2019 57th Annual Allerton Conference on Communication, Control, and Computing
(Allerton)
Allerton Park and Retreat Center
Monticello, IL, USA, September 24-27, 2019

Stochastic Gradient Descent on a Tree: an Adaptive and Robust Approach to Stochastic Convex Optimization

Sattar Vakili*, Sudeep Salgia†, and Qing Zhao†

*Prowler.io, Cambridge, UK, *sattar@prowler.io*
†School of Electrical & Computer Engineering, Cornell University, Ithaca, NY, *{ss3827,qz16}@cornell.edu*

***Abstract*— Online minimization of an unknown convex function over the interval $[0,1]$ is considered under first-order stochastic bandit feedback, which returns a random realization of the gradient of the function at each query point. Without knowing the distribution of the random gradients, a learning algorithm sequentially chooses query points with the objective of minimizing regret defined as the expected cumulative loss of the function values at the query points in excess to the minimum value of the function. An approach based on devising a biased random walk on an infinite-depth binary tree constructed through successive partitioning of the domain of the function is developed. Each move of the random walk is guided by a sequential test based on confidence bounds on the empirical mean constructed using the law of the iterated logarithm. With no tuning parameters, this learning algorithm is robust to heavy-tailed noise with infinite variance and adaptive to unknown function characteristics (specifically, convex, strongly convex, and nonsmooth). It achieves the corresponding optimal regret orders (up to a $\sqrt{\log T}$ or a $\log\log T$ factor) in each class of functions and offers better or matching regret orders than the classical stochastic gradient descent approach which requires the knowledge of the function characteristics for tuning the sequence of step-sizes.**

I. Introduction

A. Stochastic Convex Optimization

In stochastic convex optimization, the objective function $f(x)$ is a stochastic function given as the expectation over a random variable/vector ξ:

$$f(x) = \mathbb{E}[F(x,\xi)], \tag{1}$$

where the design parameter x is in a convex and compact set $\mathcal{X}$. The distribution of ξ may not be known, or even if it is known, the expectation over ξ is difficult to evaluate analytically. As a result, the objective function $f(x)$ is unknown, except for the knowledge that it is convex.

The above optimization problem can be cast as a sequential learning problem where the learner chooses a query point $x_t \in \mathcal{X}$ at each time t and observes the corresponding random loss $F(x_t,\xi_t)$ or the random gradient $G(x_t,\xi_t)$. These two feedback models are commonly referred to, respectively, as the zeroth-order and the first-order stochastic optimization. A learning policy governs the selection of the query points $\{x_t\}_{t\geq 1}$ based on past observations, with the objective that x_T converges to the minimizer $x^* = \arg\min_{x\in\mathcal{X}} f(x)$ (or $f(x_T)$ to $f(x^*)$) over a growing horizon of length T.

[0]This work was supported by the National Science Foundation under Grant CCF-1815559.

Under an online formulation of the problem, a more suitable performance measure is the cumulative regret defined as the expected cumulative loss at the query points in excess to the minimum loss: $R(T) = \mathbb{E}\left[\sum_{t=1}^{T}(F(x_t,\xi_t) - f(x^*))\right]$. Under this objective, the query process needs to balance the exploration of the input space $\mathcal{X}$ in search for x^* and the associated loss incurred during the search process. The behavior of regret $R(T)$ over a growing horizon length T is a finer measure than the convergence of x_T or $f(x_T)$. Specifically, a policy with a sublinear regret order in T implies that $f(x_T)$ converges to $f(x^*)$. The converse, however, is not true. In particular, the convergence of x_T to x^* or $f(x_T)$ to $f(x^*)$ does not imply a sublinear, let alone an optimal, order of the regret.

An example of online stochastic convex optimization is the classification of a real-time stream of random instances $\{\xi_t\}_{t\geq 1}$ with each instance given by its feature and hidden label. Without knowing the joint distribution of the feature and label, an online learning policy chooses the classifiers $\{x_t\}_{t\geq 1}$ sequentially over time to produce online classification of the streaming instances. Empirical risk minimization using mini-batching of a large data set can also be viewed as a stochastic optimization problem [4], except that the resulting expectation is with respect to the random drawing of the mini-batches (often uniform with replacement) rather than the true distribution underlying the data generation.

B. Stochastic Gradient Descent

The study of stochastic convex optimization dates back to the seminal work by Robbins and Monro in 1951 [16] under the term "stochastic approximation." The problem studied there is to approximate the root of a monotone function $g(x)$ based on successive observations of random function values at chosen query points (also known as stochastic root finding [14]). The equivalence of this problem to the first-order stochastic convex optimization is immediate when $g(x)$ is the gradient of a convex loss function $f(x)$. The zeroth-order version of the problem was studied in a follow-up work by Kiefer and Wolfowitz [8].

The stochastic gradient descent (SGD) approach developed by Robbins and Monro [16] has long become a classic and is widely used. The basic idea of SGD is to choose the next query point x_{t+1} in the opposite direction of the observed gradient while ensuring $x_{t+1} \in \mathcal{X}$ via a projection operation.

Omitting the projection operation, we can write x_{t+1} as

$$x_{t+1} = x_t - \eta_t G(x_t, \xi_t), \tag{2}$$

where η_t is a properly chosen step-size at time t. Due to the noise effect of the random gradients $G(x_t;\xi_t)$, it is necessary that the step-sizes $\{\eta_t\}_{t\geq 1}$ diminishes to zero to ensure convergence of x_t. Since $G(x_t;\xi_t)$ contains both the signal (the true gradient $g(x_t) = \mathbb{E}[G(x_t;\xi_t)]$) and noise, the diminishing rate of $\{\eta_t\}_{t\geq 1}$ in t needs to be carefully controlled to balance the tradeoff between learning rate and noise attenuation. Naturally, the optimal choice depends on how fast the gradient $g(x)$ approaches to zero as x tends to x^* and the variance of the random gradient samples.

While earlier studies on stochastic approximation focus on the convergence of x_T and $f(x_T)$ (see a survey by Lai in [11]), a series of recent work has established the regret orders of SGD for different classes of functions. As shown in Tabel I, SGD offers $\mathcal{O}(\sqrt{T}\log T)$ regret for convex functions, $\mathcal{O}(\log^2(T))$ regret for $\alpha-$strongly convex functions, and $\mathcal{O}(\log T)$ regret for functions that are non-differentiable at x^*, which are near-optimal[1] as compared to the lower bounds.

To achieve these near-optimal regret orders, however, it is necessary to know which category the underlying unknown objective function $f(x)$ belongs to, as well as nontrivial bounds on the corresponding parameters of the function characteristics (i.e., the parameter α for strong convexity and the jump in the subgradient at x^* when $f(x)$ is non-differentiable at x^*). Such information is crucial in choosing the diminishing rate of the step-sizes $\{\eta_t\}_{t\geq 1}$, and the sensitivity of SGD to model mismatch, estimation errors in the parameters, and ill-conditioning of the functions is well documented.

C. RWT: an Adaptive and Robust Approach

We show in this work that for one-dimensional problems, an alternative approach to stochastic convex optimization self adapts to the function characteristics and offers better or matching regret orders than SGD in each class of functions without assuming any knowledge on the function characteristics. It can also handle heavy-tailed noise with infinite variance, a case for which the applicability of SGD is unclear to our knowledge.

Referred to as Random Walk on a Tree (RWT), this policy was proposed by two of the authors of this paper in a prior work [19] that analyzed its regret performance for convex functions under sub-Gaussian noise distributions. In this paper, we demonstrate the adaptivity of RWT to different function characteristics and robustness to heavy-tailed noise with infinite variance. We also refine the termination thresholds in the local sequence test of RWT based on the law of the iterated logarithm, which leads to improved regret orders.

[1]A number of variants of SGD with various noise-reduction techniques exist in the literature that achieve the optimal regret order (see, for example, [15]). We consider in Table I the basic form of SGD since these noise-reduction techniques often require additional storage and computation resources and may not be suitable for online settings. An additional assumption on the smoothness of the objective function with prior knowledge on the smoothness parameter can also close the gap to the lower bounds [18].

The basic idea of RWT is to construct an infinite-depth binary tree based on successive partitioning of the input space $\mathcal{X}$. Specifically, the root of the tree corresponds to $\mathcal{X}$, which, without loss of generality, is assumed to be $[0,1]$ for the one-dimensional case. The tree grows to infinite depth based on a binary splitting of each node (i.e., the corresponding interval) that forms the two children of the node at the next level.

The query process of RWT is based on a biased random walk on this interval tree that initiates at the root node. Each move of the random walk is guided by a local sequential test based on random gradient realizations drawn from the left boundary, the middle point, and the right boundary of the interval corresponding to the current location of the random walk. The goal of the local sequential test is to determine, with a confidence level greater than $1/2$, whether there is a change of sign in the gradient in the left sub-interval or the right sub-interval of the current node. If one is true (with the chosen confidence level), the walk moves to the corresponding child that sees the sign change. For all other outcomes, the walk moves back to the parent of the current node. The stopping rule and the output of the local sequential test are based on properly constructed lower and upper confidence bounds of the empirical mean (or truncated empirical mean in the case of infinite variance) of the observed gradient realizations. A greater than $1/2$ bias of the random walk is sufficient to ensure convergence to the optimal point x^* at a geometric rate, regardless of the function characteristics.

By bounding the sample complexity of the local sequential test and analyzing the trajectory of the biased random walk, we establish the regret orders of RWT as shown in Table I for sub-Gaussian distributions (a $\log\log T$ factor is omitted; see Sec. IV for the exact orders and finite-time bounds). Similar order-optimal (up to poly-$\log T$ orders) regret performance is also established for heavy-tailed distributions with infinite variance. We are unaware of results on whether SGD can achieve sublinear regret orders under infinite noise variance.

In contrast to SGD that relies on a manually controlled sequence of step-sizes to tradeoff learning rate with noise attenuation, RWT, with no tuning parameters, self adapts to function characteristics through the local sequential test that automatically draws more or fewer samples as demanded by

	convex	strongly convex	non-differentiable at x^*
SGD	$\sqrt{T}\log T$ [18]	$\log^2 T$ [18]	$\log T$ [12]
RWT	$\sqrt{T\log T}$	$\log T$	$\log T$
Lower Bound	$\sqrt{T}$ [2]	$\log T$ [2]	$\log T$ [2]

TABLE I: Regret performance of SGD and RWT under sub-Gaussian noise.

the underlying statistical models. As shown in Table I, RWT outperforms or matches the regret orders of SGD without prior information on the function characteristics.

Another key difference between SGD and RWT is in the induced random walk in the input space $\mathcal{X}$. The unstructured moves of SGD may land at any points in $\mathcal{X}$. RWT, however, queries only a fixed set of countable number of points in $\mathcal{X}$. Furthermore, given the current location on the binary tree, the next move is restricted to only the parent and the two children of this node. This highly structured mobility allows storage-efficient caching of side observations for noise reduction at future query points.

D. Other Related Work

The classical probabilistic bisection algorithm (PBA) has been employed as a solution to stochastic root finding under a one-dimensional input space. Assuming a prior distribution of the optimal point x^*, PBA updates the belief (i.e., the posterior distribution) of x^* based on each observation and subsequently probes the median point of the belief. It was shown in [6] that the regret order of PBA is upper bounded by $\mathcal{O}(T^{0.5+\epsilon})$ for a small $\epsilon > 0$, and an $\mathcal{O}(\sqrt{T}\log T)$ regret order was conjectured.

There may appear to be a connection between RWT and PBA, since both algorithms involve a certain bisection of the input domain. These two approaches are, however, fundamentally different. First, PBA requires the knowledge on the distribution of the random gradient function to perform the belief update, while RWT operates under unknown models. Second, the belief-based bisection in PBA is on the entire input domain $\mathcal{X}$ at each query and needs to be updated based on each random observation. The interval tree in RWT is predetermined, and each move of the random walk leads to a bisection of a *sub-interval* of $\mathcal{X}$ that is shrinking in geometric rate over time with high probability. It is this zooming effect of the biased random walk that leads to a $\mathcal{O}(1)$ computation and memory complexity. For PBA, if $\mathcal{X}$ is discretized to M points for computation and storage, updating and sorting the belief would incur $\mathcal{O}(M \log M)$ computation complexity at each query and linear memory requirement. Lastly, the regret order of RWT outperforms that of PBA.

Under the zeroth-order feedback model where the decision maker has access to the function values, the problem can be viewed as a continuum-armed bandit problem, on which a vast body of results exists. In particular, the work in [1] developed an approach based on the ellipsoid algorithm that achieves an $\mathcal{O}(\sqrt{T}(\log T)^{\frac{3}{2}})$ regret when the objective function f is convex and Lipschitz. The continuum armed bandit under Lipschitz assumption (not necessarily convex) has been studied in [3], [9], [10] where higher orders of regret were shown. The $\mathcal{X}$-armed bandit introduced in [5] considered a Lipschitz function with respect to a dissimilarity function known to the learner. Under the assumption of a finite number of global optima and a particular smoothness property, an $\tilde{\mathcal{O}}(\sqrt{T})$ regret was shown. While the proposed policy in [5] uses a tree structure for updating the indexes in a bandit algorithm, it is fundamentally different from RWT in that the policy does not induce a random walk on the tree. This line of work differs from the gradient-based approach considered in this work. Nevertheless, since an $\mathcal{O}(1)$ number of samples from F can be translated to a sample from G under certain regularity assumptions, gradient-based approaches can be extended to cases where samples from F are directly fed into the learning policy.

We mention that the stochastic online learning setting considered here is different, in problem formulation, objective, and techniques, from an adversarial counterpart of the problem where the loss function is deterministic and adversarially chosen at each time t. On this line of research, see [7], [17] and references therein.

II. Problem Formulation

We aim to minimize a stochastic convex loss function $f(x)$ as given in (1). Let $g(x)$ be the gradient (or sub-gradient) of $f(x)$. Let $G(x,\xi)$ be unbiased random gradient observations with $\mathbb{E}[G(x,\xi)] = g(x)$.

Without knowing $f(x)$ or the stochastic models of $F(x,\xi)$ or $G(x,\xi)$, a learner sequentially chooses the query points $\{x_t\}_{t=1}^{\infty}$, incurs i.i.d. losses $F(x_t,\xi_t)$, and observes i.i.d. gradient samples $G(x_t,\xi_t)$. The objective is to design a learning policy π that is a mapping from past observations to the next query point to minimize the cumulative regret defined as

$$R_\pi(T) = \mathbb{E}\left[\sum_{t=1}^{T}\left(F(x_{\pi(t)},\xi_t) - F(x^*,\xi_t)\right)\right], \tag{3}$$

where $x_{\pi(t)}$ is the query point at time t under policy π.

A. Function Characteristics

The loss function f is said to be convex if and only if

$$f(y) \geq f(x) + g(x)(y-x),\ \forall x,y \in \mathcal{X}. \tag{4}$$

It is α-strongly convex (for some $\alpha > 0$) if and only if

$$f(y) \geq f(x) + g(x)(y-x) + \frac{\alpha}{2}(y-x)^2,\ \forall x,y \in \mathcal{X}. \tag{5}$$

We also consider a nonsmooth case where $f(x)$ is non-differentiable at x^*. This often occurs in optimization problems that involve L1-norm regularization or have discrete parameters [12]. For such functions, there exists a lower bound $\delta > 0$ on the magnitude of the (sub)-gradient:

$$|g(x)| \geq \delta \quad \text{for all } x \neq x^*. \tag{6}$$

In other words, the signal component in the random observations $G(x,\xi)$ does not diminish to zero as x tends to x^*, making $\log T$ regret order possible even under noise with infinite variance.

B. Noise Characteristics

The distribution of $G(x,\xi) - g(x)$ is said to be sub-Gaussian with parameter σ^2 if its moment generating function is bounded by that of a Gaussian random variable with variance σ^2:

$$\mathbb{E}\left[\exp(\lambda\left(G(x,\xi) - g(x)\right))\right] \leq \exp(\frac{\lambda^2\sigma^2}{2}). \tag{7}$$

We also consider heavy-tailed distributions where the only assumption is the existence of a b-th ($b > 1$) moment:

$$\mathbb{E}\left[|G(x,\xi)|^b\right] \le u, \tag{8}$$

for some $u > 0$. Note that this covers the class of distributions of $G(x,\xi)$ with unbounded variance.

III. Random Walk on a Tree

RWT is based on an infinite-depth binary tree with nodes representing a subinterval of $\mathcal{X}$. The 2^l nodes at depth l ($l = 0, 1, 2, \ldots$) of the tree correspond to the intervals resulting from an equal-length partition of $\mathcal{X}$, with each interval of length 2^{-l}. Each node at depth l has two children corresponding to its equal-length subintervals at depth $l+1$. Let $N_{k,l}$ ($k = 1, \ldots, 2^l$, $l = 0, 1, \ldots$) denote the kth node at depth l. We use the terms node and its corresponding interval interchangeably.

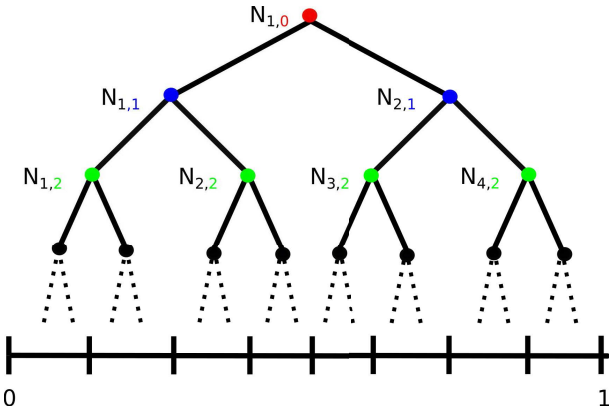

Fig. 1: The binary tree $\mathcal{T}$ representing the subintervals of $[0,1]$. At level 0, $N_{1,0}$ corresponds to the interval $[0,1]$; at level 1, $N_{1,1}$ and $N_{2,1}$, respectively, correspond to the intervals $[0, 0.5]$ and $[0.5, 1]$; and so on.

A. The Biased Random Walk on the Tree

The basic structure of RWT is to carry out a biased random walk on the interval tree. The walk starts at the root of the tree. Each move of the random walk is to one of the three adjacent nodes (i.e., the parent and the two children with the parent of the root defined as itself) of the current location. It is guided by the outputs of a confidence-bound based sequential test carried over the two boundary points and the middle point of the interval currently being visited by the random walk.

Consider a generic sampling point $x \in [0,1]$. The goal of the sequential test is to determine, at a given confidence level, whether $g(x)$ is negative or positive. If the former is true, the test module outputs -1, indicating the target x^* is more likely to lie on the right of the current sampling point x; if the latter is true, the test module outputs 1, indicating the target x^* is more likely to lie on the left of x (see the next subsection on the details of the sequential test).

Based on the binary outcomes of the local sequential tests, the random walk on the tree consists of the following loop until the end of the time horizon. Let $N_{k,l}$ denote the current location of the random walk. The boundary points and the middle point of the interval corresponding to $N_{k,l}$ are probed by the sequential test module. If the output sequence on the left boundary, middle point and the right boundary, in order, is $\{-1, 1, 1\}$ (indicating a sign change in the left subinterval), the walk moves to the left child of $N_{k,l}$. If the output sequence is $\{-1, -1, 1\}$, the walk moves to the right child of $N_{k,l}$. For all other output sequences, the walk moves back to the parent of $N_{k,l}$.

B. The Local Sequential Test

We now specify the local sequential test at a generic query point x. The test sequentially draws random gradient samples $G(x,\xi)$. After collecting each sample, it determines whether to terminate the test and if yes, which value to output. The termination and decision rules are chosen to satisfy a confidence level $1 - \check{p}$ to ensure that the resulting random walk is biased toward the target x^*, where $\check{p}$ can be any value in $(0, 1 - \frac{1}{\sqrt[3]{2}})$. By convention, we define the output of the test at $x = 0$ to be -1, and at $x = 1$ to be 1, without performing the test.

The construction of the termination rule exploits the law of the iterated logarithm and depends on the noise characteristics. We consider separately the cases of sub-Gaussian and heavy-tailed distributions.

1) Sub-Gaussian Distributions: For sub-Gaussian distributed gradient samples, the test statistics can simply be the sample mean $\overline{g}_s(x)$ given by

$$\overline{g}_s(x) = \frac{1}{s}\sum_{t=1}^{s} G(x,\xi_t). \tag{9}$$

For a given confidence level parameter $\check{p}$, the sequential test is given Fig. 2, where σ^2 is the sub-Gaussian parameter specified in (7).

▷ If $\overline{g}_s(x) > \sqrt{\frac{5\sigma^2}{s}\log\left(\frac{6\log s}{\check{p}}\right)}$, terminate; output 1.

▷ If $\overline{g}_s(x) < -\sqrt{\frac{5\sigma^2}{s}\log\left(\frac{6\log s}{\check{p}}\right)}$, terminate; output -1.

▷ Otherwise, take another sample of $G(x,\xi)$ and repeat.

Fig. 2: The sequential test at a sampling point x under sub-Gaussian noise.

2) Heavy-Tailed Distributions: For heavy-tailed distributions with a bounded b-th ($b > 1$) moment, define the truncated sample mean of gradient obtained from s observations under a given confidence-level parameter $\check{p} > 0$ as follows:

$$\widehat{g}_{s,\check{p}}(x) = \frac{1}{s}\sum_{t=1}^{s} G(x,\xi_t)\mathbb{1}\{|G(x,\xi_t)| \le B_t\}, \tag{10}$$

where

$$\begin{aligned} B_t &= B_0 \left(\frac{t}{\lambda(t)}\right)^{\frac{1}{b}}, \\ \lambda(t) &= 10^b \log\left(\frac{12\max\{\log(t),2\}}{b\sqrt{\breve{p}}}\right), \\ B_0 &= \max\left\{\left(\frac{2^{\frac{2+b}{b}}}{\lambda(1)^{\frac{2-b}{b}}}\frac{15u}{3-\sqrt{2}}\right)^{\frac{1}{b}}, \right. \\ &\qquad \left.\left(\frac{4\sqrt{2}u\log 2}{\sqrt{\log(\log 3)}}\right)^{\frac{1}{b}}\right\}. \end{aligned} \tag{11}$$

In the truncated sample mean, the t-th sample is compared to a threshold B_t and replaced with 0 if its value exceeds the threshold. The resulting sequential test is given in Fig. 3, where u is the bound on the b-th moment as given in (10).

▷ If $\widehat{g}_{s,\breve{p}}(x) > \sqrt{\frac{B_0^2}{2}s^{\frac{2-2b}{b}}\log\left(\frac{12\log s}{b\sqrt{\breve{p}}}\right)} - \frac{1}{s}\sum_{t=1}^{s}\frac{u}{B_t^{b-1}}$, terminate; output 1.

▷ If $\widehat{g}_{s,\breve{p}}(x) < -\sqrt{\frac{B_0^2}{2}s^{\frac{2-2b}{b}}\log\left(\frac{12\log s}{b\sqrt{\breve{p}}}\right)} + \frac{1}{s}\sum_{t=1}^{s}\frac{u}{B_t^{b-1}}$, terminate; output -1.

▷ Otherwise, take another sample of $G(x,\xi)$ and repeat.

Fig. 3: The sequential test at a sampling point x under heavy-tailed noise.

IV. Regret Analysis

In this section, we provide regret analysis of RWT under variant function and noise characteristics. Corresponding to the two components—the global random walk and the local sequential test—of the policy, the analysis builds on establishing the convergence rate of the random walk towards x^* and the sample complexity of the sequential test. Each is given in a lemma in the subsequent sections. All proofs are omitted due to the space limit and are available online in a full version of the paper [20].

A. The Geometric Convergence Rate of the Random Walk

Let n denote the index of the steps taken by the random walk. Let $x_{(n)}$ denote the position of the random walk after n steps. In particular, $x_{(0)}$ is the root node. Let $\Delta_{x_{(n)}} = \max_{x\in x_{(n)}}|x-x^*|$ denote the maximum distance between a point in the interval corresponding to $x_{(n)}$ and x^*. Lemma 1 establishes a high-probability upper bound on $\Delta_{x_{(n)}}$ after n steps are taken by the random walk.

Lemma 1. *With probability at least* $1-\exp(-\frac{n(2p-1)^2}{2})$*, we have*

$$\Delta_{x_{(n)}} \le 2^{-\frac{n(2p-1)}{2}}, \tag{12}$$

where $p \ge (1-\breve{p})^3 > \frac{1}{2}$ *is the bias of the walk.*

Lemma 1 shows that the random walk converges at a geometric rate to x^*. Notice that this result is independent of the characteristics of the function or noise.

B. The Sample Complexity of the Local Sequential Test

1) Sub-Gaussian Distributions: The following lemma gives an upper bound on the sample complexity and error probability of the local sequential test under sub-Gaussian distributions.

Lemma 2. *Let* $\tau(x)$ *denote the termination time of the local sequential test at an arbitrary query point* $x\in\mathcal{X}$ *as given in Fig. 2. Under sub-Gaussian distributions defined in* (7)*, the sample complexity* $\mathbb{E}[\tau(x)]$ *of the local sequential test is given by*

$$\mathbb{E}[\tau(x)] \le \frac{40\sigma^2}{g^2(x)}\log\left(\frac{12}{\sqrt{\breve{p}}}\log\left(\frac{240\sigma^2}{\sqrt{\breve{p}}g^2(x)}\right)\right)+2. \tag{13}$$

The probabilities of an incorrect test outcome under each hypothesis on the sign of $g(x)$ *are bounded as follows:*

$$\begin{aligned} &\mathbb{P}\left[\overline{g}_\tau(x) > \sqrt{\tfrac{5\sigma^2}{\tau}\log\left(\tfrac{6\log\tau}{\sqrt{\breve{p}}}\right)} \mid g(x)<0\right] \le \breve{p}, \\ &\mathbb{P}\left[\overline{g}_\tau(x) < -\sqrt{\tfrac{5\sigma^2}{\tau}\log\left(\tfrac{6\log\tau}{\sqrt{\breve{p}}}\right)} \mid g(x)>0\right] \le \breve{p}. \end{aligned} \tag{14}$$

Lemma 2 shows that the error probability of the sequential test at all query points $x\in\mathcal{X}$ is upper bounded by $\breve{p}$. The condition for the random walk to move in the right direction is that the output of all three tests carried out on the boundary points and the middle point of the current interval are correct. Thus, the probability p that the random walk moves in the right direction satisfies $p\ge(1-\breve{p})^3$ which indicates $p>\frac{1}{2}$ by the choice of $\breve{p}\in(0,1-\frac{1}{\sqrt[3]{2}})$. This ensures that the random walk is biased toward x^* as required for the geometric convergence of the random walk as specified in Lemma 1.

To bound the test error, we employ techniques similar to the ones used in the proof of the law of iterated logarithm. By bounding the error probability for geometrically increasing intervals, the total probability of error can be bounded using the union sum and the convergence for the Riemann Zeta function for index greater than 1. The upper bound $\breve{p}$ on the error probabilities is ensured by choosing appropriate constants in the termination threshold.

2) Heavy-Tailed Distributions: Analogous to Lemma 2, we have the following result on the sample complexity and error probability of the sequential test under heavy-tailed distributions.

Lemma 3. *Let* $\tau(x)$ *denote the termination time of the local sequential test at an arbitrary query point* $x\in\mathcal{X}$ *as given in Fig. 3. Under heavy-tailed distributions satisfying the bounded b-th* ($b>1$) *moment condition given in* (8)*, the sample complexity* $\mathbb{E}[\tau(x)]$ *of the local sequential test is given by*

$$\mathbb{E}[\tau(x)] \leq \frac{3b-2}{2b-2} \left(\left(\frac{2B_0^2}{g^2(x)} \log \left(\frac{9}{c_b} \log \left(\frac{36B_0^2}{g^2(x)c_b} \right) \right) \right)^{\frac{b}{2(b-1)}} + 1 \right),$$

where $c_b = (b-1)\sqrt{\check{p}}$. The probabilities of an incorrect test outcome under each hypothesis on the sign of $g(x)$ are upper bounded by $\check{p}$.

C. The Cumulative Regret

We are now ready to provide the regret performance of RWT under various cases of the function characteristics (convex, strongly convex, non-differentiable at x^*) and noise characteristics (sub-Gaussian, heavy-tailed).

1) Sub-Gaussian Distributions: The following theorem provides upper bound on regret of RWT under sub-Gaussian distributions. The regret order varies based on the function characteristics.

Theorem 1. *Let $\check{p} \in (0, 1 - \frac{1}{\sqrt[3]{2}})$ be the chosen parameter of the sequential test and p the resulting bias of the random walk. Let $g_{max} = \max_{x \in \mathcal{X}} g(x)$. For sub-Gaussian distributions with parameter σ^2, the regret of RWT is upper bounded as follows.*

- *For convex functions,*

$$R_{RWT}(T) \leq \frac{6}{2p-1} \sqrt{10\sigma^2 T \log T \log \left(\frac{12}{\sqrt{\check{p}}} \log \frac{2(2p-1)^2 T}{3 \log T \sqrt{\check{p}}} \right)} + \frac{3}{2p-1} \sqrt{2T \log T} + g_{max}(\log T + 4).$$

- *For α-strong convexity functions,*

$$R_{RWT}(T) \leq \frac{360\sigma^2 \log T}{2\alpha(2p-1)^2} \log \left(\frac{12}{\sqrt{\check{p}}} \log \frac{2(2p-1)^2 T}{3 \log T \sqrt{\check{p}}} \right) + \frac{18 \log T}{2\alpha(2p-1)^2} + g_{max}(\log T + 4). \quad (15)$$

- *For functions that are non-differentiablity at x^* with a $\delta > 0$ lower bound on the magnitude of gradient,*

$$R_{RWT}(T) \leq \frac{9 g_{max} \log T}{(2p-1)^2} \left(\frac{40\sigma^2}{\delta^2} \log \left(\frac{12}{\sqrt{\check{p}}} \log \frac{240\sigma^2}{\sqrt{\check{p}}\delta^2} \right) + 2 \right) + g_{max}(\log T + 4). \quad (16)$$

Theorem 1 shows $\mathcal{O}(\sqrt{T \log T \log \log T})$, $\mathcal{O}(\log T \log \log T)$ and $\mathcal{O}(\log T)$ regrets for objective functions $f(x)$ that are convex, α-strongly convex, and non-differentiable at x^*, respectively. Note that while the confidence parameter $\check{p}$ affects the leading constants of the regret, choosing any value in $(0, 1 - \frac{1}{\sqrt[3]{2}})$ ensures these regret orders. These (near-)optimal regret orders are thus achieved without any tuning parameter or prior knowledge of the function characteristics.

2) Heavy-Tailed Distributions: We have the following corresponding theorem for heavy-tailed distributions.

Theorem 2. *Let $\check{p} \in (0, 1 - \frac{1}{\sqrt[3]{2}})$ be the chosen parameter of the sequential test and p the resulting bias of the random walk. Let $g_{max} = \max_{x \in \mathcal{X}} g(x)$. Under heavy-tailed distributions satisfying the bounded b-th ($b > 1$) moment condition in* (8)*, the regret of RWT is upper bounded as follows.*

- *For convex functions,*

$$R_{RWT}(T) \leq \sqrt{2} B_0 \left(\frac{9(3b-2)}{(2b-2)(2p-1)^2} \right)^{\frac{b-1}{b}} T^{\frac{1}{b}} \log T^{\frac{b-1}{b}} \sqrt{\log \left(\frac{9}{c_b} \log \left(\frac{4(2b-2)(2p-1)^2 T}{c_b(3b-2)\log T} \right)^{\frac{b}{2(b-1)}} \right)} + \left(\frac{9(3b-2)}{(2b-2)(2p-1)^2} \right)^{\frac{b-1}{b}} T^{\frac{1}{b}} + g_{max}(\log T + 4).$$

- *For α-strong convexity functions,*

$$R_{RWT}(T) \leq \frac{B_0^2}{\alpha} \left(\frac{9(3b-2)}{(2b-2)(2p-1)^2} \right)^{\frac{2(b-1)}{b}} T^{\frac{2-b}{b}} \log T^{\frac{2(b-1)}{b}} \log \left(\frac{9}{c_b} \log \left(\frac{4(2b-2)(2p-1)^2 T}{c_b(3b-2)\log T} \right)^{\frac{b}{2(b-1)}} \right) + \frac{1}{2\alpha} \left(\frac{9(3b-2)}{(2b-2)(2p-1)^2} \right)^{\frac{2(b-1)}{b}} T^{\frac{2-b}{b}} + g_{max}(\log T + 4).$$

- *For functions that are non-differentiablity at x^* with a $\delta > 0$ lower bound on the magnitude of gradient,*

$$R_{RWT}(T) \leq \frac{9 g_{max} \log T (3b-2)}{(2p-1)^2(2b-2)} \left(\frac{2B_0^2}{\delta^2} \log \left(\frac{36}{4c_b} \log \left(\frac{36B_0^2}{\delta^2 c_b} \right) \right)^{\frac{b}{2(b-1)}} + 1 \right) + g_{max}(\log T + 4).$$

Theorem 2 shows $\mathcal{O}(T^{\frac{1}{b}} \log T^{\frac{b-1}{b}} (\log \log T)^{\frac{b}{2(b-1)}})$, $\mathcal{O}(T^{\frac{2-b}{b}} \log T^{\frac{2(b-1)}{b}} (\log \log T)^{\frac{b}{2(b-1)}})$ and $\mathcal{O}(\log T)$ regrets for functions that are convex, α-strongly convex, and non-differentiable at x^*, respectively. They match the corresponding lower bounds [13] (up to poly-$\log T$ factors in the first two cases).

V. Conclusion

We gave a relatively complete regret analysis of the Random-Walk-on-a-Tree (RWT) policy for stochastic convex optimization under various function and noise characteristics. Comparing with the popular SGD approach which requires careful tuning of the step-sizes based on prior knowledge of the function characteristics, RWT, with no tuning parameters, self adapts to unknown function characteristics and offers better or matching regret orders as SGD. The adaptivity is achieved via a local sequential test with termination thresholds

designed based on the law of the iterated logarithm. The highly structured random walk also enables storage-efficient local data cashing for noise reduction at future query points. We further established (near-)optimal regret orders for RWT under heavy-tailed noise with unbounded variance. This is another advantage of RWT over SGD, which, to the best of our knowledge, requires a finite variance of the noise in gradient observations. Our ongoing work on extending RWT to high-dimensional problems by integrating it with coordinate minimization has shown promising results.

REFERENCES

[1] A. Agrawal, D.P. Foster, D. Hsu, S.M. Kakade, A. Rakhlin, "Stochastic convex optimization with bandit feedback," *Advances in Neural Information Processing Systems, 24*, 2011.

[2] A. Agrawal, P. Bartlett, P. Ravikumar, and M. Wainwright, "Information-theoretic lower bounds on the oracle complexity of stochastic convex optimization," *IEEE Transactions on Information Theory,* 58(5):3235–3249, 2012.

[3] R. Agrawal, "The continuum-armed bandit problem," *SIAM journal on control and optimization*, no. 33, pp. 1926, 1995.

[4] L. Bottou, F. E. Curtis, J. Nocedal, "Optimization methods for large-scale machine learning," *Siam Review*, vol. 60, no. 2, pp. 223-311, 2018.

[5] S. Bubeck, R. Munos, G. Stolz, C. Szepesvari, "$\mathcal{X}$-armed Bandits," *Journal of Machine Learning Research 12*. pp. 1655-1695. 2011.

[6] P. I. Frazier, S. G. Henderson, R. Waeber, "Probabilistic bisection converges almost as quickly as stochastic approximation", available at *arXiv:1612.03964v1 [math.PR]*, 2016.

[7] E. Hazan, "Introduction to Online Convex Optimization," *Foundations and Trends in Optimization*, vol. 2, no. 3-4, pp 157-325, 2016.

[8] J. Kiefer, J. Wolfowitz, "Stochastic Estimation of the Maximum of a Regression Function," *Ann. Math. Statist.*, vol. 23, no. 3, pp. 462-466, 1952.

[9] R. Kleinberg, "Nearly tight bounds for the continuum-armed bandit problem," *Advances in Neural Information Processing Systems, 18*, 2005.

[10] R. Kleinberg, A. Slivkins, and E. Upfal, "Multi-armed bandits in metric spaces," *In Proceedings of the 40th annual ACM symposium on Theory of computing,* pp. 681–690. ACM, 2008.

[11] T. Lai, "Stochastic Approximation," *Annuals of Statistics,* no 31, pp. 391–406, 2003.

[12] E. Lim, "On the Convergence Rate for Stochastic Approximation in the Nonsmooth Setting," *Mathematics of Operations Research*, vol. 36, no. 3, pp. 527-537, 2011.

[13] A.S. Nemirovski and D.B. Yudin, *Problem Complexity and Method Efficiency in Optimization*, Wiley, 1983.

[14] R. Pasupathy, S. Kim, "The stochastic root-finding problem: overview, solutions, and open questions," *ACM Transactions on Modeling and Computational Simulations*, vol. 21, no. 3, pp. 19, 2011.

[15] A. Rakhlin, O. Shamir, and K. Sridharan, "Making gradient descent optimal for strongly convex stochastic optimization," *In Proceedings of ICML*, 2012.

[16] H. Robbins and S. Monro "A stochastic approximation method" *The Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, 1951.

[17] S. Shalev-Shwartz, "Online Learning and Online Convex Optimization," *Foundations and Trends in Optimization*, vol. 4, no. 2, pp 107-194, 2012.

[18] O. Shamir, T. Zhang, "Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes," *Proceedings of the 30th International Conference on Machine Learning*, 2013.

[19] S. Vakili, Q. Zhao, "A Random Walk Approach to First-Order Stochastic Convex Optimiza- tion," *International Symposium on Information Theory (ISIT)*, 2019.

[20] S. Vakili, S. Salgia, Q. Zhao, "Stochastic Gradient Descent on a Tree: an Adaptive and Robust Approach to Stochastic Convex Optimization," available at https://arxiv.org/abs/1901.05947.