Accelerated Simultaneous Perturbation Stochastic Approximation for Tracking Under Unknown-but-Bounded Disturbances

Victoria Erofeeva, Oleg Granichin, Munira Tursunova, Anna Sergeenko, and Yuming Jiang

Abstract— **In this paper, we propose an accelerated version of Simultaneous Perturbation Stochastic Approximation (Accelerated SPSA). This algorithm belongs to the class of methods used in derivative-free optimization and has proven efficacy in the problems including significant non-statistical uncertainties. We focus on analysis of Accelerated SPSA in a non-stationary setting and consider the presence of unknown-but-bounded disturbances. Research on these problems covers many directions. However, in large-scale systems, efficiency still remains a concern. It gave rise to the research where acceleration represents an objective in the algorithm's design. This problem motivated us to extend our previous research on SPSA in the direction of acceleration. We show that the proposed new accelerated version converges faster than the initial one. The validation of the algorithm is preformed in a target tracking problem.**

I. INTRODUCTION

Derivatives of a cost function are often not available in many modern optimization problems arising in signal processing, machine learning, control, and other fields. There are two possible reasons for that. First, the function may be represented by a black-box or simulation oracle as in reinforcement learning [1]. Second, it may be difficult or impractical to evaluate the gradient and/or higher order derivatives due to significant uncertainties in measurements. The described cases have increased the interest in the development of methods that doesn't rely on derivatives, i.e. derivative-free, or zeroth-order, optimization [2].

Most real-world problems include different kinds of uncertainties, e.g., noisy measurements, external disturbances or attacks. A key class of methods for derivative-free optimization under uncertainties is stochastic approximation [3], [4]. The first versions of derivative-free stochastic approximation based algorithms require substantial computational effort per iteration, which makes them undesirable for large-scale applications. In [5], the author proposed Simultaneous perturbation stochastic approximation (SPSA). The important feature of SPSA is the underlying gradient approximation that requires only two loss function measurements and does not depend on the number of parameters being optimized. The other approaches to stochastic optimization include direct-search and model based methods [6]. Usually such algorithms are applicable to problems involving only zero mean noise. Whereas in [7], it has been shown that SPSA converges in the presence of arbitrary unknown-but-bounded noise.

A fundamental assumption in stochastic approximation which has been widely adopted is that the cost function does not change throughout the time horizon. At the other hand, non-stationary systems frequently appear in practice [8]. Research on non-stationary problems covers many directions including unconstrained first-order optimization [9], stochastic [7], [10] and online convex optimization [11]. However, in large-scale problems, efficiency still remains a concern. It gave rise to the research directions where acceleration represents an objective in the algorithm's design.

In [12], the authors cover the recent advances on acceleration techniques used in convex optimization. One of the tools used to accelerate the convergence of optimization methods is momentum. Momentum-based methods were first formally studied by Polyak [13], and found their application in many practical problems. Despite the clear intuition behind the momentum methods, the proposed analysis doesn't apply for all general convex cost functions. A different approach based on algebraic arguments was proposed by Nesterov. He developed the method of estimate sequences to verify the momentum-based accelerated methods [14]. In [15], we analyzed a modified accelerated stochastic gradient method proposed for non-stationary optimization problem and built our analysis based on bounded estimate sequences.

In this paper, we continue this line of work and propose Accelerated SPSA for tracking under unknown-but-bounded noise. Overall, the contributions are as follows:

- we propose a new version of SPSA method equipped by the acceleration scheme presented in [15]. The proposed method belongs to the class of zeroth-order methods and requires additional analysis in comparison to the previous work. We obtained a bound on gradient estimates and modified the acceleration scheme based on the new results;
- previously, we considered noisy gradient measurements with the assumption that the noise has zero-mean and a known variance. In this work, we relax the assumption regarding the noise appearing in the measurements of zeroth-order oracle;
- finally, we validate the new method in a target tracking

The research was carried out within the Institute of Problems of Mechanical Engineering RAS grant supported by Russian Science Foundation (project No. 21-19-00516).

V. Erofeeva is with Skolkovo Institute of Science and Technology, Moscow, Russia v.erofeeva@skoltech.ru

O. Granichin, A. Sergeenko, M. Tursunova are with Saint Petersburg State University (Faculty of Mathematics and Mechanics, Research Laboratory for Analysis and Modeling of Social Processes) and with Institute of Problems of Mechanical Engineering RAS, St. Petersburg, Russia o.granichin@spbu.ru, a.sergeenko@spbu.ru, tursunovamunira@gmail.com

Y. Jiang is with Department of Information Security and Communication Technology, Norwegian University of Science and Technology (NTNU), NO-7491, Trondheim, Norway, e-mail: jiang@ntnu.no.

problem and show the improvement in the convergence.

The paper is organized as follows. The preliminary information is given in Section II. A formal problem setting of a time-varying mean-risk optimization and an example illustrating this problem are given in Section III. The main result including assumptions, the proposed accelerated SPSA algorithm for tracking, and its convergence properties are presented in Section IV. In Section V, the efficiency of the proposed algorithm is illustrated through the numerical simulation. Section VI concludes the paper.

II. PRELIMINARIES

Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space corresponding to sample space Ω, set of all events $\mathcal{F}$, and probability measure P. $\mathbb{E}$ denotes mathematical expectation. Let $\mathcal{F}_{t-1}$ be the σ-algebra of all probabilistic events which happened before time instant $t = 1, 2, \ldots$, $\mathbb{E}_{\mathcal{F}_{t-1}}$ denotes the conditional mathematical expectation with respect to the σ-algebra $\mathcal{F}_{t-1}$.

III. PROBLEM STATEMENT

A. Non-stationary Mean-Risk Optimization

Consider a mean-risk optimization problem:

$$\min_{\theta}\{F(\theta) = \mathbb{E}_{\mathcal{F}_{t-1}} f(\theta, \xi_t)\}, \tag{1}$$

where $\theta \in \mathbb{R}^d$ is a decision vector, ξ_t is uncertainty that belongs to set Ξ. Subsequently, we replace the notation $f(\theta, \xi_t)$ by $f_{\xi_t}(\cdot)$ emphasising that ξ_t is uncontrollable sequence. The problem (1) arises in many practical applications as well as in machine learning. The sources of uncertainty include but not limited to: estimation errors, i.e., optimization based on measured/estimated data; prediction errors, i.e., part of data doesn't exist at the moment of optimization (e.g. future demand/prices); implementation errors, i.e., discretization and model approximation errors. The uncertainty is represented by a non-controllable deterministic sequence (e.g., $\Xi = \mathbb{N}$ and $\xi_t = t$) or random sequence. In the latter case we assume that a probability distribution of ξ_t exists and may be known or unknown.

In this work, we consider zeroth-order optimization, where we have only noisy measurements of function to be optimized. In contrast to a majority of existing research, we don't have any statistical assumptions on this noise. We also assume that parameter θ cannot be directly measured. Hence, we introduce a sequence of measurement points $\mathbf{x}_1, \mathbf{x}_2, \ldots$ chosen according to an observation plan. The values $y_1, y_2, \ldots$ of the functions $f_{\xi_t}(\cdot)$ are observable at every time instant $t = 1, 2, \ldots$ with additive external *unknown-but-bounded* noise v_t

$$y_t = f_{\xi_t}(\mathbf{x}_t) + v_t. \tag{2}$$

We also assume that minimizer θ of $F(\theta)$ may vary over time. Formally, the *non-stationary mean-risk optimization problem* is as follows: estimate the time-varying minimum point θ_t of function

$$F_t(\theta) = \mathbb{E}_{\mathcal{F}_{t-1}} f_{\xi_t}(\theta) \to \min_{\theta_t}. \tag{3}$$

In the next subsection, we present an example illustrating the considered problem statement.

B. Example

Given a network consisting of $n = 3$ planar sensors identified by $i \in \mathcal{N} = \{1, 2, 3\}$. The state of sensor i is $\mathbf{s}^i \in \mathbb{R}^2$. We assume that the states are known and doesn't depend on time, i.e. the sensors are stationary. In the sensing range of the sensors, there are $m = 6$ moving planar targets identified by $l \in \mathcal{M} = \{1, 2, \ldots, 6\}$. The goal of each sensor i is to estimate the states of all targets $\mathbf{r}_t^l \in \mathbb{R}^2$ at time instant t.

Let $\theta_t = \mathrm{col}(\mathbf{r}_t^1, \ldots, \mathbf{r}_t^6) \in \mathbb{R}^{12}$ be the common state vector of all targets, $\hat{\theta}_t = \mathrm{col}(\hat{\mathbf{r}}_t^1, \ldots, \hat{\mathbf{r}}_t^6)$ be a common vector of estimates. Each target $l \in \mathcal{M}$ changes the position according to the following dynamics:

$$\mathbf{r}_t^l = \mathbf{r}_{t-1}^l + \zeta_{t-1}^l, l \in \mathcal{M}, \tag{4}$$

where ζ_{t-1}^l are random vectors uniformly distributed in a ball. We assume that at time instant t sensor i is able to measure the squared distance $\rho_t^{i,l} = \rho(\mathbf{s}^i, \mathbf{r}_t^l) = \|\mathbf{r}_t^l - \mathbf{s}^i\|^2$ to some moving target $\mathbf{r}_t^l$.

Suppose sensor i estimates the state of target l at time instant t. The sensor is able to collect the distances to the same target measured by its neighbors $j \in \mathcal{N}^i$. Denote by

$$\bar{\rho}_t^{i,j}(l) = \rho(\mathbf{s}^i, \mathbf{r}^l) - \rho(\mathbf{s}^j, \mathbf{r}^l), \ \forall j \in \mathcal{N}_t^i \tag{5}$$

a residual between a measurement of sensor i and its neighbor j for target l.

In this case, using the square difference formula we get the equations

$$\bar{\rho}_t^{i,j}(l) = (\mathbf{s}^j - \mathbf{s}^i)^{\mathrm{T}}(2\mathbf{r}^l - \mathbf{s}^j - \mathbf{s}^i), j \in \mathcal{N}_t^i.$$

This allows us to derive $C^{i,l}\mathbf{r}_t^l = D^{i,l}, \quad \mathbf{r}_t^l = [C^{i,l}]^{-1}D^{i,l}$, where $j_1, \ldots, j_{\bar{d}} \in \mathcal{N}_t^i$ and

$$C^{i,l} = 2\begin{bmatrix}(\mathbf{s}^{j_1} - \mathbf{s}^i)^{\mathrm{T}} \\ \cdots \\ (\mathbf{s}^{j_{\bar{d}}} - \mathbf{s}^i)^{\mathrm{T}}\end{bmatrix}, D^{i,l} = \begin{bmatrix}\bar{\rho}_t^{i,j_1}(l) + \|\mathbf{s}^{j_1}\|^2 - \|\mathbf{s}^i\|^2 \\ \cdots \\ \bar{\rho}_t^{i,j_{\bar{d}}}(l) + \|\mathbf{s}^{j_{\bar{d}}}\|^2 - \|\mathbf{s}^i\|^2\end{bmatrix}.$$

Each sensor sends the final measurements to a fusion center, which solves the non-stationary optimization problem based on observations:

$$y_t^{i,l} = \|\hat{\mathbf{r}}_t^l - [C^{i,l}]^{-1}D^{i,l}\|^2 + v_t^{i,l}, \tag{6}$$

where $v_t^{i,l}$ is the unknown-but-bounded noise.

The target tracking problem described above is similar to one published in [16]. The difference is that in this paper we have a hybrid system instead of a distributed one, i.e., we use both the distributed and centralized steps.

IV. MAIN RESULT

In this section, we present the main result of this paper. First, we formulate the assumptions regarding functions $F_t(\mathbf{x})$, $f_{\xi_t}(\mathbf{x})$. Then, we describe a new accelerated version of SPSA that we propose in this paper.

Assumption 1: The function $F_t(\cdot)$ is strongly convex, it has minimum point θ_t and

$$\forall \mathbf{x} \in \mathbb{R}^d \ \ \langle \mathbb{E}_{\mathcal{F}_{t-1}} \nabla f_{\xi_t}(\mathbf{x}), \mathbf{x} - \theta_t \rangle \geq \mu \|\mathbf{x} - \theta_t\|^2.$$

Assumption 2. $\forall \xi \in \Xi$, the gradient $\nabla f_\xi(\mathbf{x})$ satisfies the Lipschitz condition: $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$

$$\|\nabla f_\xi(\mathbf{x}_1) - \nabla f_\xi(\mathbf{x}_2)\| \leq L\|\mathbf{x}_1 - \mathbf{x}_2\|$$

with constant $L > 0$.

Assumption 3. For every $n \geq 0$ and $\forall \mathbf{x} \in \mathbb{R}^d$, $a, b > 0$, the drift and the gradient are bounded

$$\text{a) } \mathbb{E}_{\mathcal{F}_{n-1}} |f_{\xi_n}(\mathbf{x}) - f_{\xi_{n+1}}(\mathbf{x})| \leq a \mathbb{E}_{\mathcal{F}_{n-1}} \|\nabla f_{\xi_n}(\mathbf{x})\| + b,$$
$$\text{b) } \mathbb{E}_{\mathcal{F}_{t-1}} \|\nabla f_{\xi_t}(\mathbf{x}) - \nabla f_{\xi_{t-1}}(\mathbf{x})\| \leq c,$$
$$\text{c) } \mathbb{E}\|\nabla f_{\xi_t}(\theta_t)\|^2 \leq \delta_f^2.$$

Assumptions 1 and 2 are common in the optimization field. Assumption 3 is used in non-stationary problems.

A. Accelerated SPSA for Tracking

Let $\mathbf{\Delta}_n \in \mathbb{R}^d$, $n = 1, 2, \ldots$ be independent random variable, i.e., *simultaneous test perturbation*, drawn from Bernoulli distribution. Each component of the vector independently takes value $\pm \frac{1}{\sqrt{d}}$ with probability $\frac{1}{2}$.

Remark: we divide the iterative process into blocks and n shows the number of a current block. For example, if $n = 1$, algorithm (7) provides calculations for time instances $t = 2n - 1$ and $t = 2n$ whereas the variables with subscript n are calculated just once and used at the both time instances.

We choose initial estimate $\hat{\theta}_0 \in \mathbb{R}^d$, and parameters $\gamma_0 > 0$, $h > 0$, $\beta > 0$, $\eta \in (0, \mu)$, $\alpha_0 \in (0, 1)$. We also define $z_0 = \hat{\theta}_0$ and $H = h - \frac{h^2 L}{2}\left[(\frac{a}{2\beta} + 1)^2 + \frac{\epsilon^2}{2}\right]$, where $\epsilon > 0$. At each n, we find α_n by solving the equation

$$\alpha_n^2 = 2(\frac{2\beta H}{a} - \frac{\epsilon}{2})((1 - \alpha_n)\gamma_n + \alpha_n(\mu - \eta))$$

and $\gamma_n = (1 - \alpha_{n-1})\gamma_{n-1} + \alpha_{n-1}(\mu - \eta)$.

We consider the algorithm with two observations of functions $f_{\xi_t}(\cdot)$ for constructing sequences of measurement points $\{\mathbf{x}_t\}$ and estimates $\{\widehat{\theta}_t\}$ at $n \geq 1$:

$$\begin{cases} \tilde{\mathbf{x}}_{2n-2} = \frac{1}{\gamma_{n-1} + \alpha_n(\mu - \eta)} \Big(\alpha_n \gamma_{n-1} \mathbf{z}_{2n-2} + \gamma_n \widehat{\theta}_{2n-2}\Big), \\ \mathbf{x}_{2n} = \tilde{\mathbf{x}}_{2n-2} + \beta \mathbf{\Delta}_n, \ \mathbf{x}_{2n-1} = \tilde{\mathbf{x}}_{2n-2} - \beta \mathbf{\Delta}_n, \\ \tilde{\mathbf{x}}_{2n-1} = \tilde{\mathbf{x}}_{2n-2}, \ \widehat{\theta}_{2n-1} = \widehat{\theta}_{2n-2}, \\ \mathbf{g}_{2n} = \mathbf{\Delta}_n \frac{y_{2n} - y_{2n-1}}{2\beta}, \\ \widehat{\theta}_{2n} = \tilde{\mathbf{x}}_{2n-1} - h \mathbf{g}_{2n}, \\ \mathbf{z}_{2n} = \gamma_n^{-1} \Big[(1 - \alpha_n)\gamma_{n-1} \mathbf{z}_{2n-2} + \\ \alpha_n(\mu - \eta)\tilde{\mathbf{x}}_{2n-1} - \alpha_n \mathbf{g}_{2n}\Big]. \end{cases} \tag{7}$$

Remark: If the constants appearing in Assumptions 1-3 are unknown, we can set them to their worst case values.

B. Convergence Analysis

In this section, we provide a convergence analysis of the proposed algorithm. Let us formulate the rest of assumptions, i.e., on noise and random parameters.

Assumption 4. For $n = 1, 2, \ldots$, the successive differences $\tilde{v}_n = v_{2n} - v_{2n-1}$ of noise are bounded: $|\tilde{v}_n| \leq c_v < \infty$, or $\mathbb{E}(\tilde{v}_n)^2 \leq c_v^2$ if sequence $\{\tilde{v}_n\}$ is random.

Assumption 5: For any $n = 1, 2, \ldots$,
a) $\mathbf{\Delta}_n$ and ξ_{2n-1}, ξ_{2n} (if they are random) do not depend on σ-algebra $\mathcal{F}_{2n-2}$.
b) If $\xi_{2n-1}, \xi_{2n}, \tilde{v}_n$ are random, then random vectors $\mathbf{\Delta}_n$ and elements $\xi_{2n-1}, \xi_{2n}, \tilde{v}_n^i$ are independent.
c) $\|\mathbf{\Delta}_n\|^2 \leq c_{\mathbf{\Delta}}^2$. In our case, $\mathbf{\Delta}_n$ takes values $\pm \frac{1}{\sqrt{d}}$ with probability $\frac{1}{2}$, so we have $c_{\mathbf{\Delta}} = 1$.

The following theorem shows an upper bound of the estimation error.

Theorem 1. Let $\{A_n\}$ and $\{Z_n\}$ be the sequences in $\mathbb{R}$ defined as

$$A_0 = 0, \quad A_{n+1} = (1 - \alpha_n)[(1 - \lambda_n)a + A_n],$$
$$Z_0 = 0, \quad Z_{n+1} = (1 - \lambda_{n+1})(b + ac) + A_{n+1}c.$$

If Assumptions 1–5 are hold, algorithm (7) generates a sequence of estimates $\{\widehat{\theta}_n\}_{n=0}^{\infty}$ such that

$$\mathbb{E}_{\mathcal{F}_{n-1}} f_{\xi_n}(\widehat{\theta}_n) - f_{\xi_n}(\theta_n) \leq$$
$$\lambda_n(\phi_0(\theta_0) - f_{\xi_n}(\widehat{\theta}_n) + \Phi) + D_n,$$

where

$$D_0 = 0, \ \ D_{n+1} = (1 - \alpha_n)D_n + (1 - \alpha_n)Z_n +$$
$$\frac{\alpha_n(1 - \alpha_n)\gamma_n(\mu - \eta - 3)}{2\gamma_{n+1}} \|\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|^2 + \tilde{d}.$$

and $\phi_0(\mathbf{x}) = f_0(\widehat{\theta}_0) + \frac{\gamma_0}{2}\|\mathbf{x} - \widehat{\theta}_0\|^2$, $\Phi = \frac{\gamma_0 c^2}{2\mu^2}$, $\lambda_0 = 1$, $\lambda_t \to 0$, $\gamma_0 > 0$.

Proof: The proof is moved to Appendix. The constants are defined in the proof.

V. SIMULATION

In this section, we present a numerical experiment, which illustrates the performance of the suggested algorithm (7). Based on the example presented in Section III-B, we define a distributed network of 3 sensors tracking 6 moving targets. In this case, each sensor may have two or less active communication channels for the information exchange. Each sensor also choose a random target that it tracks at the current time instant.

We've set the following parameters of algorithm (7): $h = 0.08, \beta = 0.1, \eta = 0.95, \alpha_x = 0.1, \gamma_0 = 2.0$, $L = 2, \mu = 2, a = 2, b = 2, c = 1$. The targets start their motion at a position randomly chosen from interval $[0; 100]$. Dynamics of the targets defined in (4). We've defined ζ_t^l as a random vector uniformly distributed on the ball of radius equal to 0.2 for targets with odd identifiers and 0.6 for targets with even identifiers. This means that the targets are heterogeneous and behave differently. The sensors are stationary and their coordinates are random values uniformly distributed in interval $[100;120]$. We consider random type of noise, i.e. uniformly distributed random variable falling within the interval $[-1;1]$.

Let us consider for every target l and sensor i at each time instant t the covariance matrix of residuals $\tilde{\sum}_t^{i,l} \in \mathbb{R}^{d\times d}$ which is represented as a part of the common covariance matrix. In the simulation, the new algorithm is compared with the previous one from [7]. Figure 1 shows the typical behaviour of the averaged diagonal entries of the covariance matrix. Both presented algorithms have the same initial parameters, random values of targets and noises at each iteration. The only difference is the algorithm itself. It is well seen that the new algorithm converges faster than the previous one: while new algorithm is converged approximately by step 100, the old one converges approximately by step 500.

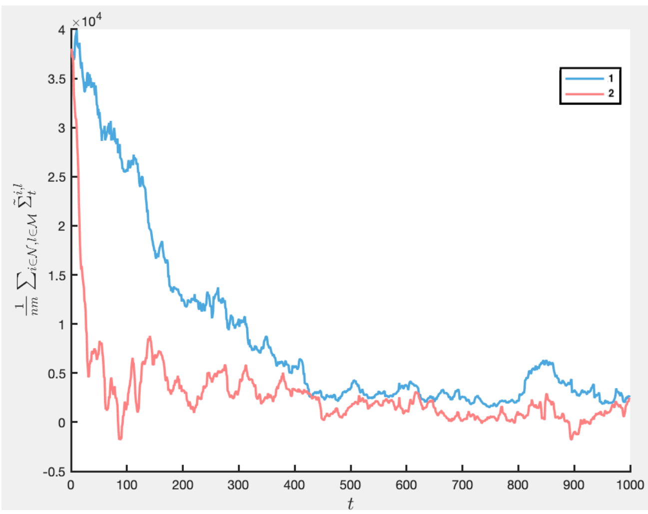

Fig. 1. Typical behaviour of the averaged entries of the covariance matrix. The blue line indicates the algorithm from [7], the red one shows the proposed new accelerated version.

VI. CONCLUSIONS

In this paper, we've proposed the Accelerated SPSA algorithm. The convergence analysis of this algorithm was carried out in non-stationary setting. We've obtained a bound on the variance of gradient estimates and modified the acceleration scheme based on the new results. We've also relaxed the assumption regarding the noise appearing in the measurements of zeroth-order oracle. Finally, we've validated the new method in the target tracking problem and showed the improvement in the convergence.

APPENDIX

In [14], Nesterov introduced a framework of estimate sequence for the development and analysis of accelerated methods. In our previous work [15], we extended it to nonstationary optimization setup. Here, we use the proposed definition of bounded estimate sequence and some lemmas to analyze Accelerated SPSA method.

Let $\bar{\mathcal{F}}_{n-1} = \sigma\{\mathcal{F}_{n-1}, v_{2n-1}, v_{2n}, \xi_{2n-1}, \xi_{2n}, \Delta_n\}$ be the σ-algebra of probabilistic events generated by $\mathcal{F}_{n-1}, v_{2n-1}, v_{2n}, \xi_{2n-1}, \xi_{2n}, \Delta_n$ and $\tilde{\mathcal{F}}_{n-1} = \sigma\{\mathcal{F}_{n-1}, v_{2n-1}, v_{2n}, \xi_{2n-1}, \xi_{2n}\}$ such that

$$\mathcal{F}_{n-1} \subset \tilde{\mathcal{F}}_{n-1} \subset \bar{\mathcal{F}}_{n-1} \subset \mathcal{F}_n.$$

Definition 1 (Bounded Estimate Sequence) [15]

Let $\phi_0(\mathbf{x})$ be a deterministic function and $\phi_t(\mathbf{x})$ be a random function depending on $\mathcal{F}_{t-1}$ for all $t \geq 1$, and $\lambda_t \geq 0$ for all $t \geq 0$. The sequence $\{(\lambda_t, \phi_t(\mathbf{x}))\}_{t=0}^{\infty}$ is called a *bounded estimate sequence* of function $f_{\xi_t}(\mathbf{x})$ if $\lambda_t \to 0$ and there exist a sequence $\{A_t\}_{t=0}^{\infty}$, $A_t \in \mathbb{R}$, and a constant $\Phi < \infty$, and for any $\mathbf{x} \in \mathbb{R}^d$ and for all $t \geq 0$ we have

$$\mathbb{E}_{\mathcal{F}_{t-1}}\phi_t(\mathbf{x}) \leq \mathbb{E}_{\mathcal{F}_{t-1}}[(1-\lambda_t)f_{\xi_t}(\mathbf{x})+ \\ A_t\|\nabla f_{\xi_t}(\mathbf{x})\| + \lambda_t(\tilde{\phi}_{0,t}(\mathbf{x}) + \Phi)], \quad (8)$$

where $\tilde{\phi}_{0,t}(\mathbf{x}) = \phi_0(\mathbf{x}) - \phi_0(\theta_t) + \phi_0(\theta_0)$, $\mathbb{E}_{\mathcal{F}_{-1}}\phi_0(\mathbf{x}) = \phi_0(\mathbf{x})$.

Here we assume $\{\lambda_t\}_{t=0}^{\infty}$ is a deterministic sequence and it doesn't depend on $\mathcal{F}_{t-1}$.

The next Lemma shows how to build the bounded estimate sequences for tracking under unknown-but-bounded noise.

Lemma 1 (Constructing a Bounded Estimate Sequence)

Assume that

1) $\{\mathbf{x}_n\}_{n=0}^{\infty}$ is an arbitrary sequence in $\mathbb{R}^d$,
2) $\phi_0(\mathbf{x})$ is defined as $\phi_0(\mathbf{x}) = \phi_0^* + \frac{\gamma_0}{2}\|\mathbf{x} - \mathbf{x}_0\|^2$,
3) coefficients $\{\alpha_n\}_{n=0}^{\infty}$ satisfy condition $\alpha_n \in [\alpha_x, 1)$,
4) $\eta = \frac{\epsilon^2(2L\beta+c)}{2}$, and $\epsilon > 0$ ensures $\eta \in (0, \mu)$,
5) $\{A_n\}_{n=0}^{\infty}$, $\{Z_n\}_{n=0}^{\infty}$ are sequences in $\mathbb{R}$ defined as

$$A_0 = 0, \quad A_{n+1} = (1-\alpha_n)[(1-\lambda_n)a + A_n],$$
$$Z_0 = 0, \quad Z_{n+1} = (1-\lambda_{n+1})(b+ac) + A_{n+1}c,$$

6) we choose $\Phi = \frac{\gamma_0 c^2}{2\mu^2}$ and $\lambda_0 = 1$.

Then the pair of sequences $\{\phi_n(\cdot)\}_{n=0}^{\infty}$ and $\{\lambda_n\}_{n=0}^{\infty}$ defined by the relations

$$\lambda_{n+1} = (1-\alpha_n)\lambda_n,$$

$$\phi_{n+1}(\mathbf{x}) = (1-\alpha_n)(\phi_n(\mathbf{x}) - Z_n)+ \\ \alpha_n[r(\mathbf{x}_n) + \langle \mathbf{g}_n, \mathbf{x} - \mathbf{x}_n\rangle + \frac{\mu-\eta}{2}\|\mathbf{x} - \mathbf{x}_n\|^2], \quad (9)$$
$$r(\mathbf{x}_n) = f_{\xi_n}(\mathbf{x}_n) - \frac{c^2}{2\eta} - a\|\nabla f_{\xi_n}(\mathbf{x}_n)\| - b$$

are bounded estimate sequences.

Proof of Lemma 1:

Let $\tilde{f}_n = f_{\xi_{2n}}(\mathbf{x}_{2n}) - f_{\xi_{2n-1}}(\mathbf{x}_{2n-1})$. For any $\mathbf{u}, \tilde{\mathbf{u}} \in \mathbb{R}^d$, using Taylor representation of $f_{\xi_t}(\mathbf{x}_t)$ for $t^{\pm} = 2n - \frac{1}{2} \pm \frac{1}{2}$, we obtain

$$f_{\xi_{t^\pm}}(\mathbf{u}) = f_{\xi_{t^\pm}}(\tilde{\mathbf{u}}) + \langle \nabla f_{\xi_{t^\pm}}(\tilde{\mathbf{u}} + \rho_{\xi_{t^\pm}}^{\pm}(\mathbf{u} - \tilde{\mathbf{u}})), \mathbf{u} - \tilde{\mathbf{u}}\rangle, \quad (10)$$

where $\rho_{\xi_{t^\pm}}^{\pm} \in (0,1)$.

Let $\mathbf{u} = \mathbf{x}_{t^\pm} = \tilde{\mathbf{x}}_{2n-2} \pm \beta\mathbf{\Delta}_n$ and $\tilde{\mathbf{u}} = \tilde{\mathbf{x}}_{2n-2}$. Based on (10), we get the following representation of difference $\tilde{f}_n$

$$\begin{gathered}\tilde{f}_n = f_{\xi_{2n}}(\mathbf{x}_{2n}) - f_{\xi_{2n-1}}(\mathbf{x}_{2n-1}) = \\ f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}) + \\ \langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2} + \rho^+_{\xi_{2n}}\beta\mathbf{\Delta}_n), \beta\mathbf{\Delta}_n\rangle + \\ \langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2} - \rho^-_{\xi_{2n-1}}\beta\mathbf{\Delta}_n), \beta\mathbf{\Delta}_n\rangle - \\ 2\langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle + 2\langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle - \\ \langle \nabla f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle + \langle \nabla f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle\end{gathered}$$

and divide it into two parts

$$\begin{gathered}\tilde{f}_n = \tilde{f}_n^{(1)} + \tilde{f}_n^{(2)}, \\ \tilde{f}_n^{(1)} = \langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2} + \rho^+_{\xi_{2n}}\beta\mathbf{\Delta}_n) - \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle + \\ \langle \nabla f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2} - \rho^-_{\xi_{2n-1}}\beta\mathbf{\Delta}_n) - \nabla f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle + \\ \langle \nabla f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}) - \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle, \\ \tilde{f}_n^{(2)} = f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}) + \\ 2\langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \beta\mathbf{\Delta}_n\rangle.\end{gathered}$$

Combining all terms, we get $\mathbf{g}_{2n}$:

$$\begin{gathered}\mathbf{g}_{2n} = \frac{1}{2\beta}(\tilde{f}_n^{(1)} + \tilde{f}_n^{(2)} + \tilde{v}_n)\mathbf{\Delta}_n = \\ \langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \mathbf{\Delta}_n\rangle\mathbf{\Delta}_n + \\ \frac{1}{2\beta}(f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}) + \tilde{f}_n^{(1)} + \tilde{v}_n)\mathbf{\Delta}_n.\end{gathered}$$

Next, consider the following product:

$$\begin{gathered}\mathbb{E}_{\mathcal{F}_{t-1}}\langle \mathbf{g}_{2n}, \theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle \leq \qquad (11) \\ \mathbb{E}_{\mathcal{F}_{t-1}}\langle\langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \mathbf{\Delta}_n\rangle\mathbf{\Delta}_n, \theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle + \\ \frac{1}{2\beta}\mathbb{E}_{\mathcal{F}_{t-1}}\langle (f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - f_{\xi_{2n-1}}(\tilde{\mathbf{x}}_{2n-2}) + \\ \tilde{v}_n)\mathbf{\Delta}_n, \theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle + \\ \mathbb{E}_{\mathcal{F}_{t-1}}\|\frac{1}{2\beta}\tilde{f}_n^{(1)}\mathbf{\Delta}_n\|\|\theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|.\end{gathered}$$

Using Assumption 5, for the first and second terms of (11), we have

$$\ldots \leq \langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle.$$

For the third term of (11), using Assumptions 2-5, we get

$$\mathbb{E}_{\mathcal{F}_{t-1}}\|\frac{1}{2\beta}\tilde{f}_n^{(1)}\mathbf{\Delta}_n\| \leq \frac{2L\beta + c}{2}.$$

Then,

$$\begin{gathered}\mathbb{E}_{\mathcal{F}_{t-1}}\|\frac{1}{2\beta}\tilde{f}_n^{(1)}\mathbf{\Delta}_n\|\|\theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\| \leq \\ \frac{2L\beta + c}{2}\|\theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\| \leq \\ \frac{\epsilon^2(2L\beta + c)}{4}\|\theta_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|^2 + \frac{2L\beta + c}{4\epsilon^2},\end{gathered}$$

where $\|\mathbf{a}\| \leq \frac{\epsilon^2}{2}\|\mathbf{a}\|^2 + \frac{1}{2\epsilon^2}$, $\epsilon > 0$.

Let $\eta = \frac{\epsilon^2(2L\beta + c)}{2}$. Assume that we can choose ϵ ensuring $\eta \in (0, \mu)$. Given η preserves the bounds for $\frac{c^2}{2\eta}$, where $2\eta \in (0, 2\mu)$ and

$$\frac{c^2}{2\eta} = \frac{c^2}{\epsilon^2(2L\beta + c)}, \;\; \epsilon^2(2L\beta + c) \in (0, 2\mu).$$

Now, we can use the conditions of Lemma 1 [15] and this completes the proof of Lemma 1. Next, we can obtain a closed form recurrence for values ϕ_n^*.

Lemma 2 (Canonical Form)

Let $\phi_0(\mathbf{x}) = \phi_0^* + \frac{\gamma_0}{2}\|\mathbf{x} - \mathbf{z}_0\|^2$. Then the process (9) preserves the canonical form of functions $\{\phi_n(\mathbf{x})\}$:

$$\phi_n(\mathbf{x}) = \phi_n^* + \frac{\gamma_n}{2}\|\mathbf{x} - \mathbf{z}_n\|^2, \qquad (12)$$

where sequences $\{\gamma_n\}$, $\{\mathbf{z}_n\}$, and $\{\phi_n^*\}$ are defined as follows:

$$\begin{gathered}\mathbf{z}_{n+1} = \gamma_{n+1}^{-1}[(1 - \alpha_n)\gamma_n\mathbf{z}_n + \\ \alpha_n(\mu - \eta)\mathbf{x}_n - \alpha_n\mathbf{g}_n(\mathbf{x}_n)], \\ \gamma_{n+1} = (1 - \alpha_n)\gamma_n + \alpha_n(\mu - \eta), \\ \phi_{n+1}^* = (1 - \alpha_n)(\phi_n^* - Z_n) - \frac{\alpha_n^2}{2\gamma_{n+1}}\|\mathbf{g}_n\|^2 + \\ \alpha_n[r(\mathbf{x}_n) + \frac{(1 - \alpha_n)\gamma_n}{\gamma_{n+1}}\Big(\langle \mathbf{g}_n, \mathbf{z}_n - \mathbf{x}_n\rangle + \\ \frac{\mu - \eta}{2}\|\mathbf{z}_n - \mathbf{x}_n\|^2\Big)]\end{gathered}$$

Proof of Lemma 2: Since we've obtained the same form used in Lemma 1 [15], the proof follows Lemma 2 in [15].

Proof of Theorem 1: Our proof rely on Lemma 3 published in [15].

Lemma 3 [15]. If $\{\lambda_n\}, \{\phi_n(x)\}$ form a bounded estimate sequence for functions $\{f_n(x)\}$ and for some sequence $\{\theta_n\}_{n=0}^\infty$ in $\mathbb{R}^q$, $\{D_n\}_{n=0}^\infty$ in $\mathbb{R}$, $D_n \geq 0$, $D_n < D_\infty < \infty$ the following inequalities hold for all $n \geq 0$:

$$\mathbb{E}f_n(\theta_n) \leq \phi_n^* + D_n = \min_{x \in \mathbb{R}^q}\phi_n(x) + D_n, \;\text{ then } \qquad (13)$$

$$\mathbb{E}f_n(\theta_n) - f_n^* \leq \lambda_n(\phi_0(\theta_0) - f^* + \Phi) + D_n \to_{n\to\infty} D_\infty.$$

Let us choose $\phi_0(\mathbf{x}) = f_0(\widehat{\theta}_0) + \frac{\gamma_0}{2}\|\mathbf{x} - \widehat{\theta}_0\|^2$. Then $f_0(\widehat{\theta}_0) = \phi_0^*$. Using Lemma 1, we have that $\{\phi_n(\cdot)\}_{n=0}^\infty$ and $\{\lambda_n\}_{n=0}^\infty$ generated by the process given in Lemma 2 form the bounded estimate sequence. We need to prove that conditions of Lemma 3 apply.

Let us prove it by induction. By choice of $\phi_0(\cdot)$, condition (13) is valid for $n = 0$. Assume that $\phi_{2n-2}^* \geq \mathbb{E}_{\mathcal{F}_{n-1}}f_{\xi_{2n-2}}(\widehat{\theta}_{2n-2}) - D$:

$$\begin{gathered}\phi_{2n}^* \geq \mathbb{E}[(1 - \alpha_n)(f_{\xi_{2n-2}}(\widehat{\theta}_{2n-2}) - D - Z_n) - \\ \frac{\alpha_n^2}{2\gamma_{n+1}}\|\mathbf{g}_{2n}\|^2 + \alpha_n[r(\tilde{\mathbf{x}}_{2n-2}) + \\ \frac{(1 - \alpha_n)\gamma_n}{\gamma_{n+1}}\Big(\langle \mathbf{g}_{2n}, \mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle + \\ \frac{\mu - \eta}{2}\|\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|^2\Big)]].\end{gathered}$$

Taking the conditional expectation over σ-algebra $\mathcal{F}_{n-1}$, by virtue of Assumptions 2-5, using triangle and Cauchy-Schwarz inequality, we get

$$\mathbb{E}_{\mathcal{F}_{n-1}}\|\mathbf{g}_{2n}\|^2 \leq \tilde{a}^2\|\nabla f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2})\|^2 + 2\tilde{a}\tilde{b}\|\nabla f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2})\| + \tilde{b}^2,$$

where $\tilde{a} = \frac{a}{2\beta} + 1$, $\tilde{b} = \frac{ac+b+c+2L\beta^2+4c\beta}{2\beta}$.

Next, we obtain

$$\mathbb{E}_{\mathcal{F}_{n-1}}\langle \mathbf{g}_{2n}, \mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle \geq \langle f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2}), \mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle - \frac{(2L\beta + c)^2 + 8c^2}{8} - \frac{3}{2}\|\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|^2.$$

Since $f_{\xi_{2n-2}}(\widehat{\theta}_{2n-2}) \geq f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2}) + \langle \nabla f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2}), \widehat{\theta}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle$, we have

$$\phi_{2n}^* \geq f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2}) - (1-\alpha_n)(D + Z_n) - \frac{\alpha_n^2\tilde{a}^2}{2\gamma_{n+1}}\|\nabla f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2})\|^2 - \frac{\alpha_n^2\tilde{a}\tilde{b} + \alpha_n\gamma_{n+1}a}{\gamma_{n+1}}\|\nabla f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2})\| + \frac{4(\eta\alpha_n^2\tilde{b}^2 + \gamma_{n+1}\alpha_n c^2 + 2\gamma_{n+1}\eta\alpha_n b) + \eta((2L\beta + c)^2 + 8c^2)}{8\gamma_{n+1}\eta} + \frac{\alpha_n(1-\alpha_n)\gamma_n(\mu - \eta - 3)}{2\gamma_{n+1}}\|\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|^2 + \langle \nabla f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2}), \frac{\alpha_n\gamma_n}{\gamma_{n+1}}(\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}) + \widehat{\theta}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\rangle.$$

Further, we obtain $\tilde{\mathbf{x}}_{2n-2}$ solving the equation:

$$\frac{\alpha_n\gamma_n}{\gamma_{n+1}}(\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}) + \widehat{\theta}_{2n-2} - \tilde{\mathbf{x}}_{2n-2} = 0.$$

At the same time, Assumption 2 gives us

$$\frac{L}{2}\|\widehat{\theta}_{2n} - \tilde{\mathbf{x}}_{2n-2}\|^2 = \frac{h^2L}{2}\|\mathbf{g}_{2n}\|^2 \geq f_{\xi_{2n}}(\widehat{\theta}_{2n}) - f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - \langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \widehat{\theta}_{2n} - \tilde{\mathbf{x}}_{2n-2}\rangle.$$

Using Assumptions 3-5, we obtain

$$\mathbb{E}_{\mathcal{F}_{n-1}} f_{\xi_{2n}}(\widehat{\theta}_{2n}) \leq \mathbb{E}_{\mathcal{F}_{n-1}}[f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - h\langle \nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}), \mathbf{g}_{2n}\rangle + \frac{h^2L}{2}\|\mathbf{g}_{2n}\|^2] \leq f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - h\|\nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2})\|^2 + \frac{h(2L\beta + c)^2}{8} + \frac{h^2L}{2}\left[\tilde{a}^2 + \frac{\epsilon^2}{2}\right]\|\nabla f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2})\|^2 + \frac{h^2L\epsilon^2(b + c\beta + 2L\beta^2)^2}{4\beta^2} + \frac{2h^2L\beta}{\epsilon^2(b + c\beta + 2L\beta^2)}.$$

Denote by $H = h - \frac{h^2L}{2}\left[\tilde{a}^2 + \frac{\epsilon^2}{2}\right]$. Now, we need to prove that $\phi_{2n}^* \geq \mathbb{E}_{\mathcal{F}_{n-1}} f_{\xi_{2n}}(\widehat{\theta}_{2n}) - D$. Collecting the terms we get:

$$\alpha_n D \geq f_{\xi_{2n}}(\tilde{\mathbf{x}}_{2n-2}) - f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2}) + (1-\alpha_n)Z_n + \left[\frac{\alpha_n^2\tilde{a}^2}{2\gamma_{n+1}} + \frac{\epsilon^2}{2} - H\right]\|\nabla f_{\xi_{2n-2}}(\tilde{\mathbf{x}}_{2n-2})\|^2 + \frac{\alpha_n(1-\alpha_n)\gamma_n(\mu - \eta - 3)}{2\gamma_{n+1}}\|\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|^2 + \tilde{d},$$

where $\tilde{d} = \frac{4(\eta\alpha_n^2\tilde{b}^2 + \gamma_{n+1}\alpha_n c^2 + 2\gamma_{n+1}\eta\alpha_n b) + \eta((2L\beta + c)^2 + 8c^2)}{8\gamma_{n+1}\eta} + \frac{h^2L\epsilon_2^2(b + c\beta + 2L\beta^2)^2}{4\beta^2} + \frac{2h^2L\beta}{\epsilon_2^2(b + c\beta + 2L\beta^2)} + \frac{h(2L\beta + c)^2}{8}$.

We find $\alpha_n \in (0, 1)$ by solving the following equation:

$$\frac{\alpha_n^2\tilde{a}^2}{2\gamma_{n+1}} + \frac{\epsilon^2}{2} - H = 0$$

and we require that $\epsilon \in (0, \sqrt{2H})$. The inequality is proved by the definition of D.

Finally, we obtain

$$D_{n+1} = (1-\alpha_n)D_n + (1-\alpha_n)Z_n + \frac{\alpha_n(1-\alpha_n)\gamma_n(\mu - \eta - 3)}{2\gamma_{n+1}}\|\mathbf{z}_{2n-2} - \tilde{\mathbf{x}}_{2n-2}\|^2 + \tilde{d}.$$

Using Lemma 3 from [15], we conclude the proof.

References

[1] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[2] J. Larson, M. Menickelly, and S. M. Wild, "Derivative-free optimization methods," *Acta Numerica*, vol. 28, pp. 287–404, 2019.

[3] H. Robbins and S. Monro, "A stochastic approximation method," *The annals of mathematical statistics*, pp. 400–407, 1951.

[4] J. Kiefer, J. Wolfowitz, *et al.*, "Stochastic estimation of the maximum of a regression function," *The Annals of Mathematical Statistics*, vol. 23, no. 3, pp. 462–466, 1952.

[5] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Transactions on Automatic Control*, vol. 37, no. 3, pp. 332–341, 1992.

[6] K.-H. Chang, "Stochastic nelder–mead simplex method–a new globally convergent direct search method for simulation optimization," *European journal of operational research*, vol. 220, no. 3, pp. 684–694, 2012.

[7] O. Granichin and N. Amelina, "Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances," *IEEE Transactions on Automatic Control*, vol. 60, no. 6, pp. 1653–1658, 2015.

[8] B. Polyak, *Introduction to Optimization*. New York: Optimization Software, 1987.

[9] A. Y. Popkov, "Gradient methods for nonstationary unconstrained optimization problems," *Automation and Remote Control*, vol. 66, no. 6, pp. 883–891, 2005.

[10] J. Zhu and J. C. Spall, "Probabilistic bounds in tracking a discrete-time varying process," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 4849–4854.

[11] O. Besbes, Y. Gur, and A. Zeevi, "Non-stationary stochastic optimization," *Operations research*, vol. 63, no. 5, pp. 1227–1244, 2015.

[12] A. d'Aspremont, D. Scieur, and A. Taylor, "Acceleration methods," *arXiv preprint arXiv:2101.09545*, 2021.

[13] B. T. Polyak, "Some methods of speeding up the convergence of iteration methods," *Ussr computational mathematics and mathematical physics*, vol. 4, no. 5, pp. 1–17, 1964.

[14] Y. Nesterov *et al.*, *Lectures on convex optimization*. Springer, 2018, vol. 137.

[15] D. Kosaty, A. Vakhitov, O. Granichin, and M. Yuchi, "Stochastic fast gradient for tracking," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 1476–1481.

[16] O. Granichin, V. Erofeeva, Y. Ivanskiy, and Y. Jiang, "Simultaneous perturbation stochastic approximation-based consensus for tracking under unknown-but-bounded disturbances," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3710–3717, 2021.