Randomized Stochastic Approximation Algorithms

Konstantin Amelin, Oleg Granichin, and Olga Granichina¹

Abstract— Multidimensional stochastic optimization plays an important role in analysis and control of many technical systems. To solve the challenging problems of multidimensional optimization, it was suggested to use the randomized algorithms of stochastic approximation with perturbed input which have simple forms and provide consistent estimates of the unknown parameters for observations under "almost arbitrary" noise. They are easily "incorporated" in the design of quantum devices to estimate gradient vector of a multi-variable function.

I. INTRODUCTION

Surprisingly, researchers did not notice for a long time that in case of noisy observations y_n , search algorithms with sequential (n = 1, 2, ...) changes in the estimate $\hat{\theta}_{n-1}$ along some random centered vector Δ_n

$$\theta_n = \theta_{n-1} - \Delta_n y_n$$

might converge to a true vector of controlled parameters θ^{\star} under not only "good", but also under "almost arbitrary" additive disturbances. This happens when observations y_n are taken at some point defined by the previous estimate θ_{n-1} and the randomized vector Δ_n which is called the simultaneous test perturbation (disturbance). Such algorithms are called randomized estimation algorithms. Their convergence under "almost arbitrary" noises is demonstrated through stochastic (probabilistic) properties of the test perturbation. In the near future, experimenters will radically change their current, sometimes, cautious attitude to stochastic algorithms and their results. Modern computing devices will be supplanted by quantum computers, which operate as stochastic systems due to the Heisenberg principle of uncertainty. By virtue of possibility of quantum parallelism, randomizedtype estimation algorithms will, most probably, form the underlying principle of future quantum computing devices.

Initially, stochastic approximation (SA) was introduced in the 1951 article in the *Annals of Mathematical Statistics* by Robbins and Monro [1] and was further developed by Kiefer and Wolfowitz [2] for optimization problems of iterative determining of the stationary point θ^* (point of a local minimum or maximum) of some twice continuously differentiable function $F(\cdot)$ when for each chosen value $x \in \mathbb{R}$ (which is called an input) one can observe with a random noise the corresponding function $F(\cdot)$ value at the point x

$$y(x) = F(x) +$$
noise.

They proved that for any $\widehat{\theta}_0 \in \mathbb{R}$ the recurrent sequence obeying the rule (procedure)

$$\widehat{\theta}_n = \widehat{\theta}_{n-1} - \alpha_n \frac{y(x_n^+) - y(x_n^-)}{2\beta_n}, \ x_n^{\pm} = \widehat{\theta}_{n-1} \pm \beta_n,$$

where $\{\alpha_n\}$ and $\{\beta_n\}$ are some given decreasing numerical sequences: $\sum_{n=1}^{\infty} \alpha_n = \infty$, $\sum_{n=1}^{\infty} \alpha_n^2 \beta_n^{-2} < \infty$, converges to the point θ^* . The main requirement is that the observation noise is conditionally zero-mean. It is usually assumed to be satisfied. It can be formulated as follows: for the *statistics* $G(x,\beta) = \frac{y(x+\beta)-y(x-\beta)}{2\beta}$, which sampled values are precisely observed or calculated, the conditional expectation is close to the gradient (derivative) of the function $F(\cdot)$: $EG(x,\beta) \approx \nabla F(x)$ for a small β (Here and further *E* is a symbol of the mathematical expectation). The Kiefer– Wolfowitz (KW) procedure originally appeared as a tool for statistical computations, it was further developed to the separate field of the control theory. Now this topic has wide variety of applications in such areas as adaptive signal processing, adaptive resource allocation in communication networks, system identification, adaptive control, *etc*.

In general, behavior of the sequence of estimates determined by the algorithm of stochastic approximation depends on the choice of the observed statistic functions $G(x,\beta)$. A convergence rate of estimates of SA algorithms seems to be the main motivation to modify original algorithms. The properties of estimates of the KW procedure and some of its generalizations were studied in details in many books (see, e.g. [3]–[6]). The estimate convergence rate depends on the smoothness of the function $F(\cdot)$. If it is twice differentiable, then the mean-square error of the conventional KW algorithm decreases as $\mathcal{O}(n^{-\frac{1}{2}})$; if it is three times differentiable, as $\mathcal{O}(n^{-\frac{2}{3}})$ [7]. V. Fabian [8] modified the KW procedure by using, besides an approximation of the first derivative, higher-order finite-difference approximations with certain weights. If the function $F(\cdot)$ has ℓ continuous derivatives then the Fabian algorithm provides the meansquare convergence rate of the order $\mathcal{O}(n^{-\frac{\ell-1}{\ell}})$ for odd ℓ . In computational terms, Fabian's algorithm is overcomplicated, the number of observations at one iteration is growing rapidly with smoothness and dimensionality; also, at each step one has to invert a matrix. If a sequence of estimates converges to the desired point of θ^* then B. Polyak [9] and A. Juditsky [10] proposed to improve the process of convergence by using averaging.

¹ This work was supported by RFBR (projects 13-07-00250) and partially supported by the Program for basic research of OEMMPU RAS 01. The authors acknowledge support of Saint-Petersburg State University for a research grants 6.50.2524.2013, 2.50.2219.2013, and 6.37.181.2014. K. Amelin, and O. Granichin and O. Granichina are with the Department of Mathematics and Mechanics, Saint Petersburg State University, 198504, Universitetskii pr. 28, St. Petersburg, Russia. Oleg Granichin is also with the Institute of Problems in Mechanical Engineering, Russian Academy of Sciences, Olga Granichina is also with the Herzen State Pedagogical University of Russia, konstantinamelin@gmail.com, oleg_granichin@mail.ru, olga_granichina@mail.ru

In many applications, the knowledge about statistic characteristics of the measurement noises may be insufficient. For example, noises are the values of an unknown deterministic function. In this case, one encounters appreciable difficulties in motivating applicability of the conventional KW procedure whose estimate often does not converge to the desired point. However, this does not suggest that in dealing with these problems one must abandon the easily representable SA algorithms. The observation can be enriched by adding into the algorithm and observation channel a new random simultaneous perturbation Δ . We modify the KW procedure using instead $G(x,\beta)$ the new randomized statistics $\tilde{G}(x,\beta,\Delta) = G(x,\beta\Delta)$ where Δ is an observed realization of the Bernoulli random variable which is equal to ± 1 with the same probability. For many practical cases, we can assume that Δ does not correlate with an observation noise, and, using the Taylor expanding formula for the function $F(\cdot)$, we obtain for new statistics $\tilde{G}(x,\beta,\Delta)$ that $E\tilde{G}(x,\beta,\Delta) =$ $\nabla F(x) + \mathcal{O}(\beta)$. For a sufficiently small β this means that statistics $\tilde{G}(x,\beta,\Delta)$ is a "good" approximation "in the mean sense" of the gradient (derivative) of the function $F(\cdot)$. A simpler statistics

$$\bar{G}(x,\beta,\Delta) = \frac{\Delta}{\beta}y(x+\beta\Delta)$$

that at each iteration (step) uses only one observation has the same property. This statistics was used in [11] for constructing a sequence of consistent estimates for a dependent observation noise (almost arbitrary noise). The essence of the perturbation Δ is an exciting action because it is used mostly to make the non-degenerate observations.

In [12] the SA algorithm was extended to the multidimensional case. When $\theta \in \mathbb{R}^d$, the conventional KW-procedure based on finite-difference approximations of the function gradient vector uses 2*d* observations at each iteration to construct the sequence of estimates (two observations for approximations of each component of the *d*-dimensional gradient vector). Let Δ be an *d*-dimensional Bernoulli random vector. The randomized statistic $\bar{G}(\mathbf{x}, \beta, \Delta)$ admits to use a computationally simpler procedure with only one measurement of the function $F(\cdot)$ [13]–[15]. The generalization of the $\tilde{G}(\mathbf{x}, \beta, \Delta)$ to the multidimensional case is

$$\tilde{G}(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Delta}) = \begin{pmatrix} \vdots \\ \frac{1}{\Delta_i} \\ \vdots \end{pmatrix} \frac{y(\mathbf{x}+\boldsymbol{\beta}\boldsymbol{\Delta}) - y(\mathbf{x}-\boldsymbol{\beta}\boldsymbol{\Delta})}{2\boldsymbol{\beta}}.$$

It was used by J. Spall [16] in the suggested *a simultaneous perturbation stochastic approximation (SPSA)* algorithm. He showed that for a large n the probabilistic distribution of appropriately scaled estimation errors is approximately normal. He used the formula obtained for the asymptotic error variance and a similar characteristic of the KW procedure to compare overall performances of algorithms. It was noticed that the SPSA algorithm has the same order of a convergence rate as the KW procedure if other things being equal. Moreover, appreciably fewer observations are used in the

multidimensional case (by the factor of *d* as $n \to \infty$). In [17], [18], algorithms, similar to SPSA, were suggested to be used for neuron networks training because they could be implemented by a small set of simple components.

It is important to note, that the asymptotic mean-square convergence rate can be increased without increasing the number of measurements of the function at each iteration in case when we use randomized SA algorithms under conditions of sufficiently smooth functions $F(\cdot)$. For the case where some generalized index of a function $F(\cdot)$ smoothness is equal to γ ($\gamma = \ell + 1$ when all partial derivatives of orders up to ℓ inclusive satisfy the Lipschitz condition), B. Polyak and A. Tsybakov [14] proposed to use statistics

$$\tilde{G}_{\gamma}(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Delta}) = \mathbf{K}_{\gamma}(\boldsymbol{\Delta}) \frac{y(\mathbf{x}+\boldsymbol{\beta}\boldsymbol{\Delta}) - y(\mathbf{x}-\boldsymbol{\beta}\boldsymbol{\Delta})}{2\boldsymbol{\beta}}$$

and

$$\bar{G}_{\gamma}(\mathbf{x},\boldsymbol{\beta},\Delta) = \frac{1}{\boldsymbol{\beta}} \mathbf{K}_{\gamma}(\Delta) y(\mathbf{x}+\boldsymbol{\beta}\Delta), \qquad (2)$$

(1)

where $\mathbf{K}_{\gamma}(\cdot)$ is a vector-function with a finite support (a differentiable kernel) determined by orthogonal Legendre polynomials of a degree smaller than γ . The corresponding randomized algorithms provide the mean-square convergence rate $\mathcal{O}(n^{-\frac{\gamma-1}{\gamma}})$ of a sequence of estimates. They demonstrated also that for a wide class of iterative algorithms this convergence rate is optimal in some asymptotically minimax sense, that is, it cannot be improved neither for any other algorithm nor for any other admissible rule of choosing the measurement points (This fact for odd ℓ was earlier established by H.-F. Chen [19]).

The algorithm with more general view of a gradient vector stochastic approximation

$$\widehat{G}_{\gamma}(\mathbf{x},\boldsymbol{\beta}^{+},\boldsymbol{\beta}^{-},\boldsymbol{\Delta}) = \mathbf{K}_{\gamma}(\boldsymbol{\Delta}) \frac{y(\mathbf{x}+\boldsymbol{\beta}^{+}\boldsymbol{\Delta}) - y(\mathbf{x}-\boldsymbol{\beta}^{-}\boldsymbol{\Delta})}{\boldsymbol{\beta}^{+}+\boldsymbol{\beta}^{-}} \quad (3)$$

is considered in [20] and in [21] with $\mathbf{K}_{\gamma}(\Delta) \equiv \Delta$. It was motivated by practical applications. The partial case with $\beta^{-} \equiv 0$ and $\mathbf{K}_{\gamma}(\Delta) \equiv \Delta$ was proposed by H.-F. Chen, T. Duncan and B. Pasik-Duncan [22].

The algorithms which are based on statistics (1)–(3) have significant advantages:

- the asymptotically optimal rate of convergence;
- the minimum amount of measurements within the current iteration;
- the consistency with almost arbitrary interference;
- the operability in nonstationary problems;
- a "natural" implementation on a quantum computer.

These properties are established in practice and theoretically substantiated. Consistency of randomized algorithms of stochastic approximation in the multidimensional case under almost arbitrary noise was established in [13], [22]– [24]. In [25], [26] the convergence study of the randomized SA algorithm with one observation under almost arbitrary noise was extended to the case of additional multiplicative noise and weakers dependence between perturbation Δ and observation noises. In [27] the detailed analysis of asymptotic properties of the resulting estimates sequence was made. Authors show that the optimal distribution for the components of Δ is the symmetric Bernoulli distribution. The effectiveness of this simple distribution is confirmed in many practice examples for the finite sample of observations too. Sometimes it is desirable to choose a different distribution in practical problems. For example, in [28] the robot controller is using a symmetric uniform distribution which consists of two parts when neighborhood of zero is removed.

Stochastic approximation algorithms were initially proven in case of the stationary functional. Randomized SA algorithms are considered for the case of time-varying quality functionals in [29]-[32]. A possible implementation of randomized stochastic approximation algorithms on a quantum computer is studied in [33]. Randomized algorithm for stochastic approximation in the machine self-learning problem is proposed and justified in [34]. An optimization of a server processing queue is considered in [35]. In [21], [36] SPSA method has been proposed for estimation of the centers of thermal updrafts for the adaptive autonomous soaring of multiple UAVs. For a multi-plant differential game a leader-follower strategies is proposed in [37] based on a randomized stochastic approximation algorithm. SPSA like method was used for an active optimization in studying the quality of educational process in high school [38]. In all these examples, traditional assumptions of an independence and zero-mean of an external noise are not held due to specifics of these problems. The randomization promotes not only a speed up of data processing but also a reduction of negative influence of an "almost" arbitrary external noise of observations.

The severe portion of applications of stochastic approximation algorithms is concerned with adaptive systems. It is based on the fact that SA algorithms have properties useful for uncertain environments. These important properties allow these algorithms to track typical behavior of such system. Furthermore, these algorithms are memory and computationally efficient which makes them applicable to a real time dynamic environment. These properties made the algorithms applicable in a such new field as soft computing where they are used for the "parameter tuning". The algorithms for neural networks training and reinforcement learning are notable among these. They are used in a popular learning paradigm for autonomous software agents with applications in e-commerce, robotics, etc. They are also widely applied in economic theory, providing a good model for collective phenomena, when they are modeled behavior of individual bounded rational agents.

This paper presents the main ideas of a randomized stochastic approximation. It is organized as follows: in Section II, we state a formal problem setting of a meanrisk optimization. Section III introduces the exciting testing perturbation as randomization and estimation algorithms. In Section IV, the convergence of estimates is studied. Section V deals with the non-constrained optimization in the context of tracking. Section VI shows a possible way of a quantum computing for the algorithm implementation.

II. MEAN-RISK OPTIMIZATION

Many practical applications need to optimize one or another mean risk functional. Although sometimes extremal values can be established analytically, the engineering systems often deal with an unknown functional whose value or gradient can be calculated at the given points.

Let $f(\theta, \mathbf{w}) : \mathbb{R}^d \times \mathbb{W} \to \mathbb{R}$, $\mathbb{W} \subset \mathbb{R}^r$, be a θ -differentiable function and let $\mathbf{x}_1, \mathbf{x}_2, \ldots$ be a sequence of measurement points chosen by the experimenter (observation plan), at which the value y_1, y_2, \ldots of the function $f(\cdot, \cdot)$ is accessible to observation at every instant $t = 1, 2, \ldots$, with additive external noise v_t

$$y_t = f(\mathbf{x}_t, \mathbf{w}_t) + v_t, \tag{4}$$

where $\mathbf{w}_t \in \mathbb{W}$, t = 1, 2, ..., are uncontrollable random variables (vectors).

Formulation of the problem. Using the observations $y_1, y_2, ...,$ construct a sequence of estimates $\{\widehat{\theta}(n)\}$ of an unknown vector θ^* minimizing the mean-risk functional

$$F(\boldsymbol{\theta}) = E_{\mathbf{w}} f(\boldsymbol{\theta}, \mathbf{w}) \to \min_{\boldsymbol{\theta}} .$$
 (5)

Here and further $E_{\mathbf{w}}$ is a symbol of the conditional mathematical expectation of respect to \mathbf{w} .

Usually, minimization of the function $F(\theta)$ is studied with a simpler observation model

$$y_t = F(\mathbf{x}_t) + v_t.$$

The generalization in the formulation (4) is motivated by three reasons. At first, it takes into account the case of multiplicative random perturbations in observations:

$$y_t = \mathbf{w}_t \bar{f}(\mathbf{x}_t) + v_t.$$

At second, it allows to separate the observation noise with "good" (e. g., zero-mean i.i.d.) statistical properties $\{w_t\}$ and arbitrary additive external noise $\{v_t\}$. (Of cause, this separation is not need when we can assume that $\{v_t\}$ is random zero-mean and i.i.d. too.) At third, we also encounter problems where the optimized functional may vary in time and its point of extremum may drift. In such cases, the problems may be posed differently depending on the optimization goals and measurable data. Usually two options of a behavior of the drift $\{\theta_t\}$ of functional's point of minimum is considered. They differ in the answer to the question: is there some random i.i.d. sequence with an expectation θ^* ? The case of the positive answer is included into (5) setting since we can consider additional disturbances $\mathbf{w}_t = \theta_t - \theta^*$. More general non-stationary problem statement is

$$F_t(\boldsymbol{\theta}) = E_{\boldsymbol{\xi}_t, \mathbf{w}} f_{\boldsymbol{\xi}_t}(\boldsymbol{\theta}, \mathbf{w}) \to \min_{\boldsymbol{\theta}}, \tag{6}$$

where $\{f_{\xi}(\theta, \mathbf{w})\}_{\xi \in \Xi}$, is a family of θ -differentiable functions: $f_{\xi}(\theta, \mathbf{w}) : \mathbb{R}^d \times \mathbb{W} \to \mathbb{R}, \ \mathbb{W} \subset \mathbb{R}^r$, and for chosen sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ we can observe with additive external noise v_t a

$$y_t = f_{\xi_t}(\mathbf{x}_t, \mathbf{w}_t) + v_t, \tag{7}$$

where $\{\xi_t\}$ is a uncontrollable sequence: $\xi_t \in \Xi$, $\mathbf{w}_t \in W$, t = 1, 2, ..., are uncontrollable random variables (vectors).

More precisely, *it is needed* to estimate the time-varying point of minimum θ_t of the function $F_t(\theta)$.

The problem (5) is a partial case of (6) when $\theta_t \equiv \theta^*$ and $\Xi = \{\xi_1\}$.

Let us state the main conditions which are usually assumed for proving of theoretical results.

SA1 The function $F(\cdot)$ is strongly convex at the first argument, i.e. it has a unique minimum point θ^*

$$\langle \mathbf{x} - \boldsymbol{\theta}^{\star}, \nabla F(\mathbf{x}) \rangle \geq \mu \| \mathbf{x} - \boldsymbol{\theta}^{\star} \|^{2}, \ \forall \mathbf{x} \in \mathbb{R}^{d}$$

with a constant $\mu > 0$.

SA2 The gradient $\nabla_{\mathbf{x}} f(\theta, \mathbf{w})$ satisfies $\forall \mathbf{w} \in \mathbb{W}$ the Lipschitz condition: $\forall \mathbf{x}', \mathbf{x}'' \in \mathbb{R}^d$

$$\|\nabla_{\mathbf{x}}f(\mathbf{x}',\mathbf{w}) - \nabla_{\mathbf{x}}f(\mathbf{x}'',\mathbf{w})\| \le M\|\mathbf{x}' - \mathbf{x}''\|_{\mathbf{x}}$$

with a constant $M > \mu$.

SA3 Operation commutativity of the differentiation on **x** and integration on **w** for the function $f(\mathbf{x}, \mathbf{w})$.

III. EXCITING TESTING PERTURBATION AS RANDOMIZATION AND ESTIMATION ALGORITHMS

Let Δ_n , n = 1, 2, ... be an observed sequence of independent random variables in \mathbb{R}^d , called the simultaneous test perturbation, with distribution function $P_n(\cdot)$ and let $\mathbf{K}_n(\cdot) : \mathbb{R}^d \to \mathbb{R}^d$, n = 1, 2, ..., be some vector functions (kernels) with compact supports which, along with distribution functions of the test perturbation, satisfy the conditions

$$\int \mathbf{K}_n(\mathbf{x}) \mathbf{P}_n(d\mathbf{x}) = 0, \quad \int \mathbf{K}_n(\mathbf{x}) \mathbf{x}^{\mathrm{T}} \mathbf{P}_n(d\mathbf{x}) = \mathbf{I},$$

$$\sup_n \int \|\mathbf{K}_n(\mathbf{x})\|^2 \mathbf{P}_n(d\mathbf{x}) < \infty, \quad n = 1, 2, \dots.$$
(8)

For example, we can choose a realization of a sequence of independent Bernoulli random vectors from \mathbb{R}^d with each component independently assuming values $\pm \frac{1}{\sqrt{d}}$ with the probabilities $\frac{1}{2}$ as a sequence $\{\Delta_n\}$ and $\mathbf{K}_n(\mathbf{x}) \equiv d\mathbf{x}$ as kernel functions.

Let us take a fixed initial vector $\hat{\theta}_0 \in \mathbb{R}^d$ and choose sequences of positive numbers $\{\alpha_n\}, \{\beta_n^+\}$ and $\{\beta_n^-\}$. We design two algorithms for constructing sequences of points of observations $\{\mathbf{x}_n\}$ and estimates $\{\hat{\theta}_n\}$. The first algorithm uses at every step (iteration) one observation

$$\begin{cases} \mathbf{x}_n = \widehat{\theta}_{n-1} + \beta_n^+ \Delta_n, y_n = f(\mathbf{x}_n, \mathbf{w}_n) + v_n, \\ \widehat{\theta}_n = \widehat{\theta}_{n-1} - \frac{\alpha_n}{\beta_n^+} \mathbf{K}_n(\Delta_n) y_n, \end{cases}$$
(9)

and the second one uses two observations

$$\begin{cases} \mathbf{x}_{2n-\frac{1}{2}\pm\frac{1}{2}} = \widehat{\theta}_{n-1} \pm \beta_n^{\pm} \Delta_n, \\ \widehat{\theta}_n = \widehat{\theta}_{n-1} - \frac{\alpha_n}{\beta_n^{+} + \beta_n^{-}} \mathbf{K}_n(\Delta_n)(y_{2n} - y_{2n-1}). \end{cases}$$
(10)

Algorithms (9) and (10) correspond to the statistics (2) and (3) described above.

IV. CONVERGENCE OF ESTIMATES

Convergence of estimates generated by the algorithms (9) and (10) is studied in details in [26] for the cases when $\beta_n^- = \beta_n^+$ or $\beta_n^- = 0$. Here we present a similar result for the general case of the algorithm (10).

Let \mathscr{F}_{n-1} be the σ -algebra of probabilistic events generated by the random variables $\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_{n-1}$ formed by the algorithm (10), and denote $\bar{v}_n = v_{2n} - v_{2n-1}$, $\bar{\mathbf{w}}_n = \begin{pmatrix} \mathbf{w}_{2n-1} \\ \mathbf{w}_{2n-1} \end{pmatrix}$, $\beta_n = (\beta^+ + \beta^-)/2$.

$$\mathbf{w}_{2n}$$
), $p_n = (p_n' + p_n)/2$.

Assume that the following condition holds

SA4 Random vectors $\bar{\mathbf{w}}_n$ and Δ_n are independent, random vectors $\bar{\mathbf{w}}_2, \bar{\mathbf{w}}_1, \ldots, \bar{\mathbf{w}}_{n-1}$ do not depend on $\bar{\mathbf{w}}_n$ and Δ_n . If values $\bar{v}_1, \ldots, \bar{v}_n, \xi_1, \ldots, \xi_{2n}$ are random then they do not depend on $\bar{\mathbf{w}}_n$ and Δ_n too.

Theorem 1. Let conditions SA1-4 and (8) be satisfied.

If $\sum_{n} \alpha_{n} = \infty$ and $\alpha_{n} \to 0$, $\beta_{n} \to 0$, $\alpha_{n}^{2} \beta_{n}^{-2} (1 + E \bar{v}_{n}^{2}) \to 0$ as $n \to \infty$,

then $E \|\widehat{\theta}_n - \theta^*\|^2 \to 0$ as $n \to \infty$. Moreover, if $\sum_n \alpha_n \beta_n^2 < \infty$ and

$$\sum_{n} \alpha_n^2 \beta_n^{-2} (1 + E_{\mathscr{F}_{n-1}} \bar{v}_n^2) < \infty, \text{ with probability } 1,$$

then $\widehat{\theta}_n \to \theta^*$ as $n \to \infty$ with probability 1.

Proof: [Sketch of the proof of Theorem 1] The proof of Theorem 1 is similar to the corresponding proof in [26]. Let be $v_n = \hat{\theta}_n - \theta^*$, By virtue the view of the algorithm (10) and observation model (4) we can get the bounds

$$E_{\mathscr{F}_{n-1}} \| \mathbf{v}_n \|^2 \le (1 - c_1 \alpha_n) \| \mathbf{v}_n \|^2 + c_2 \alpha_n^2 \beta_n^{-2} (1 + E_{\mathscr{F}_{n-1}} \bar{\mathbf{v}}_n^2)$$

with constants c_1 and c_2 which are determined by Theorem 1 conditions, and all conditions of the Robbins-Siegmund Lemma [39], that are necessary for the convergence of $\{\widehat{\theta}_n\}$ to the point θ^* as $n \to \infty$ with probability 1, are satisfied. To prove the result of Theorem 1 for the mean-square convergence, let us examine the unconditional mathematical expectation of both sides of the last inequality

$$E \|\mathbf{v}_n\|^2 \le E \|\mathbf{v}_n\|^2 (1 - c_1 \alpha_n) + c_2 \alpha_n^2 \beta_n^{-2} (1 + E \bar{v}_n^2).$$

The mean-square convergence of the sequence $\{\hat{\theta}(n)\}$ to the point θ^* is implied by Lemma 5 of [40]. This completes the sketch of the proof of Theorem 1.

Remarks. 1. Instead of mean-square boundedness of noise and disturbances we can use more weak assumptions about their statistical moments of an order ρ : $1 < \rho < 2$ (see [33]). 2. The conditions of Theorem 1 hold for the function

 $w\bar{f}(\mathbf{x})$ if the function $\bar{f}(\mathbf{x})$ satisfies conditions SA1-2.

3. The observation noise v_n in Theorem 1 can be said to be "almost arbitrary" since they may be nonrandom but independent and bounded or they may also be a realization of some stochastic process with arbitrary dependencies. In particular, to prove the results of Theorem 1, there is no need to assume that \bar{v}_n and \mathcal{F}_{n-1} are independent.

4. Though algorithms (9) and (10) may look alike, algorithm (9) is more suitable for use in realtime systems if observations contain arbitrary noise. For algorithm (10), the condition that observation noise v_{2n} and the test perturbation Δ_n are independent is rather restrictive, because the vector Δ_n is used at the previous instant 2n-1 in the system. For the algorithm (10) the noise v_{2n} and the test perturbation vector Δ_n simultaneously appear in the system and they can be regarded as independent only in the case when $\beta_n^- = 0$.

V. TRACKING

To analyze the quality of estimates we will apply the following definition for the problem of minimum tracking for mean-risk functional (6).

Definition. A sequence of estimates $\widehat{\theta}_n$ has an asymptotically efficient upper bound of residues of estimation L > 0 if $\forall \varepsilon > 0 \exists N$ such that $\sqrt{E \|\widehat{\theta}_n - \theta_n\|^2} \leq L + \varepsilon \ \forall n > N$.

In the nonstationary problem (6) stochastic approximation algorithms commonly used with constant step-sizes [6], [29], [32].

Further we will assume that the additional following conditions are true.

SA5 The gradient of ∇f_{ξ_t} is uniformly bounded at the minimum points: $E \|\nabla f_{\xi_t}(\theta_t, \mathbf{w}_t)\|^2 \leq \bar{g} < \infty$, $(\bar{g} = 0$ for the case $f_{\xi_t}(\mathbf{x}, \mathbf{w}) = F_t(\mathbf{x})$).

SA6 Drift is bounded $\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1}\| \leq \delta_{\boldsymbol{\theta}} < \infty$ and for any arbitrary point $\mathbf{x}: E_{\mathscr{F}_{t-1}} \varphi_t(\mathbf{x})^2 \leq \bar{a} \|\mathbf{x} - \boldsymbol{\theta}_{t-1}\|^2 + \bar{b}$, where $\varphi_t(\mathbf{x}) = f_{\xi_t}(\mathbf{x}, \mathbf{w}_t) - f_{\xi_{t-1}}(\mathbf{x}, \mathbf{w}_{t-1})$.

SA7 The observation noise $\{v_n\}$ satisfies: $|v_{2n} - v_{2n-1}| \le c_v < \infty$.

The following theorem shows the asymptotically efficient upper bound of residues of estimation by the algorithm (10).

Theorem 2. Let's consider the randomized SA algorithm with two measurements (10) where $\alpha_n \equiv \alpha$, $\beta_n^+ \equiv \beta$, $\beta_n^- \equiv 0$, $\mathbf{K}_n(\Delta_n) = \Delta_n$, and Δ_n be a sequence of independent identically distributed Bernoulli random vectors which components are independent and equal $\pm \frac{1}{\sqrt{d}}$ with same probability $\frac{1}{2}$.

Let the conditions SA1-7 hold.

l

If $k = 2\mu - 2\alpha(M^2 + 2\frac{\bar{\alpha}}{\beta}) < 1/\alpha$ and $\alpha < \mu/M^2$, then the sequence of estimates provided by the algorithm (10) has an asymptotically efficient upper bound which equals to

$$\bar{L} = \left(\frac{4\delta_{\theta}}{k\alpha} + \delta_{\theta}\right)^2 - \delta_{\theta}^2 + \frac{l}{k},$$
(11)
$$= 4\alpha \left(\frac{c_v^2 + 2\bar{b}}{\beta^2} + 3M^2 + 2\bar{g}\right) + 4\beta M.$$

Theorem 2 is a partial case of a more general result in [20].

Remark. The result of the Theorem 2 shows that for the case without drift ($\delta_{\theta} = 0$) the asymptotic upper bound is $\overline{L} = l/k$. It can be made infinitely small under any noise level c_{ν} , it is just needed to chose α and β sufficiently small. At the same time for the case with drift, the bigger drift norm δ_{θ} can be compensated by the choosing of the bigger stepsize α . This leads to the tradeoff between making α smaller because we have noisy observations and making it bigger because we have the drift of optimal points.

VI. ALGORITHM IMPLEMENTATION AND QUANTUM COMPUTING

Till recently quantum computer was regarded exclusively as a notional mathematical model. Of course, serious difficulties are still encountered in designing a quantum computer for everyday use. Nonetheless, intensive researches and development projects permanently starting and continuing in this field.

The representation of the algorithm (9) is associated with something well-known for those people who is familiar with fundamentals of "quantum computing". Virtually all known effective quantum algorithms implement a similar scheme:

- a preparation of an inputs "superposition",
- processing,
- measuring of a result.

Let us examine the choice of the best computer for implementation of the randomized stochastic optimization algorithm with one measurement of the penalty function per iteration. Realization of algorithm (9) on a quantum computer is described in [33]. Recently, terminology and axiomatic of quantum computation models have been greatly refined. Below we describe a method of the algorithm (9) representation for an implementation on a "hypothetical" quantum computer, i.e., a method that is consistent with the general logic of quantum computation algorithms.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function satisfying the conditions **SA1–3**. Let us assume that the quantum computer is an *r*-bit machine. The unitary operation realizing the function $f(\theta)$ on a quantum computer can be defined on all classical binary chains *x* of length *dr*, defining the argument of the function

$$U_f: |x\rangle |z\rangle \to |x\rangle |z \oplus f(x)\rangle,$$

where z is an arbitrary binary chain of length r and \oplus is a bit-by-bit operation of logical AND. This is a method of defining an operator on the basis vectors. On all other vectors, the operator is continued linearly. Clearly, the operator constructed in this way is invertible and acts in a complex space of dimension 2^{dr+r} .

We estimate the minimum of a function, using algorithm (9) which iteration has a form

$$x = \widehat{\theta} + \beta \Delta, \ y = f(x) + v, \ \widehat{\theta} := \widehat{\theta} - \gamma(x - \widehat{\theta})y,$$

where $\gamma = \alpha / \beta^2$.

To feed the computer input, let us prepare a superposition of 2d perturbed values of the current estimate vector

$$x = \frac{1}{2^{\frac{d}{2}}} \sum_{\Delta_i \in \{-1,+1\}^d} |\widehat{\theta} + \beta \Delta_i\rangle = H_\beta |\widehat{\theta}\rangle,$$

where ± 1 are regarded as *r*-digit numbers, H_{β} is the corresponding unitary operator which is familiar to the traditional in quantum computing Hadamard transformation.

Applying the unitary operator U_f to $|x\rangle|0\rangle$, we obtain

$$U_f|x
angle|0
angle = rac{1}{2^d}\sum_{\Delta_i\in\{-1,+1\}^d}|\widehat{ heta}+eta\Delta_i
angle|f(\widehat{ heta}+eta\Delta_i)
angle.$$

By general properties of the quantum computation model, after a state measurement, obtain with probability $\frac{1}{2d}$ a vector

$$|\widehat{\theta} + \beta \Delta_i\rangle|f(\widehat{\theta} + \beta \Delta_i)\rangle, \ \Delta_i \in \{-1, +1\}^d$$

Using first dr digits of this vector, we can easily determine a random perturbation vector Δ_n . According to algorithm (9), its coordinates must be multiplied by the corresponding value of the loss function at a perturbed point, i.e., by the value at the last r digits of the measurement result.



Fig. 1. The quantum circuit for "on the fly" computing of the gradient.

Fig. 1 shows the quantum circuit for "on the fly" computing of the function f gradient.

REFERENCES

- H. Robbins and S. Monro, "A stochastic approximation method," *The* Annals of Mathematical Statistics, pp. 400–407, 1951.
- [2] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *The Annals of Mathematical Statistics*, vol. 23, no. 3, pp. 462–466, 1952.
- [3] J. C. Spall, Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control.
- [4] H. J. Kushner and G. Yin, Stochastic Approximation and Recursive Algorithms and Applications. New York, Springer–Verlag, 2003.
- [5] O. Granichin and B. Polyak, Randomized Algorithms of an Estimation and Optimization Under Almost Arbitrary Noises. Moscow: Nauka, 2003.
- [6] V. S. Borkar, Stochastic Approximation: a Dynamical Systems Viewpoint. Cambridge University Press Cambridge, 2008.
- [7] M. Wasan, Stochastic Approximation. Cambridge University Press, Cambridge, England, 1969.
- [8] V. Fabian, "Stochastic approximation of minima with improved asymptotic speed," *The Annals of Mathematical Statistics*, vol. 38, no. 1, pp. 191–200, 1967.
- [9] B. T. Polyak, "New stochastic approximation type procedures," Automation and Remote Control, vol. 7, pp. 98–107, 1990.
- [10] A. Juditsky, "A stochastic estimation algorithm with observation averaging," *IEEE Transactions on Automatic Control*, vol. 38, no. 5, pp. 794–798, 1993.
- [11] O. Granichin, "A stochastic recursive procedure with correlated noise in the observation, that employs trial perturbations at the input," *Vestnik Leningrad University: Math*, vol. 22, no. 1, pp. 27–31, 1989.
- [12] J. R. Blum, "Multidimensional stochastic approximation methods," *The Annals of Mathematical Statistics*, vol. 25, no. 4, pp. 737–744, 1954.
- [13] O. Granichin, "Procedure of stochastic approximation with disturbances at the input," *Automation and Remote Control*, vol. 53, no. 2, part 1, pp. 232–237, 1992.
- [14] B. T. Polyak and A. B. Tsybakov, "Optimal order of accuracy of search algorithms in stochastic optimization," *Problemy Peredachi Informatsii*, vol. 26, no. 2, pp. 45–53, 1990.
- [15] J. C. Spall, "A one-measurement form of simultaneous perturbation stochastic approximation," *Automatica*, vol. 33, no. 1, pp. 109–112, 1997.
- [16] —, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Transactions on Automatic Control*, vol. 37, no. 3, pp. 332–341, 1992.
- [17] J. Alespector, "A parallel gradient descent method for learning in analog vlsi neural networks," *Advances in Neural Information Processing Systems*, vol. 5, pp. 836–844, 1993.
- [18] Y. Maeda and Y. Kanata, "Learning rules for recurrent neural networks using perturbation and their application to neuro-control," *Transactions Institute of Electrical Engineers of Japan*, vol. 113, pp. 402–402, 1993.

- [19] H. Chen, "Lower rate of convergence for locating a maximum of a function," *The Annals of Statistics*, vol. 16, no. 3, pp. 1330–1334, 1988.
- [20] O. Granichin, "Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances (submitted)," *IEEE Transactions on Automatic Control*, 2014.
- [21] K. Amelin, N. Amelina, O. Granichin, O. Granichina, and B. Andrievsky, "Randomized algorithm for uavs group flight optimization," in *Proc. of the 11th IFAC Int. Workshop on Adaptation and Learning* in Control and Signal Processing (ALCOS2013), 2013, pp. 205–208.
- [22] H. Chen, T. E. Duncan, and B. Pasik-Duncan, "A kiefer-wolfowitz algorithm with randomized differences," *IEEE Transactions on Automatic Control*, vol. 44, no. 3, pp. 442–453, 1999.
- [23] O. Granichin, "Unknown function minimum point estimation under dependent noise," *Problems Inform. Transmission*, vol. 28, no. 2, pp. 16–20, 1992.
- [24] B. Polyak and A. Tsybakov, "On stochastic approximation with arbitrary noise (the kw case)," *Topics in Nonparametric Estimation. Adv. in Soviet Mathematics. Am. Math. Soc.*, no. 12, pp. 107–113, 1992.
- [25] O. Granichin, "Stochastic Approximation under Dependent Noises, Detecting Signals and Adaptive Control," in *Approximation, Probability, and Related Fields*, Anastassiou, G and Rachev, ST, Ed. Plenum, 1994, pp. 247–271.
- [26] —, "Randomized algorithms for stochastic approximation under arbitrary disturbances," *Automation and Remote Control*, vol. 63, no. 2, pp. 209–219, 2002.
- [27] P. Sadegh and J. C. Spall, "Optimal random perturbations for stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Transactions on Automatic Control*, vol. 43, no. 10, pp. 1480–1484, 1998.
- [28] Y. Maeda and R. J. De Figueiredo, "Learning rules for neuro-controller via simultaneous perturbation," *IEEE Transactions on Neural Net*works, vol. 8, no. 5, pp. 1119–1130, 1997.
- [29] O. Granichin, L. Gurevich, and A. Vakhitov, "Discrete-time minimum tracking based on stochastic approximation algorithm with randomized differences," in *Proc. of the 48th IEEE Conference on Decision and Control (CDC2009)*, 2009, pp. 5763–5767.
- [30] J. C. Spall and J. A. Cristion, "A neural network controller for systems with unmodeled dynamics with applications to wastewater treatment," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 27, no. 3, pp. 369–375, 1997.
- [31] —, "Model-free control of nonlinear stochastic systems with discrete-time measurements," *IEEE Transactions on Automatic Control*, vol. 43, no. 9, pp. 1198–1210, 1998.
- [32] A. Vakhitov, O. Granichin, and L. Gurevich, "Algorithm for stochastic approximation with trial input perturbation in the nonstationary problem of optimization," *Automation and Remote Control*, vol. 70, no. 11, pp. 1827–1835, 2009.
- [33] A. Vakhitov, O. Granichin, and S. Sysoev, "A randomized stochastic optimization algorithm: Its estimation accuracy," *Automation and Remote Control*, vol. 67, no. 4, pp. 589–597, 2006.
- [34] O. Granichin and O. Izmakova, "A randomized stochastic approximation algorithm for self-learning," *Automation and Remote Control*, vol. 66, no. 8, pp. 1239–1248, 2005.
- [35] Y. Volkovich and O. Granichin, "Adaptive optimization of a server which processes queue of tasks," *Stohasticheskaya Optimizaciya v Informatike*, vol. 1, no. 1, pp. 17–28, 2005.
- [36] C. Antal, O. Granichin, and S. Levi, "Adaptive Autonomous Soaring of Multiple UAVs Using Simultaneous Perturbation Stochastic Approximation," in *Proc. of the 49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 3656–3661.
- [37] M. Jimenez-Lizarraga, A. Poznyak, and M. Alcorta, "Leader-follower strategies for a multi-plant differential game," in *Proc. of the American Control Conference*, 2008, pp. 3839–3844.
- [38] O. Granichina, "Mathematical models of the quality of educational process in high school with an active optimization," *Stohasticheskaya Optimizaciya v Informatike*, vol. 2, no. 1, pp. 77–108, 2006.
- [39] H. Robbins and D. Siegmund, "A convergence theorem for nonnegative almost supermartingales and some applications," *Optimizing methods* in statistics, pp. 233–257, 1971.
- [40] B. Polyak, *Introduction to optimization*. Optimization Software, Publications Division (New York), 1987.