Journal of Machine Learning Research 18 (2017) 1-11 Submitted 12/16; Published 5/17

An Optimal Algorithm for Bandit and Zero-Order Convex Optimization with Two-Point Feedback

Ohad Shamir OHAD.SHAMIR@WEIZMANN.AC.IL
Department of Computer Science and Applied Mathematics
Weizmann Institute of Science
Rehovot 7610001, Israel

Editor: Alexander Rakhlin

Abstract

We consider the closely related problems of bandit convex optimization with two-point feedback, and zero-order stochastic convex optimization with two function evaluations per round. We provide a simple algorithm and analysis which is optimal for convex Lipschitz functions. This improves on Duchi et al. (2015), which only provides an optimal result for smooth functions; Moreover, the algorithm and analysis are simpler, and readily extend to non-Euclidean problems. The algorithm is based on a small but surprisingly powerful modification of the gradient estimator.

Keywords: zero-order optimization, bandit optimization, stochastic optimization, gradient estimator

1. Introduction

We consider the problem of bandit convex optimization with two-point feedback Agarwal et al. (2010). This problem can be defined as a repeated game between a learner and an adversary as follows: At each round t, the adversary picks a convex function f_t on $\mathbb{R}^d$, which is not revealed to the learner. The learner then chooses a point $\mathbf{w}_t$ from some known and closed convex set $\mathcal{W} \subseteq \mathbb{R}^d$, and suffers a loss $f_t(\mathbf{w}_t)$. As feedback, the learner may choose two points $\mathbf{w}'_t, \mathbf{w}''_t \in \mathcal{W}$ and receive[1] $f_t(\mathbf{w}'_t), f_t(\mathbf{w}''_t)$. The learner's goal is to minimize average regret, defined as

$$\frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{w}_t) - \min_{\mathbf{w}\in\mathcal{W}} \frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{w}).$$

In this paper, we focus on obtaining bounds on the expected average regret (with respect to the learner's randomness).

A closely-related and easier setting is zero-order stochastic convex optimization. In this setting, our goal is to approximately solve $F(\mathbf{w}) = \min_{\mathbf{w}\in\mathcal{W}} \mathbb{E}_{\xi}[f(\mathbf{w};\xi)]$, given limited access to $\{f(\cdot;\xi_t)\}_{t=1}^{T}$ where ξ_t are i.i.d. instantiations. Specifically, we assume that each

1. This is slightly different than the model of Agarwal et al. (2010), where the learner only chooses $\mathbf{w}'_t, \mathbf{w}''_t$ and the loss is $\frac{1}{2}\left(f_t(\mathbf{w}'_t) + f_t(\mathbf{w}''_t)\right)$. However, our results and analysis can be easily translated to their setting, and the model we discuss translates more directly to the zero-order stochastic optimization considered later.

$f(\cdot, \xi_t)$ is not directly observed, but rather can be queried at two points. This models situations where computing gradients directly is complicated or infeasible. It is well-known (Cesa-Bianchi et al., 2004) that given an algorithm with expected average regret R_T in the bandit optimization setting above, if we feed it with the functions $f_t(\mathbf{w}) = f(\mathbf{w}; \xi_t)$, then the average $\bar{\mathbf{w}}_T = \frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t$ of the points generated satisfies the following bound on the expected optimization error:

$$\mathbb{E}[F(\bar{\mathbf{w}}_T)] - \min_{\mathbf{w}\in\mathcal{W}} F(\mathbf{w}) \leq R_T.$$

Thus, an algorithm for bandit optimization can be converted to an algorithm for zero-order stochastic optimization with similar guarantees.

The bandit optimization setting with two-point feedback was proposed and studied in Agarwal et al. (2010). Independently, Nesterov (2011) and considered two-point methods for stochastic optimization. Both papers are based on randomized gradient estimates which are then fed into standard first-order algorithms (e.g. gradient descent, or more generally mirror descent). However, the regret/error guarantees in both papers were suboptimal in terms of the dependence on the dimension. Recently, Duchi et al. (2015) considered a similar approach for the stochastic optimization setting, attaining an optimal error guarantee when $f(\cdot; \xi)$ is a smooth function (differentiable and with Lipschitz-continuous gradients). Related results in the smooth case were also obtained by Ghadimi and Lan (2013). However, to tackle the general case, where $f(\cdot; \xi)$ may be non-smooth, Duchi et al. (2015) resorted to a non-trivial smoothing scheme and a significantly more involved analysis. The resulting bounds have additional factors (logarithmic in the dimension) compared to the guarantees in the smooth case. Moreover, an analysis is only provided for Euclidean problems (where the domain $\mathcal{W}$ and Lipschitz parameter of f_t scale with the L_2 norm).

In this note, we present and analyze a simple algorithm with the following properties:

- For Euclidean problems, it is optimal up to constants for both smooth and non-smooth functions. This closes the gap between the smooth and non-smooth Euclidean problems in this setting.

- The algorithm and analysis are readily applicable to non-Euclidean problems. We give an example for the 1-norm, with the resulting bound optimal up to logarithmic factors.

- The algorithm and analysis are simpler than those proposed in Duchi et al. (2015). They apply equally to the bandit and zero-order optimization setting, and can be readily extended using standard techniques, e.g. improved bounds for strongly-convex functions; regret/error bounds holding with high-probability rather than just in expectation; and improved bounds if allowed $k > 2$ observations per round instead of just two (Hazan et al., 2007; Shalev-Shwartz, 2007; Agarwal et al., 2010).

Like previous algorithms, our algorithm is based on a random gradient estimator, which given a function f and point $\mathbf{w}$, queries f at two random locations close to $\mathbf{w}$, and computes a random vector whose expectation is a gradient of a smoothed version of f. The papers Nesterov (2011); Duchi et al. (2015); Ghadimi and Lan (2013) essentially use the estimator which queries at $\mathbf{w}$ and $\mathbf{w} + \delta\mathbf{u}$ (where $\mathbf{u}$ is a random unit vector and $\delta > 0$ is a small parameter), and returns

$$\frac{d}{\delta}\left(f(\mathbf{w}+\delta\mathbf{u}) - f(\mathbf{w})\right)\mathbf{u}. \tag{1}$$

The intuition is readily seen in the one-dimensional ($d = 1$) case, where the expectation of this expression equals

$$\frac{1}{2\delta}\left(f(w+\delta) - f(w-\delta)\right), \tag{2}$$

which indeed approximates the derivative of f (assuming f is differentiable) at w, if δ is small enough.

In contrast, our algorithm uses a slightly different estimator (also used in Agarwal et al., 2010), which queries at $\mathbf{w} - \delta\mathbf{u}, \mathbf{w} + \delta\mathbf{u}$, and returns

$$\frac{d}{2\delta}\left(f(\mathbf{w}+\delta\mathbf{u}) - f(\mathbf{w}-\delta\mathbf{u})\right)\mathbf{u}. \tag{3}$$

Again, the intuition is readily seen in the case $d = 1$, where the expectation of this expression also equals Eq. (2).

When δ is sufficiently small and f is differentiable at $\mathbf{w}$, both estimators compute a good approximation of the true gradient $\nabla f(\mathbf{w})$. However, when f is not differentiable, the variance of the estimator in Eq. (1) can be quadratic in the dimension d, as pointed out by Duchi et al. (2015): For example, for $f(\mathbf{w}) = \|\mathbf{w}\|_2$ and $\mathbf{w} = 0$, the second moment equals

$$\mathbb{E}\left[\left\|\frac{d}{\delta}\left(f(\delta\mathbf{u}) - f(\mathbf{0})\right)\mathbf{u}\right\|_2^2\right] = \mathbb{E}\left[d^2\|\mathbf{u}\|_2^4\right] = d^2.$$

Since the performance of the algorithm crucially depends on the second moment of the gradient estimate, this leads to a highly sub-optimal guarantee. In Duchi et al. (2015), this was handled by adding an additional random perturbation and using a more involved analysis. Surprisingly, it turns out that the slightly different estimator in Eq. (3) does not suffer from this problem, and its second moment is essentially *linear* in the dimension d.

We note that in this work, we assume that $\mathbf{u}$ is a random unit vector, similar to previous works. However, our results can be readily extended to other distributions, such as uniform in the Euclidean unit ball, or a Gaussian distribution.

2. Algorithm and Main Results

We consider the algorithm described in Figure 1, which performs standard mirror descent using a randomized gradient estimator $\tilde{\mathbf{g}}_t$ of a (smoothed) version of f_t at point $\mathbf{w}_t$. Following Duchi et al. (2015), we assume that one can indeed query f_t at any point $\mathbf{w}_t + \delta_t\mathbf{u}_t$ as specified in the algorithm[2].

The analysis of the algorithm is presented in the following theorem:

2. This may require us to query at a distance δ_t outside $\mathcal{W}$. If we must query within $\mathcal{W}$, then a standard technique (see Agarwal et al., 2010) is to simply run the algorithm on a slightly smaller set $(1-\epsilon)\mathcal{W}$, where $\epsilon > 0$ is sufficiently large so that $\mathbf{w}_t + \delta_t\mathbf{u}_t$ must be in $\mathcal{W}$. Since the formal guarantee in Thm. 1 holds for arbitrarily small δ_t, and each f_t is Lipschitz, we can generally take δ_t (and hence ϵ) sufficiently small so that the additional regret/error incurred is arbitrarily small.

Algorithm 1 Two-Point Bandit Convex Optimization Algorithm

Input: Step size η, function $r : \mathcal{W} \mapsto \mathbb{R}$, exploration parameters $\delta_t > 0$
Initialize $\boldsymbol{\theta}_1 = \mathbf{0}$.
for $t = 1, \ldots, T-1$ **do**
  Predict $\mathbf{w}_t = \arg\max_{\mathbf{w}\in\mathcal{W}} \langle \boldsymbol{\theta}_t, \mathbf{w}\rangle - r(\mathbf{w})$
  Sample $\mathbf{u}_t$ uniformly from the Euclidean unit sphere $\{\mathbf{w} : \|\mathbf{w}\|_2 = 1\}$
  Query $f_t(\mathbf{w}_t + \delta_t \mathbf{u}_t)$ and $f_t(\mathbf{w}_t - \delta_t \mathbf{u}_t)$
  Set $\tilde{\mathbf{g}}_t = \frac{d}{2\delta_t}\left(f_t(\mathbf{w}_t + \delta_t\mathbf{u}_t) - f_t(\mathbf{w}_t - \delta_t\mathbf{u}_t)\right)\mathbf{u}_t$
  Update $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta\tilde{\mathbf{g}}_t$
end for

Theorem 1 *Assume the following conditions hold:*

1. *r is 1-strongly convex with respect to a norm $\|\cdot\|$, and $\sup_{\mathbf{w}\in\mathcal{W}} r(\mathbf{w}) \leq R^2$ for some $R < \infty$.*
2. *f_t is convex and G_2-Lipschitz with respect to the 2-norm $\|\cdot\|_2$.*
3. *The dual norm $\|\cdot\|_*$ of $\|\cdot\|$ is such that $\sqrt[4]{\mathbb{E}_{\mathbf{u}_t}\|\mathbf{u}_t\|_*^4} \leq p_*$ for some $p_* < \infty$.*

If $\eta = \frac{R}{p_* G_2 \sqrt{dT}}$, and δ_t chosen such that $\delta_t \leq p_* R\sqrt{\frac{d}{T}}$, then the sequence $\mathbf{w}_1, \ldots, \mathbf{w}_T$ generated by the algorithm satisfies the following for any T and $\mathbf{w}^* \in \mathcal{W}$:

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{w}_t) - \frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{w}^*)\right] \leq c\; p_* G_2 R\sqrt{\frac{d}{T}},$$

where c is some numerical constant.

We note that condition 1 is standard in the analysis of the mirror-descent method (see the specific corollaries below), whereas conditions 2 and 3 are needed to ensure that the variance of our gradient estimator is controlled.

As mentioned earlier, the bound on the average regret which appears in Thm. 1 immediately implies a similar bound on the error in a stochastic optimization setting, for the average point $\bar{\mathbf{w}}_T = \frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t$. We note that the result is robust to the choice of η, and is the same up to constants as long as $\eta = \Theta(R/p_* G_2\sqrt{dT})$. Also, the constant c, while always strictly positive, shrinks as $\delta_t \to 0$ (see the proof below for details).

As a first application of the theorem, let us consider the case where $\|\cdot\|$ is the Euclidean norm $\|\cdot\|_2$. In this case, we can take $r(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|_2^2$, and the algorithm reduces to a standard variant of online gradient descent, defined as $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \tilde{\mathbf{g}}_t$ and $\mathbf{w}_t = \arg\min_{\mathbf{w}\in\mathcal{W}} \|w - \boldsymbol{\theta}_t\|_2$. In this case, we get the following corollary:

Corollary 2 *Suppose f_t for all t is G_2-Lipschitz with respect to the Euclidean norm, and $\mathcal{W} \subseteq \{\mathbf{w} : \|\mathbf{w}\|_2 \leq R\}$. Then using $\|\cdot\| = \|\cdot\|_2$ and $r(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|_2^2$, it holds for some constant c and any $\mathbf{w}^* \in \mathcal{W}$ that*

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{w}_t) - \frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{w}^*)\right] \leq c\; G_2 R\sqrt{\frac{d}{T}},$$

The proof is immediately obtained from Thm. 1, noting that $p_* = 1$ in our case. This bound matches (up to constants) the lower bound in Duchi et al. (2015), hence closing the gap between upper and lower bounds in this setting.

As a second application, let us consider the case where $\|\cdot\|$ is the 1-norm, $\|\cdot\|_1$, the domain $\mathcal{W}$ is the simplex in $\mathbb{R}^d$, $d > 1$ (although our result easily extends to any subset of the 1-norm unit ball), and we use a standard entropic regularizer:

Corollary 3 *Suppose f_t for all t is G_1-Lipschitz with respect to the L_1 norm. Then using $\|\cdot\| = \|\cdot\|_1$ and $r(\mathbf{w}) = \sum_{i=1}^d w_i \log(dw_i)$, it holds for some constant c and any $\mathbf{w}^* \in \mathcal{W}$ that*

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^T f_t(\mathbf{w}_t) - \frac{1}{T}\sum_{t=1}^T f_t(\mathbf{w}^*)\right] \le c\ G_1\sqrt{\frac{d\log^2(d)}{T}}.$$

This bound matches (this time up to a factor polylogarithmic in d) the lower bound in Duchi et al. (2015) for this setting.

Proof The function r is 1-strongly convex with respect to the 1-norm (see for instance Shalev-Shwartz, 2012, Example 2.5), and has value at most $\log(d)$ on the simplex. Also, if f_t is G_1-Lipschitz with respect to the 1-norm, then it must be $\sqrt{d}G_1$-Lipschitz with respect to the Euclidean norm. Finally, to satisfy condition 3 in Thm. 1, we upper bound $\sqrt[4]{\mathbb{E}[\|\mathbf{u}_t\|_\infty^4]}$ using the following lemma, whose proof is given in the appendix:

Lemma 4 *If $\mathbf{u}$ is uniformly distributed on the unit sphere in $\mathbb{R}^d$, $d > 1$, then $\sqrt[4]{\mathbb{E}[\|\mathbf{u}\|_\infty^4]} \le c\sqrt{\frac{\log(d)}{d}}$ where c is a positive numerical constant independent of d.*

Plugging these observations into Thm. 1 leads to the desired result. ∎

Finally, we make two additional remarks on possible extensions and improvements to Thm. 1.

Remark 5 (Querying at $k > 2$ points) *If the algorithm is allowed to query f_t at $k > 2$, then it can be modified to attain an improved regret bound, by computing $\lfloor k/2 \rfloor$ independent estimates of $\tilde{g}_t$ at every round (using a freshly sampled $\mathbf{u}_t$ each time), and using their average. This leads to a new gradient estimator $\tilde{\mathbf{g}}_t^k$, which satisfies $\mathbb{E}[\|\tilde{\mathbf{g}}_t^k\|^2] \le \frac{1}{k}\mathbb{E}[\|\tilde{\mathbf{g}}_t\|^2] + \|\mathbb{E}[\tilde{\mathbf{g}}_t]\|^2$. Based on the proof of Thm. 1, it is easily verified that this leads to an average expected regret bound of $\frac{cG_2R}{\sqrt{T}}\left(1 + p_*\sqrt{d/k}\right)$ for some numerical constant c.*

Remark 6 (Non-Euclidean Geometries) *When considering norms other than the Euclidean norm, it is tempting to conjecture that our algorithm and analysis can be improved, by sampling $\mathbf{u}_t$ from a distribution adapted to the geometry of that norm (not necessarily the Euclidean ball), and assuming f_t is Lipschitz w.r.t. the dual norm. However, adapting the proof (and in particular getting appropriate versions of Lemma 8 and Lemma 9) does not appear straightforward, and the potential performance improvement is currently unclear.*

3. Proof of Theorem 1

As discussed in the introduction, the key to getting improved results compared to previous papers is the use of a slightly different random gradient estimator, which turns out to have significantly less variance. The formal proof relies on a few simple lemmas listed below. The key lemma is Lemma 10, which establishes the improved variance behavior.

Lemma 7 *For any $\mathbf{w}^* \in \mathcal{W}$, it holds that*

$$\sum_{t=1}^{T} \langle \tilde{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{w}^* \rangle \leq \frac{1}{\eta} R^2 + \eta \sum_{t=1}^{T} \|\tilde{\mathbf{g}}_t\|_*^2.$$

This lemma is the canonical result on the convergence of online mirror descent, and the proof is standard (see e.g. Shalev-Shwartz, 2012).

Lemma 8 *Define the function*

$$\hat{f}_t(\mathbf{w}) = \mathbb{E}_{\mathbf{u}_t} \left[f_t(\mathbf{w} + \delta_t \mathbf{u}_t) \right],$$

over $\mathcal{W}$, where $\mathbf{u}_t$ is a vector picked uniformly at random from the Euclidean unit sphere. Then the function is convex, Lipschitz with constant G_2, satisfies

$$\sup_{\mathbf{w} \in \mathcal{W}} |\hat{f}_t(\mathbf{w}) - f_t(\mathbf{w})| \leq \delta_t G_2,$$

and is differentiable with the following gradient:

$$\nabla \hat{f}_t(\mathbf{w}) = \mathbb{E}_{\mathbf{u}_t} \left[\frac{d}{\delta_t} f_t(\mathbf{w} + \delta_t \mathbf{u}_t) \mathbf{u}_t \right].$$

Proof The fact that the function is convex and Lipschitz is immediate from its definition and the assumptions in the theorem. The inequality follows from $\mathbf{u}_t$ being a unit vector and that f_t is assumed to be G_2-Lipschitz with respect to the 2-norm. The differentiability property follows from Lemma 2.1 in Flaxman et al. (2005). ■

Lemma 9 *For any function g which is L-Lipschitz with respect to the 2-norm, it holds that if $\mathbf{u}$ is uniformly distributed on the Euclidean unit sphere, then*

$$\sqrt{\mathbb{E}\left[(g(\mathbf{u}) - \mathbb{E}[g(\mathbf{u})])^4 \right]} \leq c \frac{L^2}{d}.$$

for some numerical constant c.

Proof A standard result on the concentration of Lipschitz functions on the Euclidean unit sphere implies that

$$\Pr(|g(\mathbf{u}) - \mathbb{E}[g(\mathbf{u})]| > t) \leq 2 \exp\left(-c' d t^2 / L^2\right)$$

for some numerical constant $c' > 0$ (see the proof of Proposition 2.10 and Corollary 2.6 in Ledoux, 2005). Therefore,

$$\sqrt{\mathbb{E}\left[(g(\mathbf{u})-\mathbb{E}[g(\mathbf{u})])^4\right]} = \sqrt{\int_{t=0}^{\infty}\Pr\left((g(\mathbf{u})-\mathbb{E}[g(\mathbf{u})])^4 > t\right)dt}$$
$$= \sqrt{\int_{t=0}^{\infty}\Pr\left(|g(\mathbf{u})-\mathbb{E}[g(\mathbf{u})]| > \sqrt[4]{t}\right)dt} \leq \sqrt{\int_{t=0}^{\infty} 2\exp\left(-\frac{c'd\sqrt{t}}{L^2}\right)dt} = \sqrt{2\frac{L^4}{(c'd)^2}},$$

where in the last step we used the fact that $\int_{x=0}^{\infty}\exp(-\sqrt{x})dx = 2$. The expression above equals cL^2/d for some numerical constant c. ■

Lemma 10 *It holds that $\mathbb{E}[\tilde{\mathbf{g}}_t|\mathbf{w}_t] = \nabla\hat{f}_t(\mathbf{w}_t)$ (where $\hat{f}_t(\cdot)$ is as defined in Lemma 8), and $\mathbb{E}[\|\tilde{\mathbf{g}}_t\|^2|\mathbf{w}_t] \leq cdp_*^2G_2^2$ for some numerical constant c.*

Proof For simplicity of notation, we drop the t subscript. Since $\mathbf{u}$ has a symmetric distribution around the origin,

$$\begin{aligned}
\mathbb{E}[\tilde{\mathbf{g}}|\mathbf{w}] &= \mathbb{E}_{\mathbf{u}}\left[\frac{d}{2\delta}\left(f(\mathbf{w}+\delta\mathbf{u})-f(\mathbf{w}-\delta\mathbf{u})\right)\mathbf{u}\right]\\
&= \mathbb{E}_{\mathbf{u}}\left[\frac{d}{2\delta}\left(f(\mathbf{w}+\delta\mathbf{u})\right)\mathbf{u}\right] + \mathbb{E}_{\mathbf{u}}\left[\frac{d}{2\delta}f(\mathbf{w}-\delta\mathbf{u})(-\mathbf{u})\right]\\
&= \mathbb{E}_{\mathbf{u}}\left[\frac{d}{2\delta}\left(f(\mathbf{w}+\delta\mathbf{u})\right)\mathbf{u}\right] + \mathbb{E}_{\mathbf{u}}\left[\frac{d}{2\delta}f(\mathbf{w}+\delta\mathbf{u})(\mathbf{u})\right]\\
&= \mathbb{E}_{\mathbf{u}}\left[\frac{d}{\delta}f(\mathbf{w}+\delta\mathbf{u})\mathbf{u}\right]
\end{aligned}$$

which equals $\nabla\hat{f}(\mathbf{w})$ by Lemma 8.

As to the second part of the lemma, we have the following, where α is an arbitrary parameter and where we use the elementary inequality $(a-b)^2 \leq 2(a^2+b^2)$.

$$\begin{aligned}
\mathbb{E}[\|\tilde{\mathbf{g}}\|_*^2|\mathbf{w}] &= \mathbb{E}_{\mathbf{u}}\left[\left\|\frac{d}{2\delta}\left(f(\mathbf{w}+\delta\mathbf{u})-f(\mathbf{w}-\delta\mathbf{u})\right)\mathbf{u}\right\|_*^2\right]\\
&= \frac{d^2}{4\delta^2}\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left(f(\mathbf{w}+\delta\mathbf{u})-f(\mathbf{w}-\delta\mathbf{u})\right)^2\right]\\
&= \frac{d^2}{4\delta^2}\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left((f(\mathbf{w}+\delta\mathbf{u})-\alpha)-(f(\mathbf{w}-\delta\mathbf{u})-\alpha)\right)^2\right]\\
&\leq \frac{d^2}{2\delta^2}\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left((f(\mathbf{w}+\delta\mathbf{u})-\alpha)^2+(f(\mathbf{w}-\delta\mathbf{u})-\alpha)^2\right)\right]\\
&= \frac{d^2}{2\delta^2}\left(\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left(f(\mathbf{w}+\delta\mathbf{u})-\alpha\right)^2\right]+\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left(f(\mathbf{w}-\delta\mathbf{u})-\alpha\right)^2\right]\right).
\end{aligned}$$

Again using the symmetric distribution of $\mathbf{u}$, this equals

$$\frac{d^2}{2\delta^2}\left(\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left(f(\mathbf{w}+\delta\mathbf{u})-\alpha\right)^2\right]+\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left(f(\mathbf{w}+\delta\mathbf{u})-\alpha\right)^2\right]\right)$$
$$=\frac{d^2}{\delta^2}\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^2\left(f(\mathbf{w}+\delta\mathbf{u})-\alpha\right)^2\right].$$

Applying Cauchy-Schwartz and using the condition $\sqrt[4]{\mathbb{E}_{\mathbf{u}}\|\mathbf{u}\|_*^4}\leq p_*$ stated in the theorem, we get the upper bound

$$\frac{d^2}{\delta^2}\sqrt{\mathbb{E}_{\mathbf{u}}\left[\|\mathbf{u}\|_*^4\right]}\sqrt{\mathbb{E}_{\mathbf{u}}\left[\left(f(\mathbf{w}+\delta\mathbf{u})-\alpha\right)^4\right]} \;=\; \frac{p_*^2d^2}{\delta^2}\sqrt{\mathbb{E}_{\mathbf{u}}\left[\left(f(\mathbf{w}+\delta\mathbf{u})-\alpha\right)^4\right]}.$$

In particular, taking $\alpha=\mathbb{E}_{\mathbf{u}}[f(\mathbf{w}+\delta\mathbf{u})]$ and using Lemma 9 (noting that $f(\mathbf{w}+\delta\mathbf{u})$ is $G_2\delta$-Lipschitz w.r.t. $\mathbf{u}$ in terms of the 2-norm), this is at most $\frac{p_*^2d^2}{\delta^2}c\frac{(G_2\delta)^2}{d}=cdp_*^2G_2^2$ as required. ■

We are now ready to prove the theorem. Taking expectations on both sides of the inequality in Lemma 7, we have

$$\mathbb{E}\left[\sum_{t=1}^T\langle\tilde{\mathbf{g}}_t,\mathbf{w}_t-\mathbf{w}^*\rangle\right] \;\leq\; \frac{1}{\eta}R^2+\eta\sum_{t=1}^T\mathbb{E}\left[\|\tilde{\mathbf{g}}_t\|_*^2\right] \;=\; \frac{1}{\eta}R^2+\eta\sum_{t=1}^T\mathbb{E}\left[\mathbb{E}\left[\|\tilde{\mathbf{g}}_t\|_*^2|\mathbf{w}_t\right]\right]. \quad (4)$$

Using Lemma 10, the right hand side is at most

$$\frac{1}{\eta}R^2+\eta cdp_*^2G_2^2T$$

The left hand side of Eq. (4), by Lemma 10 and convexity of $\hat{f}_t$, equals

$$\mathbb{E}\left[\sum_{t=1}^T\langle\mathbb{E}[\tilde{\mathbf{g}}_t|\mathbf{w}_t],\mathbf{w}_t-\mathbf{w}^*\rangle\right]=\mathbb{E}\left[\sum_{t=1}^T\langle\nabla\hat{f}_t(\mathbf{w}_t),\mathbf{w}_t-\mathbf{w}^*\rangle\right]\geq\mathbb{E}\left[\sum_{t=1}^T\left(\hat{f}_t(\mathbf{w}_t)-\hat{f}_t(\mathbf{w}^*)\right)\right].$$

By Lemma 8, this is at least

$$\mathbb{E}\left[\sum_{t=1}^T\left(f_t(\mathbf{w}_t)-f_t(\mathbf{w}^*)\right)\right]-2G_2\sum_{t=1}^T\delta_t.$$

Combining these inequalities and plugging back into Eq. (4), we get

$$\mathbb{E}\left[\sum_{t=1}^T\left(f_t(\mathbf{w}_t)-f_t(\mathbf{w}^*)\right)\right]\leq 2G_2\sum_{t=1}^T\delta_t+\frac{1}{\eta}R^2+cdp_*^2G_2^2\eta T.$$

Choosing $\eta=R/(p_*G_2\sqrt{dT})$, and any $\delta_t\leq p_*R\sqrt{d/T}$, we get

$$\mathbb{E}\left[\sum_{t=1}^T\left(f_t(\mathbf{w}_t)-f_t(\mathbf{w}^*)\right)\right]\leq(c+3)p_*G_2R\sqrt{dT}.$$

Dividing both sides by T, the result follows.

Acknowledgments

This research was supported in part by an Israel Science Foundation grant 425/13 and an FP7 Marie Curie CIG grant. We thank the anonymous reviewers for several helpful comments.

Appendix A. Proof of Lemma 4

We note that the distribution of $\|\mathbf{u}\|_\infty^4$ is identical to that of $\frac{\|\mathbf{n}\|_\infty^4}{\|\mathbf{n}\|_2^4}$, where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, I_d)$ is a standard Gaussian random vector. Moreover, by a standard concentration bound on the norm of Gaussian random vectors (e.g. Corollary 2.3 in Barvinok, 2005, with $\epsilon = 1/2$):

$$\max\left\{\Pr\left(\|\mathbf{n}\|_2 \leq \sqrt{\frac{d}{2}}\right), \Pr\left(\|\mathbf{n}\|_2 \geq \sqrt{2d}\right)\right\} \leq \exp\left(-\frac{d}{16}\right).$$

Finally, for any value of $\mathbf{n}$, we always have $\frac{\|\mathbf{n}\|_\infty}{\|\mathbf{n}\|_2} \leq 1$, since the Euclidean norm is always larger than the infinity norm. Combining these observations, and using $\mathbf{1}_A$ for the indicator function of the event A, we have

$$\begin{aligned}
\mathbb{E}[\|\mathbf{u}\|_\infty^4] &= \mathbb{E}\left[\frac{\|\mathbf{n}\|_\infty^4}{\|\mathbf{n}\|_2^4}\right] \\
&= \Pr\left(\|\mathbf{n}\|_2 \leq \sqrt{\frac{d}{2}}\right) \mathbb{E}\left[\frac{\|\mathbf{n}\|_\infty^4}{\|\mathbf{n}\|_2^4} \,\middle|\, \|\mathbf{n}\|_2 \leq \sqrt{\frac{d}{2}}\right] \\
&\quad + \Pr\left(\|\mathbf{n}\|_2 > \sqrt{\frac{d}{2}}\right) \mathbb{E}\left[\frac{\|\mathbf{n}\|_\infty^4}{\|\mathbf{n}\|_2^4} \,\middle|\, \|\mathbf{n}\|_2 > \sqrt{\frac{d}{2}}\right] \\
&\leq \exp\left(-\frac{d}{16}\right) * 1 + \Pr\left(\|\mathbf{n}\|_2 > \sqrt{\frac{d}{2}}\right) \mathbb{E}\left[\frac{\|\mathbf{n}\|_\infty^4}{\left(\sqrt{d/2}\right)^4} \,\middle|\, \|\mathbf{n}\|_2 > \sqrt{\frac{d}{2}}\right] \\
&= \exp\left(-\frac{d}{16}\right) + \left(\frac{2}{d}\right)^2 \mathbb{E}\left[\|\mathbf{n}\|_\infty^4 \mathbf{1}_{\|\mathbf{n}\|_2 > \sqrt{d/2}}\right] \\
&\leq \exp\left(-\frac{d}{16}\right) + \frac{4}{d^2}\mathbb{E}\left[\|\mathbf{n}\|_\infty^4\right]. \qquad (5)
\end{aligned}$$

Thus, it remains to upper bound $\mathbb{E}\left[\|\mathbf{n}\|_\infty^4\right]$ where $\mathbf{n}$ is a standard Gaussian random variable. Letting $\mathbf{n} = (n_1, \ldots, n_d)$, and noting that $n_1, \ldots, n_d$ are independent and identically distributed standard Gaussian random variables, we have for any scalar $z \geq 1$ that

$$\begin{aligned}\Pr(\|\mathbf{n}\|_\infty \leq z) &= \prod_{i=1}^{n} \Pr(|n_i| \leq z) \;=\; (\Pr(|n_1| \leq z))^d \\ &= (1 - \Pr(|n_1| > z))^d \;\overset{(1)}{\geq}\; 1 - d\Pr(|n_1| > z) \\ &= 1 - 2d\Pr(n_1 > z) \;\overset{(2)}{\geq}\; 1 - d\exp(-z^2/2),\end{aligned}$$

where (1) is Bernoulli's inequality, and (2) is using a standard tail bound for a Gaussian random variable. In particular, the above implies that

$$\Pr\left(\|\mathbf{n}\|_\infty > z\right) \leq d\exp(-z^2/2).$$

Therefore, for an arbitrary positive scalar $r \geq 1$,

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{n}\|_\infty^4\right] &= \int_{z=0}^{\infty} \Pr\left(\|\mathbf{n}\|_\infty^4 > z\right) dz \\ &\leq \int_{z=0}^{r} 1dz + \int_{z=r}^{\infty} \Pr\left(\|\mathbf{n}\|_\infty > \sqrt[4]{z}\right) dz \\ &\leq r + \int_{z=r}^{\infty} d\exp\left(-\frac{\sqrt{z}}{2}\right) dz \\ &= r + 4d(2+\sqrt{r})\exp\left(-\frac{\sqrt{r}}{2}\right).\end{aligned}$$

In particular, plugging $r = 4\log^2(d)$ (which is larger than 1, since we assume $d > 1$), we get $4(2 + 2\log(d) + \log^2(d))$. Plugging this back into Eq. (5), we get that

$$\mathbb{E}[\|\mathbf{u}\|_\infty^4] \;\leq\; \exp\left(-\frac{d}{16}\right) + 16\frac{2 + 2\log(d) + \log^2(d)}{d^2},$$

which can be shown to be at most $c'\left(\frac{\log(d)}{d}\right)^2$ for all $d > 1$, where $c' < 150$ is a numerical constant. In particular, this means that $\sqrt[4]{\mathbb{E}[\|\mathbf{u}\|_\infty^4]} \leq \sqrt[4]{c'}\sqrt{\frac{\log(d)}{d}}$ as required.

References

A. Agarwal, O. Dekel, and L. Xiao. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *Conference on Learning Theory (COLT)*, 2010.

A. Barvinok. Measure concentration lecture notes. `http://www.math.lsa.umich.edu/~barvinok/total710.pdf`, 2005.

N. Cesa-Bianchi, A. Conconi, and C. Gentile. On the generalization ability of on-line learning algorithms. *Information Theory, IEEE Transactions on*, 50(9):2050–2057, 2004.

J. Duchi, M. Jordan, M. Wainwright, and A. Wibisono. Optimal rates for zero-order optimization: the power of two function evaluations. *Information Theory, IEEE Transactions on*, 61(5):2788–2806, May 2015.

A. Flaxman, A. Kalai, and B. McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005.

S. Ghadimi and G. Lan. Stochastic first- and zeroth-order methods for nonconvex stochastic programming. *SIAM Journal on Optimization*, 23(4):2341–2368, 2013.

E. Hazan, A. Agarwal, and S. Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169–192, 2007.

M. Ledoux. *The concentration of measure phenomenon*, volume 89. American Mathematical Soc., 2005.

Y. Nesterov. Random gradient-free minimization of convex functions. Technical Report 2011/16, ECORE, 2011.

S. Shalev-Shwartz. *Online learning: Theory, algorithms, and applications*. PhD thesis, The Hebrew University, 2007.

S. Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2), 2012.