0^{K+1}\frac{1}{\sqrt{x+1}}dx \leq \sum_{k=0}^{K}\alpha_k \leq \int_1^{K+1}\frac{1}{\sqrt{x}}dx + 1 = 2\sqrt{K+1} - 1. \tag{30}$$

Now we can take that $K = 31$.

Next, we find the number of iterations for the initial modification. The estimation algorithm is analogous to the algorithm for the STP method, but, in this instance, the variable $\widetilde{p}_k$ is introduced. This is necessary to indicate whether the method is taking a step or remaining in a fixed position:

$$\widetilde{p}_k = \begin{cases} 1, \text{ with probability } p, \\ 0, \text{ with probability } 1 - p \end{cases}$$

Then similar to the equation (29)

$$f\left(x_0 - \sum_{k=0}^{K} \widetilde{p}_k \alpha_k\right) \le \varepsilon, \qquad \sum_{k=0}^{K} \widetilde{p}_k \alpha_k \le x_0. \tag{31}$$

It can be observed that if all the values of $\widetilde{p}_k$ are set to the same value, p, the result is the same. The equation (31) is transformed into the following equation:

$$f\left(x_0 - p\sum_{k=0}^{K} \alpha_k\right) \le \varepsilon.$$

Using the equation (30), we find the required number of iterations for the initial modification of the oracle for STP.

For the last two modifications we get

$$f\left(x_0 - \sum_{k=0}^{K} \widetilde{p}_k \alpha_k\right) \le \varepsilon, \qquad \sum_{k=0}^{K} \widetilde{p}_k \alpha_k \le x_0, \tag{32}$$

where

$$\widetilde{p}_k = \begin{cases} 1, \text{ with probability } p_k, \\ 0, \text{ with probability } 1 - p_k \end{cases}$$

The variable p_k depends on the iteration:

$$p_k = \tfrac{1}{1+\exp\left(f(x^k - \alpha_k) - f(x^k)\right)} \qquad \text{or} \qquad p_k = \tfrac{1}{1+\exp\left(f(x^k - \alpha_k) - f(x^k) + \Delta_k\right)},$$

respectively.

For the first formulation, let us consider the special case of the step

$$\begin{aligned}\mathbb{E}[x_{k+1}|x_k] &= (1 - p_k)x_k + p_k(x_k - \alpha_k) \\ &= x_k - \tfrac{1}{\sqrt{1+k}} \tfrac{1}{(1+\exp\left((x_k-\alpha_k)^2 - x_k^2\right))} = x_k - \tfrac{1}{\sqrt{1+k}(1+\exp\left(\frac{1}{1+k} - \frac{2x_k}{\sqrt{1+k}}\right))}.\end{aligned}$$

In order to evaluate the second summand from above, we evaluate x_k from below, $\widetilde{x}_k = x_0 - \sum\limits_{n=0}^{k} \frac{1}{\sqrt{1+n}}$, for the usual STP. Next, we use the same evaluation by the integral. Thus we obtain:

$$x_k \ge x_0 - \sum_{n=0}^{k} \frac{1}{\sqrt{1+n}} > x_0 - \int\limits_{1}^{k+1} \tfrac{dx}{\sqrt{x}} - 1 = x_0 - 2\sqrt{k+1} + 1$$

Now we get the final result:

$$\mathbb{E}[x_{k+1}|x_k] > x_k - \frac{1}{\sqrt{1+k}(1 + \exp\left(\frac{1}{1+k} - \frac{2(x_0+1)}{\sqrt{1+k}} + 4\right))}$$

Thus obtained an estimate for one iteration such that the second summand is independent of x_k. This will allow us to estimate the required number of iterations:

If $\quad f\left(x_0 - \sum_{k=0}^{K} \frac{1}{\sqrt{1+k}(1+\exp\left(\frac{1}{1+k} - \frac{2(x_0+1)}{\sqrt{1+k}} + 4\right))}\right) \leq \varepsilon,$

then $\quad \mathbb{E}(f(x_0 - \sum_{k=0}^{K} \widetilde{p}_k \alpha_k)) \leq \varepsilon.$

In the case of noise, we exploit the fact that this random variable is distributed according to the sub-Gaussian distribution, which introduces an additional summand in the exponent in the aforementioned setting:

$$\mathbb{E}[x_{k+1}|x_k] > x_k - \frac{1}{\sqrt{1+k}(1+\exp\left(\frac{1}{1+k} - \frac{2(x_0+1)}{\sqrt{1+k}} + 4 + \sigma^2/2 + \delta\right))}$$

This subsection concludes with the last result:

If $\quad f\left(x_0 - \sum_{k=0}^{K} \frac{1}{\sqrt{1+k}(1+\exp\left(\frac{1}{1+k} - \frac{2(x_0+1)}{\sqrt{1+k}} + 4 + \sigma^2/2 + \delta\right))}\right) \leq \varepsilon,$

then $\quad \mathbb{E}(f(x_0 - \sum_{k=0}^{K} \widetilde{p}_k \alpha_k)) \leq \varepsilon.$

We also consider the simpler case where the step size α_k is taken from the theorem for strongly convex target functions (see Subsection 4.3). Given that $\alpha_k = \frac{C_k^1 \mu_{\mathcal{D}}}{C_k^2 L} \sqrt{2\lambda(f(x^k) - f(x^*))}$ (see Theorem 19), moreover, in our case $\alpha_k \sim x^k$, it can be demonstrated that the STP will converge in two iterations. Furthermore, if a step is taken with some probability, then the number of iterations required to achieve accuracy ε is inversely proportional to the probability of taking the step.

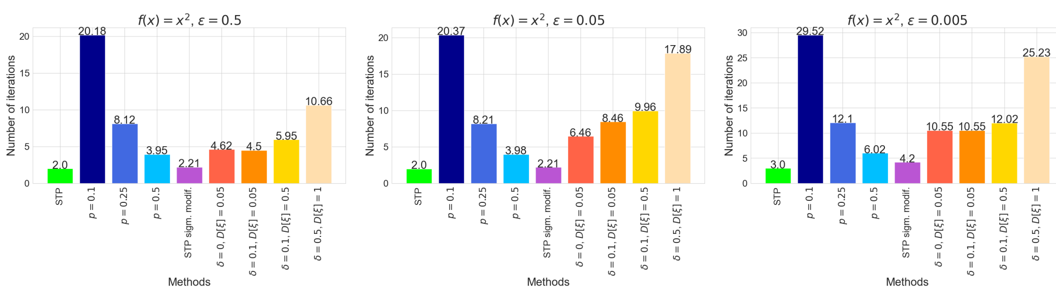

Figure 3: The number of iterations needed to achieve accuracy ε, α_k for SC problem

A.3 The impact of the sigmoid probability

The impact of the sigmoid is illustrated in the plot below. It can be observed that as the degree exponent decreases, the probability of undertaking a method step increases.

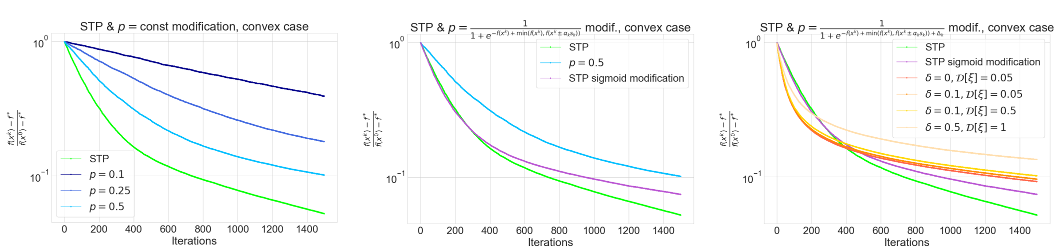

Figure 4: The impact of the sigmoid probability function

A.4 SOME ADDITIONAL RESULTS FOR THE PRINCIPAL MODIFICATION OF THE ORACLE FOR STP

In order to provide some estimates on the mathematical expectation of $\frac{1}{1+e^{\Delta_k}}$, we attempt to make assumptions about the distribution of the random variable Δ_k. This is done in accordance with the principles set out in Lemma 10.

If Δ is uniformly distributed on the segment $[-a, a]$, $a \in \mathbb{R}$:

$$\mathbb{E}\left[\frac{1}{1+e^x}\right] = \int\limits_{-\infty}^{\infty} \frac{1}{1+e^x} \frac{\mathbb{1}_{[-a,a]}(x)}{2a} dx = \frac{1}{2a} \int\limits_{-a}^{a} \frac{1}{1+e^x} dx = \frac{a-\ln(e^a+1)+a+\ln(e^{-a}+1)}{2a} = \frac{1}{2}.$$

If Δ is normally distributed:

$$\mathbb{E}\left[\frac{1}{1+e^x}\right] = \int\limits_{-\infty}^{\infty} \frac{1}{1+e^x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{x}{\sigma})^2} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{\infty} \frac{dx}{e^{x^2/2\sigma^2}(1+e^x)} = \frac{1}{\sqrt{2\pi\sigma^2}} \sqrt{\frac{\pi\sigma^2}{2}} = \frac{1}{2}.$$

With all of the above written, let us rewrite equation (20).

Claim 21. If the variable Δ is uniformly distributed on the segment $[-a, a]$ or if it is normally distributed, then the equation (7) from the Lemma 10 takes the following form:

$$\mathbb{E}\left[f(x^{k+1})|x^k\right] \leq f(x^k) - \frac{1}{2}\mu_{\mathcal{D}}\alpha_k\|\nabla f(x^k)\|_{\mathcal{D}} + \frac{L}{4}\alpha_k^2. \tag{33}$$

Moreover, if Δ random variable is distributed in such a way that its distribution function is symmetric, as, for example, for normal or uniform distributions, then for it equation (20) takes the form of equation (33).