

# Parameter-Free Online Convex Optimization with Sub-Exponential Noise

**Kwang-Sung Jun**

*Boston University, Boston, MA*

KJUN@BU.EDU

**Francesco Orabona**

*Boston University, Boston, MA*

FRANCESCO@ORABONA.COM

**Editors:** Alina Beygelzimer and Daniel Hsu

## Abstract

We consider the problem of unconstrained online convex optimization (OCO) with sub-exponential noise, a strictly more general problem than the standard OCO. In this setting, the learner receives a subgradient of the loss functions corrupted by sub-exponential noise and strives to achieve optimal regret guarantee, without knowledge of the competitor norm, i.e., in a parameter-free way. Recently, [Cutkosky and Boahen \(COLT 2017\)](#) proved that, given unbounded subgradients, it is impossible to guarantee a sublinear regret due to an exponential penalty. This paper shows that it is possible to go around the lower bound by allowing the observed subgradients to be unbounded via stochastic noise. However, the presence of unbounded noise in unconstrained OCO is challenging; existing algorithms do not provide near-optimal regret bounds or fail to have a guarantee. So, we design a novel parameter-free OCO algorithm for Banach space, which we call BANCO, via a reduction to betting on noisy coins. We show that BANCO achieves the optimal regret rate in our problem. Finally, we show the application of our results to obtain a parameter-free locally private stochastic subgradient descent algorithm, and the connection to the law of iterated logarithms.

**Keywords:** Parameter-free, online convex optimization, unconstrained, differentially-private stochastic subgradient descent

## 1. Introduction

In this paper, we are interested in the problem of unconstrained Online Convex Optimization (OCO) with sub-exponential noise. In the standard unconstrained OCO problem, at each round  $t$ , an algorithm chooses an iterate  $\mathbf{w}_t \in \mathbb{R}^d$  and then receives a *negative* subgradient  $\mathbf{g}_t \in -\partial\ell_t(\mathbf{w}_t)$  of a convex loss function  $\ell_t(\mathbf{x})$  given by an adversary.<sup>1</sup> The goal of the learner is to minimize the regret defined by the difference between the cumulative loss of the learner and that of the unknown, arbitrary comparator  $\mathbf{u}$ :

$$\text{Regret}_T(\mathbf{u}) = \sum_{t=1}^T \ell_t(\mathbf{w}_t) - \sum_{t=1}^T \ell_t(\mathbf{u}).$$

Departing from the standard setup, we consider a game where the learner receives a *noisy* version  $\hat{\mathbf{g}}_t$  of  $\mathbf{g}_t$ . Specifically, we assume that the noise  $\hat{\mathbf{g}}_t - \mathbf{g}_t$  is sub-exponential. Note that such a setting nicely mirrors the one of optimization of a fixed convex function with a stochastic first-order oracle.

1. The notation  $\mathbf{g}_t$  is a mnemonic for “gain” since the subgradients correspond to losses in online linear games.

The presence of noise implies that  $\mathbf{w}_t$ , a function of the past noisy subgradients, is also stochastic. Thus, it is natural to minimize the *expected* regret:

$$\mathbb{E}[\text{Regret}_T(\mathbf{u})] = \mathbb{E}\left[\sum_{t=1}^T \ell_t(\mathbf{w}_t) - \sum_{t=1}^T \ell_t(\mathbf{u})\right]. \quad (1)$$

We will define more formally the setting and noise in Section 2. Our goal is to achieve expected regret bounds that have optimal dependency on  $\|\mathbf{u}\|$  and  $T$ , that is the so-called *parameter-free* or *adaptive* OCO algorithms (Foster et al., 2015; Orabona and Pál, 2016; Foster et al., 2017; Cutkosky and Boahen, 2017; Kołowski, 2017; Cutkosky and Orabona, 2018; Foster et al., 2018).

Our problem is motivated by a recent lower bound result on the unconstrained OCO showing that, without prior information on the largest subgradient, parameter-free algorithms are doomed to suffer an exponential penalty  $\exp(\max_t L_t/L_{t-1})$ , where  $L_t$  is dual norm of the largest subgradient up to time  $t$  (Cutkosky and Boahen, 2017). Given such a catastrophic negative result that implies the excessive power of the adversary, one may ask the following question: under what condition on the game can the learner minimize regret efficiently with unbounded subgradients? Our study provides a positive answer by allowing subgradients observed by the learner to be unbounded via *stochasticity*, which limits the adversarial power without restricting observed subgradients to be bounded.

In order to develop low-regret algorithms for noisy OCO, it is tempting to directly use existing algorithms and their guarantees. However, these attempts either result in a suboptimal dependence on  $\|\mathbf{u}\|$  in the regret, namely  $\|\mathbf{u}\|^2$ , or do not lead to nontrivial regret bounds (see Section 3 for details). This motivates the following question: does there exist an unconstrained noisy OCO algorithm whose expected regret scales as optimally with  $\|\mathbf{u}\|$  and  $T$ ? We answer this question in the affirmative by proposing a new Betting Algorithm for Noisy COins (BANCO). BANCO enjoys expected regret

$$O\left(\|\mathbf{u}\|\sqrt{(G^2 + \sigma^2)T \log(1 + \|\mathbf{u}\|T)}\right)$$

in a smooth Banach space, where  $G$  is the bound on the *expected* negative subgradients  $\mathbf{g}_t$  and  $\sigma^2$  is the variance of the noisy negative subgradients  $\hat{\mathbf{g}}_t$ . Our result reveals that, despite the noisy and unbounded nature of the feedback, it is possible to adapt to the unknown and best-in-hindsight comparator just as in the noise-free environments, in expectation.

BANCO is constructed via a natural extension of the coin betting framework (Orabona and Pál, 2016), where we reduce noisy OCO to a 1-d game of betting money on *noisy* coin flips to maximize one's expected wealth. The noisy OCO in Banach space is then reduced to the 1-d coin betting, equipped with any *constrained* noisy OCO learner in a black-box manner. We describe the coin betting view and its extension to Banach OCO in Section 4 and 5 respectively. Furthermore, we further show that the dependence on the variance  $\sigma^2$  cannot be improved, also matching the dependence on  $\|\mathbf{u}\|$  up to logarithmic factors. We stress that, combining our lower bound and the existing ones in the literature, our regret upper bound is unimprovable. We discuss details on lower bounds in Section 6.

Finally, in Section 7, we show some consequences of our results. Indeed, the noisy OCO problem and its algorithms have numerous applications as learning with noisy observations is a dominating paradigm of machine learning. First, we show that our noisy OCO algorithm can be directly used for locally differentially-private stochastic subgradient descent (SGD). In fact, in private SGD noise is added on the subgradients to guarantee privacy, perfectly fitting our framework. In particular, we achieve the first parameter-free locally private SGD algorithm. Second, we show that our algorithmic construction reveals a tight connection to concentration inequalities. Specifically, we show that our

algorithm implies a Banach valued concentration inequality that matches the rate of the law of the iterated logarithm. The connection is made through a simple observation that a noisy coin betting potential directly implies a supermartingale, which is then combined with Doob’s inequality to show concentration inequalities that hold for any time step  $t$ .

We conclude our paper with open problems in Section 8.

## 2. Problem Definition and Preliminaries

In this section, we describe our notations, formally define the problem, and provide background on coin betting.

**Notations.** The dual of a Banach space  $V$  over a field  $F$ , denoted by  $V^*$ , is the set of all continuous linear maps  $V \rightarrow F$ . We use the notation  $\langle \mathbf{v}, \mathbf{w} \rangle$  to indicate the application of a dual vector  $\mathbf{v} \in V^*$  to a vector  $\mathbf{w} \in V$ .  $V^*$  is also a Banach space with the dual norm:  $\|\mathbf{v}\|_* = \sup_{\mathbf{w} \in V, \|\mathbf{w}\| \leq 1} \langle \mathbf{v}, \mathbf{w} \rangle$ . We abbreviate  $\mathbf{x}_1, \dots, \mathbf{x}_t$  by  $\mathbf{x}_{1:t}$ .

**Online convex optimization with noise.** In OCO with noise, as introduced in the introduction, the learner receives a noisy version  $\hat{\mathbf{g}}_t \in V^*$  of the negative subgradient  $\mathbf{g}_t \in V^*$ . Since the learner’s predictions  $\mathbf{w}_t \in V$  are a function of past noisy subgradients, the regret is also stochastic. Therefore, our goal is to minimize the *expected* regret defined in (1).

We assume that the true subgradients are bounded by  $G$ :  $\|\mathbf{g}_t\|_* \leq G$ . Furthermore, the noise  $\boldsymbol{\xi}_t := \hat{\mathbf{g}}_t - \mathbb{E}_t[\hat{\mathbf{g}}_t]$  is conditionally zero-mean and has conditional finite variance measured with the dual norm:

$$\mathbb{E}[\|\boldsymbol{\xi}_t\|_*^2 \mid \boldsymbol{\xi}_{1:t-1}] \leq \sigma^2, \forall t, \quad (2)$$

for some  $\sigma > 0$ . Hereafter, we use the notation  $\mathbb{E}_t$  to denote  $\mathbb{E}[\cdot \mid \boldsymbol{\xi}_{1:t-1}]$ . We also assume a tail condition such that  $\boldsymbol{\xi}_t$  is conditionally sub-exponential with parameters  $(\sigma_{\text{ID}}^2, b)$ :<sup>2</sup>

$$\max_{\mathbf{a}: \|\mathbf{a}\| \leq 1} \mathbb{E}_t[\exp(\beta \langle \boldsymbol{\xi}_t, \mathbf{a} \rangle)] \leq \exp\left(\frac{\beta^2 \sigma_{\text{ID}}^2}{2}\right), \forall |\beta| \leq \frac{1}{b}. \quad (3)$$

One can show that, when (3) is achieved with equality, we have  $\sigma_{\text{ID}}^2 \leq \sigma^2$ . The intuition of the condition above is that the tail of the noise  $\boldsymbol{\xi}_t$  behaves well in any direction; a similar form of condition for sub-Gaussian vectors can be found in Hsu et al. (2012). This noise definition covers a wide range of distributions, including Gaussian and Laplace. Consider the L2 norm for simplicity. If  $d = 1$ , we have  $\sigma^2 = \sigma_{\text{ID}}^2$ . This is not true in general and the relationship depends on the noise distribution and the norm being considered. If  $\boldsymbol{\xi}_t \sim \mathcal{N}(0, s^2 \mathbf{I})$ , then one can see that  $\sigma_{\text{ID}}^2 = s^2$  and  $\sigma^2 = ds^2$ . As another example, the Laplace mechanism noise used in differentially-private learning satisfies the tail condition above; see Section 7.1.

**OCO as betting on noisy coins.** One recent framework for unconstrained OCO is coin betting, which views the OCO game as maximizing a gambler’s wealth via repeated betting on *adversarial* coin flips (McMahan and Abernethy, 2013; Orabona and Pál, 2016). This framework provides a straightforward way to design algorithms that achieve optimal regret bounds with respect to any competitor, without imposing a bounded set for the competitor nor any parameter to tune, i.e., parameter-free. Consider 1d OCO with  $G = 1$  for simplicity. The gambler starts with the initial endowment  $\text{Wealth}_0 = \tau$  for some  $\tau > 0$ . In each iteration  $t$ , the gambler determines how much money to bet and whether to bet on heads (+1) or tails (−1), which is encoded as  $|w_t|$  and  $\text{sign}(w_t)$

2.  $\beta$  is often qualified as  $|\beta| < \frac{1}{b}$  in the literature. Our qualification is merely for ease of exposition.

respectively. After the adversary's (continuous) coin outcome  $g_t \in [-1, 1]$  is revealed, the gambler's wealth, denoted by  $\text{Wealth}_t$ , is updated additively:  $\text{Wealth}_t = \text{Wealth}_{t-1} + g_t w_t$ . That is, the gambler makes (loses) money when she gets the coin side correct (incorrect), and the amount of return (loss) is determined by  $|g_t w_t|$  (respectively). Developing successful strategies critically rely on designing a potential function  $F_t(x)$  and an appropriate betting amount  $w_t$  such that

$$\text{Wealth}_0 = F_0(0) \quad \text{and} \quad F_{t-1}(x) + g_t w_t \geq F_t(x + g_t), \forall t. \quad (4)$$

One can show that the two properties above imply  $\text{Wealth}_t \geq F_t(\sum_{s=1}^t g_s)$  (the derivation is similar to (7) below). McMahan and Orabona (2014, Theorem 1) show that a lower bound on  $\text{Wealth}_T$  is equivalent to an upper bound on the linearized regret w.r.t. a comparator  $u$ ,  $\sum_{t=1}^T g_t(u - w_t)$ , which reveals a tight connection between coin betting and OCO.

In this paper, we extend the coin betting problem to noisy coin outcomes. Specifically, the gambler observes a noisy version of the coin outcome  $\hat{g}_t \in \mathbb{R}$  rather than  $g_t = \mathbb{E}[\hat{g}_t]$ . While the extension appears obvious, the existing coin betting strategies (e.g. Orabona and Pál, 2016; Orabona and Tommasi, 2017) cannot be applied to the noisy setting; their design ensures that the wealth never goes below 0 w.p. 1, which cannot be true for our setting as the coin outcome can be arbitrarily bad.

To cope with noisy coins, we develop a *noisy coin betting* framework. The key idea is that, although we cannot guarantee the nonnegativity of wealth, we can guarantee it for the *expected* wealth. Departing from the conditions for noise-free coin betting (4), we assume that  $F_t$  and  $w_t$  satisfy the betting relationship in conditional expectation

$$F_{t-1}(x) + w_t g_t \geq \mathbb{E}_t[F_t(x + \hat{g}_t)]. \quad (5)$$

This immediately implies that

$$\mathbb{E} \text{Wealth}_t \geq \mathbb{E} \left[ F_t \left( \sum_{s=1}^t \hat{g}_s \right) \right]. \quad (6)$$

In fact, by induction, assume that (6) holds for  $t - 1$ . Then,

$$\begin{aligned} \mathbb{E} \text{Wealth}_t &= \mathbb{E}[\text{Wealth}_{t-1} + \mathbb{E}_t \hat{g}_t w_t] \stackrel{(a)}{\geq} \mathbb{E} \left[ F_{t-1} \left( \sum_{s=1}^{t-1} \hat{g}_s \right) + \mathbb{E}_t \hat{g}_t w_t \right] \\ &= \mathbb{E} \left[ F_{t-1} \left( \sum_{s=1}^{t-1} \hat{g}_s \right) + g_t w_t \right] \stackrel{(b)}{\geq} \mathbb{E} \left[ F_t \left( \sum_{s=1}^t \hat{g}_s \right) \right], \end{aligned} \quad (7)$$

where (a) is by the inductive hypothesis and (b) is by (5).

### 3. The Devil is in the Details: Failing Approaches

As a warm-up, we discuss how one might attempt to extend existing algorithms for the noisy setting and why these approaches would fail. For simplicity, consider that  $V = \mathbb{R}^d$ , the norm is the L2 norm, and  $G = 1$ . For this, we need algorithms that enjoy regret bounds *without requiring a subgradient bound as an input*. For example, one can apply online subgradient descent (OGD), which guarantees a regret bound w.r.t. the noisy subgradients:

$$\hat{R}_T^{\text{Lin}}(\mathbf{u}) := \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} - \mathbf{w}_t \rangle = \frac{\|\mathbf{u}\|^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^T \|\hat{\mathbf{g}}_t\|^2.$$

Notice that  $\hat{R}_T^{\text{Lin}}(\mathbf{u})$  itself does not bound  $\text{Regret}_T(\mathbf{u})$  and one must turn to either *expected* or *high probability* regret bounds. With the choice of the step size  $\eta = 1/\sqrt{(\sigma^2 + 1)T}$ , we have an expected regret bound:

$$\mathbb{E}[\text{Regret}_T(\mathbf{u})] \stackrel{(a)}{\leq} \mathbb{E}\left[\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u} - \mathbf{w}_t \rangle\right] \stackrel{(b)}{=} \mathbb{E}\left[\sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} - \mathbf{w}_t \rangle\right] = O\left(\left(\|\mathbf{u}\|^2 + 1\right)\sqrt{(\sigma^2 + 1)T}\right),$$

where (a) is by convexity and (b) is by the tower rule. However, the dependence on the unknown comparator  $\mathbf{u}$  is  $\|\mathbf{u}\|^2$ , which is much larger than the best known rate, which is  $\|\mathbf{u}\|\sqrt{\log(1 + \|\mathbf{u}\|)}$  (McMahan and Orabona, 2014). While there exist algorithms that almost achieve this rate w.r.t.  $\|\mathbf{u}\|$  without requiring a bound on  $\hat{\mathbf{g}}_{1:T}$  as input (e.g., Cutkosky and Boahen (2017) with  $\gamma \approx \frac{1}{2}$ ), the lower bound of Cutkosky and Boahen (2017) implies that the overall regret bound cannot be sublinear.

Another attempt is to leverage the fact that the noisy subgradients are bounded with high probability. Consider for example a 1d OCO problem with  $(\sigma_{\text{ID}}^2, 0)$ -sub-exponential noise in which case  $\sigma = \sigma_{\text{ID}}$ . Let  $E_1$  be the event that  $|\hat{g}_t| \leq g_t + \sigma\sqrt{\log(T/\delta)}$  for all  $t \leq T$  (omitting constants), which satisfies  $\mathbb{P}(-E_1) \leq \delta$ . Using the standard parameter-free OCO algorithms such as the one in McMahan and Orabona (2014), one may obtain the following bound under the event  $E_1$ :

$$\hat{R}_T^{\text{Lin}}(u) = O\left(|u| \left(G + \sigma\sqrt{\log(T/\delta)}\right) \sqrt{T \log(1 + |u|)}\right),$$

which is, again, *not* an upper bound on  $\text{Regret}_T(u)$ , not even under  $E_1$ .<sup>3</sup> Define the linearized regret:  $R_T^{\text{Lin}}(u) = \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u} - \mathbf{w}_t \rangle$ . In a special case where there exists  $c > 0$  such that  $\hat{R}_T^{\text{Lin}}(u) \leq c|u|T$  (though we explain below this is unrealistic), one may have an expected regret bound as follows:

$$\begin{aligned} \mathbb{E} \text{Regret}_T(u) &\leq \mathbb{E}[R_T^{\text{Lin}}(u)] = \mathbb{E}[\hat{R}_T^{\text{Lin}}(u)] \\ &= \mathbb{E}[\hat{R}_T^{\text{Lin}}(u)|E_1] \cdot \mathbb{P}(E_1) + \mathbb{E}[\hat{R}_T^{\text{Lin}}(u)|-E_1] \cdot \mathbb{P}(-E_1) \\ &= O\left(|u| \left(G + \sigma\sqrt{\log(T/\delta)}\right) \sqrt{T \log(1 + |u|)}\right) + c|u|T\delta. \end{aligned}$$

Indeed, the assumption  $\hat{R}_T^{\text{Lin}}(u) \leq c|u|T$  would be true for *constrained* OCO with *bounded* noise  $\xi_t$ . However, our case is neither constrained nor with bounded noise. For a fixed  $u$ , if  $u - w_T > 0$ , then  $\hat{g}_T$  can be arbitrarily large, making the regret much larger than  $c|u|T$  for any  $c$ . Such an issue caused by unbounded noise poses a significant challenge in designing unconstrained algorithms adapting to the unknown comparator  $\mathbf{u}$  under noisy feedback.

Finally, we remark that, for linear losses, the standard OGD can have an expected regret that does not scale with  $\sigma$ . This, however, does not generalize to generic convex losses. In fact, our lower bound result in Section 6 shows that the factor  $\sigma$  in the expected regret bound cannot be avoided in general. We elaborate more on this in Appendix A.

#### 4. One-dimensional Betting Algorithm with Noisy Coins

In this section, we show how to construct noisy coin betting potentials. We focus on potential functions  $F_t$  and associated betting strategy  $w_t$  defined as follows:

$$F_t(x) = \int f_t(x, \beta) d\pi(\beta), \quad \text{and} \quad w_t = \int \beta f_t\left(\sum_{i=1}^{t-1} \hat{g}_i, \beta\right) d\pi(\beta),$$

for some functions  $f_t(x, \beta)$ , and a prior  $\pi(\beta)$ . This defines a family of noisy coin betting potentials, parameterized by the prior  $\pi$ . While this kind of potentials have been used by Chernov and Vovk (2010); Koolen and van Erven (2015) for parameter-free algorithms for learning with expert advice,

3. One may attempt to derive a high probability regret bound via a decomposition  $\sum_{t=1}^T \langle g_t, u - w_t \rangle = \sum_{t=1}^T \langle g_t - \hat{g}_t, u - w_t \rangle + \sum_{t=1}^T \langle \hat{g}_t, u - w_t \rangle$ . However, the first summation involves  $w_t$  that is unbounded, and analyzing the behavior of  $w_t$  appears nontrivial. We leave high probability regret bounds as future work.

---

**Algorithm 1** Betting Algorithm for Noisy COins (BANCO)
 

---

**Require:** sub-exponential parameters  $(\sigma^2, b)$ , expected subgradient bound  $G$ , initial money  $\tau$ .  
**for**  $t = 1$  **to**  $T$  **do**  
     Play  $w_t = \tau \int_{-a}^a \beta \exp\left(\beta \sum_{s=1}^{t-1} \hat{g}_s - \beta^2(t-1)\left(\frac{\sigma^2}{2} + G^2\right)\right) d\pi(\beta)$  where  $a = \min\left(\frac{k_1}{G}, \frac{1}{b}\right)$ .  
     Receive  $\hat{g}_t \in \mathbb{R}$ .  
**end for**

---

our key novelty lies in blending the effect of sub-exponential noise into the potential naturally, making it amenable to analysis.

Our construction is based on the following key inequality for sub-exponential random variables.

**Lemma 1** *Let  $\hat{g}$  be a  $(\sigma^2, b)$ -sub-exponential random variable, with mean  $g$  such that  $|g| \leq G$ . Let  $k_1$  satisfy*

$$1 - k_1 = \exp(-k_1 - k_1^2), \quad (8)$$

that is  $k_1 = 0.683803\dots$ . Then, for any  $\beta$  such that  $|\beta| \leq \min(k_1/G, 1/b)$ , we have

$$1 + \beta \mathbb{E}_{\hat{g}}[\hat{g}] \geq \mathbb{E}_{\hat{g}} \exp\left(\beta \hat{g} - \beta^2 \left(\frac{\sigma^2}{2} + G^2\right)\right). \quad (9)$$

**Proof** Given that  $|\beta| \leq \min(k_1/G, 1/b)$ , we have  $\beta g \geq -k_1$  and  $1 + \beta g \geq e^{\beta g - \beta^2 g^2}$ . Then,

$$\begin{aligned} 1 + \beta \mathbb{E}_{\hat{g}}[\hat{g}] &= 1 + \beta g \geq \exp(\beta g - \beta^2 g^2) \geq \mathbb{E}_{\hat{g}} \exp\left(\beta \hat{g} - \beta^2 \left(\frac{\sigma^2}{2} + g^2\right)\right) \\ &\geq \mathbb{E}_{\hat{g}} \exp\left(\beta \hat{g} - \beta^2 \left(\frac{\sigma^2}{2} + G^2\right)\right), \end{aligned}$$

where the second inequality is due to  $\mathbb{E} \exp(\beta(\hat{g} - g)) \leq \exp(\beta^2 \sigma^2 / 2)$  for all  $|\beta| \leq \frac{1}{b}$ . ■

From this lemma, multiplying the right hand side of the equation for  $i = 1$  to  $t$ , it is natural to define our noisy coin betting potential as

$$F_t(x) := \tau \int_{-a}^a \exp\left(\beta x - \beta^2 t \left(\frac{\sigma^2}{2} + G^2\right)\right) d\pi(\beta), \quad (10)$$

and associated prediction strategy

$$w_t = \tau \int_{-a}^a \beta \exp\left(\beta \sum_{s=1}^{t-1} \hat{g}_s - \beta^2(t-1)\left(\frac{\sigma^2}{2} + G^2\right)\right) d\pi(\beta),$$

where  $a \leq \min(k_1/G, 1/b)$  and  $\pi$  has support in  $[-a, a]$ . In this way, we obtain our Betting Algorithm for Noisy COins (BANCO) and summarize it in Algorithm 1. In the following theorem we show that (10) satisfy our assumptions.

**Theorem 2** *Let  $w_t$  be computed by Algorithm 1. Then  $F_t$  in (10) is a noisy betting potential.*

**Proof** From the definition it is obvious that  $F_0(0) = \tau$ . We then have to show that  $\mathbb{E}_t F_t(x + \hat{g}_t) \leq F_{t-1}(x) + g_t w_t$ . Hence, consider

$$\mathbb{E}_t F_t(x + g_t) = \tau \mathbb{E}_t \int_{-a}^a \exp\left(\beta(x + \hat{g}_t) - \beta^2 t \left(\frac{\sigma^2}{2} + G^2\right)\right) d\pi(\beta)$$

$$\begin{aligned}
 &= \tau \mathbb{E}_t \int_{-a}^a \exp(\beta \hat{g}_t - \beta^2(\sigma^2/2 + G^2)) \exp(\beta x - \beta^2(t-1)(\sigma^2/2 + G^2)) d\pi(\beta) \\
 &\stackrel{(a)}{\leq} \tau \int_{-a}^a (1 + \beta \mathbb{E}_t \hat{g}_t) \exp(\beta x - \beta^2(t-1)(\sigma^2/2 + G^2)) d\pi(\beta) \stackrel{(b)}{=} F_{t-1}(x) + g_t w_t,
 \end{aligned}$$

where (a) is due to (9) and (b) is by Fubini's theorem.  $\blacksquare$

In the standard coin betting, a lower bound on the wealth is equivalent to an upper bound on the regret for linearized losses by (McMahan and Orabona, 2014, Theorem 1). We extend this result to the *expected* wealth and linearized regret, proof in Appendix B.

**Theorem 3 (Reward-Regret relationship)** *Let  $V, V^*$  be a pair of dual vector spaces. Let  $F : V^* \rightarrow \mathbb{R} \cup \{+\infty\}$  be a proper convex lower semi-continuous function and let  $F^* : V \rightarrow \mathbb{R} \cup \{+\infty\}$  be its Fenchel conjugate. Let  $w_1, w_2, \dots, w_T \in V$  and  $\tau \in \mathbb{R}$ . Then,*

$$\underbrace{\tau + \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{g}_t, w_t \rangle \right]}_{\mathbb{E}[\text{Wealth}_T]} \geq \mathbb{E} \left[ F \left( \sum_{t=1}^T \hat{g}_t \right) \right] \quad \text{iff} \quad \forall u \in V, \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \hat{g}_t, u - w_t \rangle \right]}_{\mathbb{E}[\hat{R}_T^{\text{lin}}(u)]} \leq F^*(u) + \tau.$$

Hence, to obtain a regret bound from the above theorem, we just need to compute the Fenchel conjugate of the noisy coin betting potential  $F_T$ .

To construct a specific algorithm, it remains to choose the prior  $\pi$ . While one can choose any prior, it is preferred to have a closed form expression for  $w_t$ . We choose  $\text{Uniform}[-a, a]$  for simplicity, which results in

$$w_t = \tau \frac{\sqrt{\pi} \exp\left(\frac{(\sum_{s=1}^{t-1} \hat{g}_s)^2}{4t(\sigma^2/2 + G^2)}\right) \left[ \text{erf}\left(\frac{2at(\sigma^2/2 + G^2) - \sum_{s=1}^{t-1} \hat{g}_s}{2\sqrt{t(\sigma^2/2 + G^2)}}\right) + \text{erf}\left(\frac{2at(\sigma^2/2 + G^2) + \sum_{s=1}^{t-1} \hat{g}_s}{2\sqrt{t(\sigma^2/2 + G^2)}}\right) \right]}{4a\sqrt{t(\sigma^2/2 + G^2)}}.$$

Note that a similar prediction strategy was also proposed in (Koolen and van Erven, 2015). It is easy to verify that another choice that results in a closed form update with an equivalent wealth guarantee is with a Gaussian prior centered at zero. For improving numerical precision for computing  $w_t$  above, we refer to (Koolen, 2015).

In the following theorem we calculate the Fenchel conjugate of this potential function from which the regret bound immediately follows by Theorem 3, proof in Appendix C.

**Theorem 4** *Let  $F(x) = \tau \int_{-a}^a \exp(\beta x - \beta^2 S) d\pi(\beta)$  where  $\pi(\beta)$  is  $\text{Uniform}[-a, a]$ . Then,*

$$F^*(u) \leq \max \left\{ |u| \sqrt{2S \ln \left( 1 + \frac{16ea^2 S^2 u^2}{\tau^2} \right)}, \frac{8}{3a} |u| \ln \left( \frac{32}{3ea\tau} |u| \right) \right\}.$$

Applying the two theorems above with  $S = T(\sigma^2/2 + G^2)$  and  $a = \min(k_1/G, 1/b)$ , where  $k_1$  is defined in (8), we have the expected regret guarantee of BANCO:

$$\begin{aligned}
 \mathbb{E}[\text{Regret}_T(u)] \leq \tau + |u| \max \left\{ \sqrt{2 \left( G^2 + \frac{\sigma^2}{2} \right) T \ln \left( 1 + 16e \min \left( \frac{k_1}{G}, \frac{1}{b} \right)^2 T^2 \left( G^2 + \frac{\sigma^2}{2} \right)^2 \frac{u^2}{\tau^2} \right)}, \right. \\
 \left. \frac{8}{3} \max \left( \frac{G}{k_1}, b \right) \ln \left( 32 \max \left( \frac{G}{k_1}, b \right) \frac{|u|}{3e\tau} \right) \right\}.
 \end{aligned}$$

---

**Algorithm 2** BANCO in Banach Spaces
 

---

**Require:** Banach space  $V$ , learner  $\mathcal{A}_D$  with domain being the unit ball  $S \subset V$   
**for**  $t = 1$  **to**  $T$  **do**  
     Get point  $w_t \in \mathbb{R}$  from BANCO, Algorithm 1  
     Get point  $\mathbf{y}_t \in S$  from  $\mathcal{A}_D$   
     Play  $\mathbf{x}_t = w_t \mathbf{y}_t \in V$   
     Receive a noisy negative subgradient  $\hat{\mathbf{g}}_t$  such that  $\mathbb{E}[\hat{\mathbf{g}}_t] \in -\partial \ell_t(\mathbf{x}_t)$   
     Set  $s_t = \langle \hat{\mathbf{g}}_t, \mathbf{y}_t \rangle$   
     Send  $s_t$  to BANCO, Algorithm 1  
     Send  $\hat{\mathbf{g}}_t$  as the  $t$ -th negative subgradient to  $\mathcal{A}_D$   
**end for**

---

## 5. Banach Online Convex Optimization with Noise

In this section, we extend the parameter-free algorithm, BANCO, to Banach spaces. Attempting to extend the 1d algorithm to higher dimensional spaces would require an ad hoc analysis specialized to the particular algorithm. Instead, we leverage a black-box reduction: we take any *constrained* noisy OCO algorithm for Banach space and turn it into an *unconstrained* one via BANCO.

Let  $V$  be a Banach space and the negative subgradients  $\hat{\mathbf{g}}_t \in V^*$  satisfy  $\|\mathbb{E}_t \hat{\mathbf{g}}_t\|_* \leq G$ . Define  $S$  to be the unit ball in  $V$ . We summarize our reduction in Algorithm 2, which is a direct extension of Cutkosky and Orabona (2018) for noisy subgradients. The key feature of the algorithm is a black-box reduction that takes two learners: (i) the 1d coin-betting that predicts the *magnitude*  $w_t \in \mathbb{R}$  and (ii) a  $d$ -dimensional learner  $\mathcal{A}_D$  that predicts the *direction*  $\mathbf{y}_t \in S$ . The reduction then makes the combined prediction by  $\mathbf{x}_t = w_t \mathbf{y}_t$ . After receiving the noisy negative subgradient  $\hat{\mathbf{g}}_t$  evaluated at  $\mathbf{x}_t$ , we feed  $s_t = \langle \hat{\mathbf{g}}_t, \mathbf{y}_t \rangle$  into the coin-betting algorithm and  $\hat{\mathbf{g}}_t$  into  $\mathcal{A}_D$  as the subgradient.

Theorem 5 below shows that the expected regret of Algorithm 2 is nicely decomposed into two expected regrets, each from the noisy coin betting algorithm and  $\mathcal{A}_D$ . The fact that we require the expected regret of  $\mathcal{A}_D$  w.r.t. the *unit norm comparator* frees us from tuning the parameter of  $\mathcal{A}_D$  for the optimal step size, delegating the burden of adaptation to the noisy coin betting algorithm. The proof is simple and immediate from Cutkosky and Orabona (2018), but for completeness we report it in Appendix D.

**Theorem 5** Suppose  $\mathcal{A}_D$  obtains expected regret  $R_T^D(\mathbf{u}) := \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} - \mathbf{y}_t \rangle$  for any competitor  $\mathbf{u}$  in the unit ball  $S \subset V$  and the coin betting algorithm obtains expected regret  $R_T^M(v) := \sum_{t=1}^T s_t \cdot (v - w_t)$  for any competitor  $v \in \mathbb{R}$ . Then, Algorithm 2 guarantees

$$\mathbb{E} \text{Regret}_T(\mathbf{u}) \leq R_T^M(\|\mathbf{u}\|) + \|\mathbf{u}\| R_T^D(\mathbf{u}/\|\mathbf{u}\|),$$

where we define  $\mathbf{u}/\|\mathbf{u}\| = \mathbf{0}$  when  $\mathbf{u} = \mathbf{0}$ .

Note that the loss  $\langle \hat{\mathbf{g}}_t, \mathbf{y}_t \rangle = \langle \mathbf{g}_t, \mathbf{y}_t \rangle + \langle \boldsymbol{\xi}_t, \mathbf{y}_t \rangle$  fits the 1d noisy OCO setting exactly. To see this,  $|\langle \mathbf{y}_t, \mathbf{g}_t \rangle| \leq \|\mathbf{g}_t\|_* \leq G$ . Furthermore, the random variable  $\langle \mathbf{y}_t, \boldsymbol{\xi}_t \rangle \mid \boldsymbol{\xi}_{1:t-1}$  is  $(\sigma_{\text{ID}}^2, b)$ -sub-exponential since

$$\forall |\nu| \leq 1/b, \quad \mathbb{E}_t \exp(\nu \langle \boldsymbol{\xi}_t, \mathbf{y}_t \rangle) \leq \exp(\nu^2 \sigma_{\text{ID}}^2 / 2),$$

where we use the fact  $\|\mathbf{y}_t\| \leq 1$  and our noise assumption (3).

For  $\mathcal{A}_D$ , one can invoke any algorithm for the Banach space of interest (Srebro et al., 2011). In particular, if  $V$  is  $(2, \lambda)$ -uniformly convex (Pinelis, 2015), we can use online mirror descent with



stepsizes  $\eta_t = \frac{\sqrt{\lambda}}{\sqrt{\sum_{s=1}^t \|\hat{\mathbf{g}}_s\|_*^2}}$  and predictions projected onto the unit ball  $S$ . One can then immediately obtain the expected regret bound with noisy subgradients:

$$\begin{aligned} R_T^D\left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\right) &\leq \mathbb{E}\left[\sum_{t=1}^T \left\langle \mathbf{g}_t, \frac{\mathbf{u}}{\|\mathbf{u}\|} - \mathbf{y}_t \right\rangle\right] = \mathbb{E}\left[\sum_{t=1}^T \left\langle \hat{\mathbf{g}}_t, \frac{\mathbf{u}}{\|\mathbf{u}\|} - \mathbf{y}_t \right\rangle\right] = O\left(\mathbb{E}\left[\frac{1}{\sqrt{\lambda}} \sqrt{\sum_{t=1}^T \|\hat{\mathbf{g}}_t\|_*^2}\right]\right) \\ &\stackrel{(a)}{\leq} O\left(\frac{1}{\sqrt{\lambda}} \sqrt{\sum_{t=1}^T (\mathbb{E}\|\mathbf{g}_t\|_*^2 + \sigma^2)}\right), \end{aligned}$$

where (a) uses Jensen's inequality and the fact that  $\mathbb{E}[\|\hat{\mathbf{g}}_t\|_*^2] \leq 2\mathbb{E}[\|\mathbf{g}_t\|_*^2] + 2\sigma^2$ .

Finally, Algorithm 2 equipped the uniform prior in the noisy coin betting algorithm and  $\mathcal{A}_D$  chosen as above enjoys the following expected regret bound:

$$\begin{aligned} \mathbb{E} \text{Regret}_T(\mathbf{u}) &= O\left(\|\mathbf{u}\| \max\left\{(G+b) \ln \frac{\|\mathbf{u}\|(G+b)}{\tau}, \sqrt{(G^2 + \sigma_{\text{ID}}^2)T \ln\left(\frac{\|\mathbf{u}\|(G^2 + \sigma_{\text{ID}}^2)T}{\tau} + 1\right)}\right\}\right. \\ &\quad \left. + \frac{\|\mathbf{u}\|}{\sqrt{\lambda}} \sqrt{\sum_{t=1}^T (\mathbb{E}\|\mathbf{g}_t\|_*^2 + \sigma^2) + \tau}\right). \end{aligned}$$

Examples of  $(2, \lambda)$ -uniformly convex Banach space include Hilbert spaces with 2-norm (in which case  $\lambda = 1$ ), as well as with  $p$ -norm with  $p \in (1, 2]$  (in which case  $\lambda = p - 1$ ). The runtime of Algorithm 2 is dominated by the direction learner  $\mathcal{A}_D$  since the runtime of BANCO does not scale with  $d$ . In other words, the black-box reduction adds little computational overhead while adapting to the unknown best-in-hindsight comparator from noisy feedback.

## 6. Lower bound

In this section, we investigate lower bounds on the noisy OCO problem. Theorem 6 shows that our dependence on the noise variance  $\sigma^2$  is unimprovable in general.

**Theorem 6** *Let  $\sigma \geq 2$ ,  $p \geq 1$ . Let  $q$  satisfy  $1/q = 1 - 1/p$ . Denote by  $\nabla \hat{\ell}_t(\mathbf{x})$  a noisy subgradient of  $\ell_t(\mathbf{x})$ . For any algorithm, there exists a noisy OCO instance with 1-Lipschitz loss functions w.r.t.  $p$ -norm and  $\mathbb{E}\|\nabla \hat{\ell}_t(\mathbf{x}) - \nabla \ell_t(\mathbf{x})\|_q^2 \leq \sigma^2$  and a comparator  $u$  s.t.*

$$\begin{aligned} p \geq 2 &\implies \mathbb{E} \text{Regret}_T(u) \geq \min\left\{c_0 \sigma \|\mathbf{u}\|_p d^{\frac{1}{2} - \frac{1}{p}} \sqrt{T}, \frac{1}{18} \|\mathbf{u}\|_p d^{-\frac{1}{p}} T\right\} \quad \text{and} \\ p \in [1, 2] &\implies \mathbb{E} \text{Regret}_T(u) \geq \min\left\{c_0 \sigma \|\mathbf{u}\|_p \sqrt{T}, \frac{1}{18} \|\mathbf{u}\|_p T\right\}, \end{aligned}$$

where  $c_0$  is a universal constant.

The main argument of the proof is based on a carefully constructed stochastic optimization instance, which is connected to online convex optimization through the online-to-batch conversion (Littlestone, 1989); see Appendix E for details.

Note that our lower bound's dependence on  $\|\mathbf{u}\|$  mismatches our upper bound by a factor of  $\sqrt{\log(1 + \|\mathbf{u}\|)}$ . The reason is that the constructed problem class for the proof is an easier optimization problem where the learner knows the norm of the best competitor  $\mathbf{u}$ . One may attempt to extend the lower bound of Orabona (2013) to the noisy setting, which has the right dependence

on  $\|\mathbf{u}\|$ . However, their construction is based on linear losses in which there exists a learner whose expected regret does not scale with  $\sigma$ , as we show in Appendix A.

Nevertheless, we claim that the expected regret of the noisy OCO is

$$\Omega((G + \sigma)\|\mathbf{u}\|\sqrt{T \log(1 + \|\mathbf{u}\|)}).$$

This claim is based on the lower bound  $\Omega(G\|\mathbf{u}\|\sqrt{T \log(1 + \|\mathbf{u}\|)})$  for noise-free unconstrained OCO (Orabona, 2013, Theorem 2). Specifically, suppose there exists an algorithm  $\mathcal{A}$  achieving a strictly better order of regret bound than  $G\|\mathbf{u}\|\sqrt{T \log(1 + \|\mathbf{u}\|)}$  in the noisy setting. We can then solve the standard noise-free problem by adding some infinitesimal noise to the observed (non-noisy) gradients by ourselves and feeding that noisy gradients to  $\mathcal{A}$ . This leads to a better regret bound than the lower bound for the noise-free problem, which is a contradiction.

## 7. Applications

We discuss two applications of our results to domains beyond the one of online learning.

### 7.1. Parameter-Free Locally Differentially Private SGD

In this section, we describe the application of our algorithm to the locally differentially private SGD (Duchi et al., 2014; Song et al., 2015). An  $\epsilon$ -differentially private algorithm must guarantee that the log-likelihood ratio of the outputs of the algorithm under two databases differing in a single individual's data is smaller than  $\epsilon$  (Dwork et al., 2006). In the stricter definition of local differential privacy (Wasserman and Zhou, 2010; Kasiviswanathan et al., 2011; Duchi et al., 2014; Song et al., 2015) instead an untrusted algorithm is allowed to access a perturbed version of a sensitive dataset only through a sanitization interface. In particular, the sanitization mechanism must guarantee that the log-likelihood ratio of the data of two individuals  $i$  and  $j$  is smaller than  $\epsilon$ .

**Definition 1 (Local Differential Privacy (Duchi et al., 2014; Song et al., 2015))** *Let  $D = (X_1, \dots, X_n)$  be a sensitive dataset where each  $X_i$  corresponds to data about individual  $i$ . A randomized sanitization mechanism  $M$  which outputs a disguised version  $S = (U_1, \dots, U_n)$  of  $D$  is said to provide  $\epsilon$ -local differential privacy to individual  $i$ , if*

$$\sup_S \sup_{x, x' \in D} \frac{\mathbb{P}[U_i \in S | X_i = x]}{\mathbb{P}[U_i \in S | X_i = x']} \leq \exp(\epsilon),$$

where the randomization is taken over the randomization in the sanitization mechanism.

The local differential setting can be specialized to SGD (Song et al., 2015). Consider the minimization of function  $H(\mathbf{w}) = \mathbb{E}_{\mathbf{x} \sim \rho_X} [h(\mathbf{w}, \mathbf{x})]$ , where  $h(\mathbf{w}, \mathbf{x})$  is convex in the first argument and  $\mathbf{x}$  represents sensitive data about one individual. The sanitization mechanism becomes the noisy subgradient oracle that returns  $\mathcal{G}(\mathbf{w}) \in \partial h(\mathbf{w}, \mathbf{x}) + \boldsymbol{\xi}_t$  when queried on  $\mathbf{w}$ , where  $\mathbf{x}$  is coming i.i.d. from  $\rho_X$  and the noise  $\boldsymbol{\xi}_t$  guarantees the local differential privacy (Song et al., 2015).

We now apply the results from Section 5, to show a parameter-free locally differential private SGD algorithm. Consider the Laplace sanitization mechanism that adds noise with probability density function  $\rho_\xi(\mathbf{z}) \propto \exp(-\frac{\epsilon}{2}\|\mathbf{z}\|_2)$ . In words, the noise added to the subgradients makes them very similar to one another. Song et al. (2015) proved that this mechanism is  $\epsilon$ -local differentially private. Also, the noise is zero-mean and they proved that  $\mathbb{E}[\|\boldsymbol{\xi}_t\|_2^2] \leq \frac{4(d^2+d)}{\epsilon^2}$ , satisfying (2). We now prove that the Laplace mechanism also satisfies the sub-exponential noise assumption (3). The proof is rather technical, hence we defer it to Appendix G.

**Lemma 7** Let  $\xi \in \mathbb{R}^d$  a random variable drawn from the density  $\rho_\xi(\mathbf{z}) \propto \exp(-\frac{\epsilon}{2}\|\mathbf{z}\|_2)$ . Then

$$\max_{\|\mathbf{a}\| \leq 1} \mathbb{E}_t [\exp(\beta \langle \xi_t, \mathbf{a} \rangle)] \leq \exp\left(\frac{9d^2\beta^2}{\epsilon^2}\right), \quad \forall |\beta| \leq \frac{\epsilon}{4}.$$

Theorem 5 in conjunction with the online-to-batch conversion (Littlestone, 1989) directly implies the convergence guarantee of a differentially private version of BANCO as stated in the following corollary.

**Corollary 1** Assume  $h(\mathbf{w}, \mathbf{x})$  convex in the first argument  $\mathbf{w} \in \mathbb{R}^d$  and with its subgradients have L2 norm bounded by 1, where the subgradient is with respect to the first argument. Set the uniform prior in BANCO, Algorithm 1, and  $\mathcal{A}_D$  being projected OGD with stepsizes  $\eta_t = 1/\sqrt{\sum_{s=1}^t \|\hat{\mathbf{g}}_s\|_2^2}$  in Algorithm 2 for  $T$  iterations on the sequence of losses  $\ell_t(\mathbf{w}) = h(\mathbf{w}, \mathbf{x}_t)$ , where  $\mathbf{x}_t$  are coming i.i.d. from a distribution  $\rho_X$ . Set  $\hat{\mathbf{g}}_t = \mathbf{g}_t + \xi_t$ , where  $\xi_t \in \mathbb{R}^d$  is drawn from the density  $\rho_\xi(\mathbf{z}) \propto \exp(-\frac{\epsilon}{2}\|\mathbf{z}\|_2)$ . Then, for any  $\mathbf{w}^* \in \mathbb{R}^d$ , we have

$$\mathbb{E} \left[ H \left( \frac{1}{T} \sum_{t=1}^T \mathbf{w}_t \right) \right] - H(\mathbf{w}^*) \leq O \left( \frac{d\|\mathbf{w}^*\|_2}{\epsilon\sqrt{T}} \sqrt{\ln \left( 1 + \frac{d^2\|\mathbf{w}^*\|_2 T}{\epsilon^2\tau} \right)} + \frac{\tau}{T} \right).$$

This convergence rate matches the one for locally private SGD in Wu et al. (2017) up to polylogarithmic terms, with the important difference that we do not need to assume the knowledge of the norm of the optimal solution  $\mathbf{w}^*$  to tune the stepsizes.

## 7.2. Noisy Coin Betting Implies the Law of Iterated Logarithms in Banach Spaces

There is tight connection between concentration inequalities in Banach spaces and online linear optimization algorithms unveiled by Rakhlin and Sridharan (2017). They showed that online mirror descent with adaptive stepsizes gives rise to self-normalized concentration inequality for martingales. Hence, it is natural to ask what kind of concentration can be derived from the noisy coin betting algorithms. Here, we show that there is a connection between the law of iterated logarithms for sub-Gaussian RVs in Banach spaces and Algorithm 2. The exact same reasoning holds for sub-exponential RVs, but we consider the sub-Gaussian case for ease of exposition.

First, consider the one-dimensional case. It is immediate to see that, setting  $g_t = 0$ , we have that (5) implies that  $\mathbb{E}_t[F_t(\sum_{i=1}^t \hat{g}_i)] \leq F_{t-1}(\sum_{i=1}^{t-1} \hat{g}_i)$ , that is  $F_t(\sum_{i=1}^t \hat{g}_i)$  is a supermartingale. Hence, we can use Doob's inequality (Durrett, 2010, Exercise 5.7.1) to have

$$\mathbb{P} \left[ \max_t F_t \left( \sum_{i=1}^t \hat{g}_i \right) \geq \frac{1}{\delta} \right] \leq \delta \mathbb{E}[F_0(0)] = \tau\delta. \quad (11)$$

This inequality allows immediately to derive a concentration inequality. The only missing ingredient is the correct prior on the betting fraction  $\beta$  that gives us the optimal bound. We derive it in the following lemma, whose proof is in Appendix F.

**Lemma 8** Set  $\tau = 1$  and let  $\pi(\beta) = \frac{1}{2\pi|\beta|(\ln^2(\sigma_{\text{ID}}|\beta|)+1)}$  be the prior. Assume  $d = 1$ ,  $g_t = 0, \forall t$ . Let  $\xi_t$  be sub-Gaussian (i.e.,  $b = 0$ ). Then,

$$F_t \left( \sum_{s=1}^t \hat{g}_s \right) \geq \frac{\exp \left( \frac{(\sum_{s=1}^t \hat{g}_s)^2}{2t\sigma_{\text{ID}}^2} \right)}{2\pi\sqrt{e} \frac{\sum_{s=1}^t \hat{g}_s}{\sqrt{t\sigma_{\text{ID}}^2}} \left( \ln^2 \frac{\sum_{s=1}^t \hat{g}_s}{t\sigma_{\text{ID}}} + 1 \right)}.$$

Furthermore, the noisy coin betting potential  $F_t$  implies

$$\mathbb{P} \left( \sup_t \left| \sum_{s=1}^t \hat{\mathbf{g}}_s \right| \geq \sigma_{\text{ID}} \sqrt{2t \ln \left( \left( \frac{6\pi\sqrt{e}}{\delta} \right)^{3/2} \cdot (\ln^2(\sqrt{t}) + 1) \right)} \right) \leq \delta .$$

We remark that the choice of prior in Lemma 8 resembles  $\frac{1}{\beta \ln^2(\beta)}$  used by Chernov and Vovk (2010) and Koolen and van Erven (2015), but their choice does not work when the range of  $\beta$  is unbounded.

We now show that the reduction in Algorithm 2 implies a Banach-valued martingale concentration inequality. Specifically, for the Banach space being  $(2, \lambda)$ -uniformly convex and with the choice of OMD described in Section 5 as  $\mathcal{A}_{\text{D}}$ , we have  $\sum_{s=1}^t \langle \hat{\mathbf{g}}_s, \mathbf{u} - \mathbf{y}_s \rangle \leq \sqrt{\frac{2}{\lambda} \sum_{s=1}^t \|\hat{\mathbf{g}}_s\|_*^2}$  for all  $\mathbf{u}$  in the unit ball  $S \subset V$  w.p. 1. This implies, by the definition of the dual norm,

$$\left\| \sum_{s=1}^t \hat{\mathbf{g}}_s \right\|_* \leq \sqrt{\frac{2}{\lambda} \sum_{s=1}^t \|\hat{\mathbf{g}}_s\|_*^2} + \sum_{s=1}^t \langle \hat{\mathbf{g}}_s, \mathbf{y}_s \rangle .$$

Since  $\langle \mathbf{y}_s, \hat{\mathbf{g}}_s \rangle$  is the loss fed into BANCO, Lemma 8 implies that

$$\mathbb{P} \left( \max_t \left\| \sum_{s=1}^t \hat{\mathbf{g}}_s \right\|_* \geq \sqrt{\frac{2}{\lambda} \sum_{s=1}^t \|\hat{\mathbf{g}}_s\|_*^2} + \sigma_{\text{ID}} \sqrt{2t \ln \left( \left( \frac{6\pi\sqrt{e}}{\delta} \right)^{3/2} (\ln^2(\sqrt{t}) + 1) \right)} \right) \leq \delta .$$

## 8. Conclusion and Future Work

In this paper, we introduced the unconstrained OCO problem with subgradients corrupted by sub-exponential noise, motivated by a recent pessimistic results on learning with unbounded subgradients. Straightforward extensions of existing algorithms do not result in optimal regret rates. Hence, we proposed a new algorithm called BANCO via the noisy coin betting framework, which achieves the same optimal minimax regret rate as in the noise-free unconstrained OCO w.r.t. the comparator  $\|\mathbf{u}\|$  and the horizon  $T$ . Our lower bound on the noise level  $\sigma$  implies that the regret bound of BANCO is optimal up to constant factors. Numerous applications follow naturally including differential privacy, which provides the first parameter-free subgradient descent algorithm for differential privacy.

Our study opens up numerous research directions. First, one immediate difference in our upper bound from the standard noise-free OCO algorithms is that we do not have a data-dependent regret bound; we have  $(G^2 + \sigma^2)T$  rather than  $\mathbb{E}[\sum_{t=1}^T \|\hat{\mathbf{g}}_t\|_*^2]$ . It would be interesting to investigate whether data-dependent bounds are possible. Second, it would be desirable not to require the knowledge of the noise through  $(\sigma^2, b)$ . While there are cases where the noise is known ahead of time, such as in private SGD, in the vast majority of applications data arrives through a noisy channel with an unknown noise. Third, it would be interesting to consider more general noise conditions such as heavy-tailed distributions. Finally, high probability regret bounds would be a straightforward research direction.

## Acknowledgments

We would like to thank Adam Smith for his valuable feedback on differentially-private SGDs.

## References

- A. Agarwal, P. L. Bartlett, P. Ravikumar, and M. J. Wainwright. Information-theoretic lower bounds on the oracle complexity of stochastic convex optimization. *IEEE Trans. on Information Theory*, 58(5):3235, 2012.
- A. Chernov and V. Vovk. Prediction with advice of unknown number of experts. In *Proc. of the Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 117–125, 2010.
- A. Cutkosky and K. Boahen. Online learning without prior information. In *Proc. of the Conference On Learning Theory (COLT)*, pages 643–677, 2017.
- A. Cutkosky and F. Orabona. Black-box reductions for parameter-free online learning in Banach spaces. In *Proc. of the Conference On Learning Theory (COLT)*, pages 1493–1529, 2018.
- J. C. Duchi, M. I. Jordan, and M. J. Wainwright. Privacy aware learning. *Journal of the ACM*, 61(6): 38, 2014.
- R. Durrett. *Probability: theory and examples*. Cambridge University Press, 2010.
- C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006.
- D. J. Foster, A. Rakhlin, and K. Sridharan. Adaptive online learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 3375–3383, 2015.
- D. J. Foster, S. Kale, M. Mohri, and K. Sridharan. Parameter-free online learning via model selection. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 6020–6030, 2017.
- D. J. Foster, A. Rakhlin, and K. Sridharan. Online learning: Sufficient statistics and the Burkholder method. In *Proc. of the Conference On Learning Theory (COLT)*, pages 3028–3064, 2018.
- Daniel Hsu, Sham Kakade, Tong Zhang, et al. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 17, 2012.
- S. P. Kasiviswanathan, H. K. Lee, K. Nissim, S. Raskhodnikova, and A. Smith. What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826, 2011.
- W. M. Koolen and T. van Erven. Second-order quantile methods for experts and combinatorial games. In *Proc. of the Conference On Learning Theory (COLT)*, pages 1155–1175, 2015.
- Wouter M. Koolen. Implementing squint, 2015. URL [http://blog.wouterkoolen.info/Squint\\_implementation/post.html](http://blog.wouterkoolen.info/Squint_implementation/post.html).
- W. Kotłowski. Scale-invariant unconstrained online learning. In *Proc. of the International Conference on Algorithmic Learning Theory (ALT)*, 2017.
- N. Littlestone. From on-line to batch learning. In *Proc. of the Conference On Learning Theory (COLT)*, pages 269–284, 1989.
- Y. L. Luke. Inequalities for generalized hypergeometric functions. *Journal of Approximation Theory*, 5(1):41–65, 1972.

- J. Matoušek. *Lectures on discrete geometry*, volume 212. Springer New York, 2002.
- H. B. McMahan and J. Abernethy. Minimax optimal algorithms for unconstrained linear optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2724–2732, 2013.
- H. B. McMahan and F. Orabona. Unconstrained online linear learning in Hilbert spaces: Minimax algorithms and normal approximations. In *Proc. of the Conference On Learning Theory (COLT)*, pages 1020–1039, 2014.
- F. Orabona. Dimension-free exponentiated gradient. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 1806–1814, 2013.
- F. Orabona and D. Pál. Coin betting and parameter-free online learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 577–585. 2016.
- F. Orabona and T. Tommasi. Training deep networks without learning rates through coin betting. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2160–2170, 2017.
- I. Pinelis. Rosenthal-type inequalities for martingales in 2-smooth Banach spaces. *Theory of Probability & Its Applications*, 59(4):699–706, 2015.
- A. Rakhlin and K. Sridharan. On equivalence of martingale tail bounds and deterministic regret inequalities. In *Proc. of the Conference On Learning Theory (COLT)*, pages 1704–1722, 2017.
- S. Song, K. Chaudhuri, and A. Sarwate. Learning from data with heterogeneous noise using SGD. In *Proc. of International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 894–902, 2015.
- N. Srebro, K. Sridharan, and A. Tewari. On the universality of online mirror descent. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2645–2653, 2011.
- L. Wasserman and S. Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010.
- X. Wu, F. Li, A. Kumar, K. Chaudhuri, S. Jha, and J. Naughton. Bolt-on differential privacy for scalable stochastic gradient descent-based analytics. In *Proc. of the 2017 ACM International Conference on Management of Data*, pages 1307–1322. ACM, 2017.

## Appendix A. OGD with linear losses

We show that for linear losses OGD's expected regret does not scale with the noise level  $\sigma$ .

Consider the linear losses  $\ell_t(\mathbf{x}) = -\langle \mathbf{g}_t, \mathbf{x} \rangle$ . Let  $G = 1$  for simplicity. Assume that the loss functions are set before the game starts. That is,  $\mathbf{g}_t$ 's are deterministic. The standard OGD makes predictions by  $\mathbf{w}_t = \eta \sum_{s=1}^{t-1} \hat{\mathbf{g}}_s$ . Let  $\mathbf{w}_t^*$  be the prediction that OGD would have made in the noise-free setting:  $\mathbf{w}_t^* = \eta \sum_{s=1}^{t-1} \mathbf{g}_s$ . It is easy to see that  $\mathbb{E} \mathbf{w}_t = \mathbb{E} \eta \sum_{s=1}^{t-1} \hat{\mathbf{g}}_s = \eta \sum_{s=1}^{t-1} \mathbf{g}_s = \mathbf{w}_t^*$ . Therefore, the expected regret of OGD satisfies

$$\mathbb{E} \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} - \mathbf{w}_t \rangle = \mathbb{E} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u} - \mathbf{w}_t \rangle = \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u} - \mathbf{w}_t^* \rangle.$$

Therefore, let alone the data-dependent regret, OGD has a regret bound of  $O((\|\mathbf{u}\|^2 + 1)\sqrt{T})$  with a tuned  $\eta$ . Interestingly, the regret bound does not involve  $\sigma$ . However, one cannot expect to be free from  $\sigma$  in general. Indeed, our lower bound in Theorem 6 shows that the factor  $\sigma$  must be present in general.

## Appendix B. Proof of Theorem 3

From one direction of the equivalence, we have

$$\begin{aligned} \mathbb{E}[\hat{R}_T^{\text{Lin}}(\mathbf{u})] &= \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} - \mathbf{w}_t \rangle \right] = \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} \rangle - \text{Wealth}_T + \tau \right] \\ &\leq \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} \rangle - F \left( \sum_{t=1}^T \hat{\mathbf{g}}_t \right) + \tau \right] \\ &\leq \mathbb{E} \left[ \max_{\mathbf{x} \in V^*} \langle \mathbf{x}, \mathbf{u} \rangle - F(\mathbf{x}) + \tau \right] = F_T^*(\mathbf{u}) + \tau. \end{aligned}$$

In the other direction, we have that

$$\text{Wealth}_T - \tau = \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} \rangle - \hat{R}_T^{\text{Lin}}(\mathbf{u}), \quad \forall \mathbf{u} \in V,$$

that implies

$$\mathbb{E}[\text{Wealth}_T - \tau] = \mathbb{E} \left[ \max_{\mathbf{u} \in V} \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} \rangle - \text{Regret}_T(\mathbf{u}) \right].$$

Hence, we have

$$\mathbb{E}[\text{Wealth}_T - \tau] \geq \mathbb{E} \left[ \max_{\mathbf{u} \in V} \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} \rangle - F^*(\mathbf{u}) - \tau \right] = \mathbb{E} \left[ F \left( \sum_{t=1}^T \hat{\mathbf{g}}_t \right) - \tau \right].$$

## Appendix C. Proof of Theorem 4

From the definition of the Fenchel duality we have

$$f^*(u) = \max_{\theta} u\theta - f(\theta) = u\theta^* - f(\theta^*),$$

where  $\theta^* = \arg \max_{\theta} u\theta - f(\theta)$ . Define  $\beta^* = \arg \max_{\beta} \exp(\beta\theta^* - \beta^2 S)$ , that is  $\beta^* = \frac{\theta^*}{2S}$ . Assume that  $\theta^* \geq 0$ . The reasoning is analogous for  $\theta^* < 0$ . In fact, one can show that the function is even.

We perform a case-by-case analysis. We first assume that  $\theta^* \leq \sqrt{2S}$ . Then,

$$f^*(u) \leq u\theta^* - f(\theta^*) \leq |u|\sqrt{2S},$$

from which the stated bound follows. Hence, we can safely assume  $\theta^* > \sqrt{2S}$ , which is equivalent to  $\beta^* \geq \frac{1}{\sqrt{2S}}$ . Let  $[v_1, v_2] \subseteq [-a, a \wedge \beta^*]$ . Recall that we use the uniform prior:  $\pi(\beta) = 1/(2a), \forall \beta \in [-a, a]$ . The following inequality becomes useful:

$$\begin{aligned} f(\theta^*) &= \frac{\tau}{2a} \int_{-a}^a \exp(\beta\theta^* - \beta^2 S) d\beta \geq \frac{\tau}{2a} \int_{v_1}^{v_2} \exp(\beta\theta^* - \beta^2 S) d\beta \\ &\geq \frac{v_2 - v_1}{2a} \tau \exp(v_1\theta^* - v_1^2 S). \end{aligned} \quad (12)$$

**Case 1:**  $\beta^* \leq a$ .

Using (12) with  $v_1 = \beta^* - \frac{1}{\sqrt{2S}}$  and  $v_2 = \beta^*$ , we have

$$f(\theta^*) = \frac{\tau}{2a} \int_{-a}^a \exp(\beta\theta^* - \beta^2 S) d\beta \geq \frac{\tau}{2a\sqrt{2S}} \exp\left(\frac{(\theta^*)^2}{4S} - \frac{1}{2}\right).$$

Hence, we have

$$\begin{aligned} f^*(u) &\leq |u|\theta^* - \frac{\tau}{2a\sqrt{2S}} \exp\left(\frac{(\theta^*)^2}{4S} - \frac{1}{2}\right) \\ &\leq \max_x x|u| - \frac{\tau}{2a\sqrt{2S}} \exp\left(\frac{x^2}{4S} - \frac{1}{2}\right). \end{aligned}$$

To solve the problem above, we consider the following stylized problem:

$$\max_x x|u| - A \exp(Bx^2 - C).$$

We see by setting the gradient to zero that  $A(2Bx) \exp(Bx^2 - C) = |u| \implies 4A^2 B^2 x^2 \exp(2Bx^2 - 2C) = u^2$ . Letting  $z = 2Bx^2$  and  $D = \frac{u^2}{2A^2 B e^{-2C}}$ , we have  $z \exp(z) = D$ . Using Lambert function, we have  $z = W_0(D)$  and so  $x = \sqrt{\frac{W_0(D)}{2B}}$ , which we call  $x^*$ . We use the upper bound on  $W_0(y) \leq \ln(1+y)$  for  $y > 0$  by [Orabona and Pál \(2016, Lemma 17\)](#). Then, plugging in  $A = \frac{\tau}{2a\sqrt{2S}}$ ,  $B = 1/(4S)$ , and  $C = 1/2$ ,

$$f^*(u) \leq x^*|u| \leq |u| \sqrt{2S \ln\left(1 + \frac{16ea^2 S^2 u^2}{\tau^2}\right)}.$$

**Case 2:**  $\beta^* > a$ .

In this case, we have  $\theta^* > 2Sa$ . Then, choose  $v_1 = a - c$  and  $v_2 = a$  to arrive at

$$\begin{aligned} f(\theta^*) &= \frac{1}{2a} \int_{-a}^a \tau \exp(\beta\theta^* - \beta^2 S) d\beta \geq \frac{v_2 - v_1}{2a} \tau \exp(v_1\theta^* - v_1^2 S) \\ &\geq \frac{c}{2a} \tau \exp\left(v_1\theta^* - v_1^2 \frac{\theta^*}{2a}\right) = \frac{c}{2a} \tau \exp(\theta^* Q), \end{aligned}$$

where  $Q = v_1 - \frac{v_1^2}{2a}$ . Using  $\theta^* > 0$ ,

$$f^*(u) \leq u\theta^* - \frac{c}{2a} \tau \exp(\theta^* Q) \leq \max_{\theta} |u|\theta - \frac{c}{2a} \tau \exp(\theta Q) = \frac{|u|}{Q} \ln\left(|u| \frac{2a}{ec\tau Q}\right).$$

Setting  $c = a/2$ , we have  $Q = \frac{3}{8}a$ , which leads to  $f^*(u) = \frac{8}{3a}|u| \ln\left(\frac{32}{3ea\tau}|u|\right)$ .



## Appendix D. Proof of Theorem 5

Observe that  $|s_t| \leq \|\hat{\mathbf{g}}_t\|_* \|\mathbf{y}_t\| \leq \|\hat{\mathbf{g}}_t\|_*$  since  $\|\mathbf{y}_t\| \leq 1$  for all  $t$ . Furthermore,

$$\begin{aligned}
 \mathbb{E} \text{Regret}_T(\mathbf{u}) &\leq \mathbb{E} \left[ \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u} - \mathbf{x}_t \rangle \right] = \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} - \mathbf{x}_t \rangle \right] \\
 &= \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} \rangle - \langle \hat{\mathbf{g}}_t, w_t \mathbf{y}_t \rangle \right] \\
 &= \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} \rangle - \langle \hat{\mathbf{g}}_t, \mathbf{y}_t \rangle \|\mathbf{u}\| + \langle \hat{\mathbf{g}}_t, \mathbf{y}_t \rangle \|\mathbf{u}\| - \langle \hat{\mathbf{g}}_t, \mathbf{y}_t \rangle w_t \right] \\
 &\leq \|\mathbf{u}\| \mathbb{E} \left[ \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u} / \|\mathbf{u}\| \rangle - \langle \hat{\mathbf{g}}_t, \mathbf{y}_t \rangle \right] + R_T^M(\|\mathbf{u}\|) \\
 &\leq \|\mathbf{u}\| R_T^D(\mathbf{u} / \|\mathbf{u}\|) + R_T^M(\|\mathbf{u}\|).
 \end{aligned}$$

## Appendix E. Proof of Theorem 6

It is not hard to see that a stochastic optimization lower bound imply an online learning lower bounds. This is due to the online to batch conversion (Littlestone, 1989) which implies stochastic optimization is “not harder” than online learning. Specifically, suppose we have a lower bound on the convergence of stochastic optimization for convex functions:  $\mathbb{E} F(\mathbf{x}_T) - F(\mathbf{x}^*) \geq c/\sqrt{T}$ . Then, we can claim a lower bound in the online convex optimization:  $\mathbb{E} \sum_{t=1}^T f_t(\mathbf{x}'_t) - f_t(\mathbf{u}) \geq c/\sqrt{T}$ . Here is a proof: Suppose a better rate is possible in online learning with some method:  $\mathbb{E} \sum_{t=1}^T f_t(\mathbf{x}'_t) - f_t(\mathbf{u}) < c/\sqrt{T}$ . One can then perform online learning with  $f_t = F$  where the online learner acquires noisy version  $\hat{f}_t$ . With the online-to-batch conversion, this solves the stochastic optimization with a better rate:  $\mathbb{E} F(\frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t) - F(\mathbf{x}^*) < c/\sqrt{T}$ , which is a contradiction.

Therefore, it suffices to show a lower bound on stochastic optimization. Before presenting the lower bound statement, we describe the problem setup. We closely follow the setup of Agarwal et al. (2012). Let  $\mathbb{S} \subseteq \mathbb{R}^d$ . Let the function class  $\mathcal{F}$  consists of functions  $f : \mathbb{S} \rightarrow \mathbb{R}$  that are convex and 1-Lipschitz w.r.t.  $\ell_p$ -norm:  $|f(\mathbf{x}) - f(\mathbf{y})| \leq L \|\mathbf{x} - \mathbf{y}\|_p, \forall \mathbf{x}, \mathbf{y} \in \mathbb{S}$ . An algorithm  $\mathcal{M}$  has access to  $T$  calls of the first order oracle and outputs  $\mathbf{x}_T$  after  $T$  calls to the oracle (hereafter, we color definitions with light blue for the benefit of readers). The oracle  $\psi_\sigma(\mathbf{x}, f)$  takes  $\mathbf{x} \in \mathbb{S}$  and returns  $(\hat{f}(\mathbf{x}), \hat{\mathbf{z}}(\mathbf{x}))$  where  $\hat{f}(\mathbf{x})$  is the noisy function value and  $\hat{\mathbf{z}}(\mathbf{x})$  is a noisy subgradient such that  $\mathbb{E} \hat{\mathbf{z}}(\mathbf{x}) \in \partial f(\mathbf{x})$ . The oracle guarantees a noise condition  $\mathbb{E} \|\hat{\mathbf{z}}(\mathbf{x}) - \mathbb{E} \hat{\mathbf{z}}(\mathbf{x})\|_q^2 \leq \sigma^2$ . Our goal is to find a lower bound on

$$\epsilon^* := \inf_{\mathcal{M}} \sup_{f \in \mathcal{F}} \mathbb{E} [f(\mathbf{x}_T) - f(\mathbf{x}_f^*)],$$

where  $\mathbf{x}_f^*$  is the minimizer of  $f$ . The quantity  $\epsilon^*$  depends on  $d, T, \sigma$ , and  $\mathbb{S}$ .

Let  $\mathbb{B}_\infty(r)$  be the  $\infty$ -norm ball with radius  $r$ . We present our stochastic optimization lower bound in Theorem 9 below. The difference from the lower bound in Agarwal et al. (2012) is that the bound therein is that (i) they obscure the dependence on the noise  $\sigma$  by equating it to the Lipschitz constant and (ii) they assume uncentered second moment noise bound  $\mathbb{E} \|\hat{\mathbf{z}}(\mathbf{x})\|_q^2 \leq \sigma^2$  rather than the variance of  $\|\hat{\mathbf{z}}(\mathbf{x})\|_q$ . Departing from the prior work, we consider a different function class that keeps the Lipschitz constant at 1 while allowing the noise level  $\sigma$  to be arbitrarily large.

**Theorem 9** *Let  $r$  be the largest number such that  $\mathbb{B}_\infty(r) \subseteq \mathbb{S}$ . Let  $\sigma \geq 2$ . Then, there exists a universal constant  $c_0$  such that*

$$q \in [1, 2] \implies \epsilon^* \geq \min \left\{ c_0 \sigma \frac{r\sqrt{d}}{\sqrt{T}}, \frac{r}{18} \right\} \quad \text{and}$$

$$q \geq 2 \implies \epsilon^* \geq \min \left\{ c_0 \sigma \frac{rd^{1/q-1}}{\sqrt{T}}, \frac{rd^{1/q-1}}{18} \right\}.$$

**Proof** The proof closely follows [Agarwal et al. \(2012\)](#), but we consider a different function class. The key idea is to construct a function class such that identification of the target function is equivalent to identification of coefficients  $\{\alpha_i \in [0, 1]\}$ ,  $i \in \{1, \dots, d\}$ , on a set of basis functions. Furthermore, the construction defines an oracle such that each query amounts to revealing a coin outcome  $\{0, 1\} \sim \text{Bernoulli}(\alpha_i)$  for some  $i$ 's (details vary for different  $q$ 's). Then, the number of observations in statistical estimation is directly connected to the number of oracle calls, allowing a statistical lower bound to imply an iteration complexity of stochastic optimization.

Let  $\mathcal{V} \subseteq \{\pm 1\}^d$  has  $M$  distinct vertices of  $d$ -dimensional hypercube such that (i)  $\mathcal{V}$  is  $\frac{d}{4}$ -packing w.r.t. hamming distance (i.e.,  $\sum_i \mathbb{1}\{\alpha_i \neq \beta_i\} \geq \frac{d}{4}, \forall \alpha \neq \beta \in \mathcal{V}$ ) and (ii)  $M \geq (2/\sqrt{e})^{d/2} \approx 1.1^d$ . Such a packing is known to be possible ([Matoušek, 2002](#)). We define the function class  $\mathcal{G}(\delta)$  that consists of

$$\forall \alpha \in \mathcal{V}, \quad g_\alpha(\mathbf{x}) := \frac{c}{d} \sum_i \left( \left( \frac{1}{2} + \alpha_i \delta \right) f_i^+(\mathbf{x}) + \left( \frac{1}{2} - \alpha_i \delta \right) f_i^-(\mathbf{x}) \right) \quad \text{where}$$

$$f_i^+(\mathbf{x}) := \begin{cases} -x_i + r(-\sigma - 1) & \text{if } x_i \leq -r \\ \sigma x_i & \text{if } -r \leq x_i \leq r \\ x_i + r(\sigma - 1) & \text{if } r \leq x_i \end{cases} \quad \text{and} \quad f_i^-(\mathbf{x}) := \begin{cases} -x_i + r(\sigma - 1) & \text{if } x_i \leq -r \\ -\sigma x_i & \text{if } -r \leq x_i \leq r \\ x_i + r(-\sigma - 1) & \text{if } r \leq x_i \end{cases}.$$

We assume that  $\delta \leq \frac{1}{2\sigma}$ , which ensures the convexity of  $g_\alpha$ .

**Case 1:**  $q \in [1, 2]$ .

For this case, we assume an oracle that first chooses  $I \in [d]$  uniformly at random, draw  $b_I \in \{0, 1\}$  with  $\text{Ber}(1/2 + \alpha_I \delta)$ , and then return the function value and the subgradient of

$$\hat{g}_\alpha(\mathbf{x}) = c(b_I f_I^+(\mathbf{x}) + (1 - b_I) f_I^-(\mathbf{x})).$$

Thus, the learner only sees either  $c f_I^+(\mathbf{x})$  or  $c f_I^-(\mathbf{x})$ , and the function value and the subgradient are unbiased. Denote by  $\hat{z}_\alpha(\mathbf{x})$  be the noisy subgradient returned by the oracle such that  $z_\alpha(\mathbf{x}) := \mathbb{E}[\hat{z}_\alpha(\mathbf{x})] \in \partial g_\alpha(\mathbf{x})$ .

Some facts on the subgradient norms:

- $\|z_\alpha(\mathbf{x})\|_q^2 \leq \max\left\{\frac{c^2}{d^2}, \frac{4c^2\delta^2\sigma^2}{d^2}\right\} \|\mathbf{1}\|_q^2 = c^2 d^{(2/q)-2}$ .
- $\mathbb{E} \|\hat{z}_\alpha(\mathbf{x})\|_q^2 \leq c^2 \sigma^2$ .
- $\mathbb{E} \|\hat{z}_\alpha(\mathbf{x}) - z_\alpha(\mathbf{x})\|_q^2 \leq 2 \mathbb{E}[\|\hat{z}_\alpha(\mathbf{x})\|_q^2] + 2 \|z_\alpha(\mathbf{x})\|_q^2 \leq 2(c^2 \sigma^2 + c^2 d^{(2/q)-2})$ .

By setting  $c = 1/2$ ,  $g_\alpha(\mathbf{x})$  is 1-Lipschitz and the noise variance is bounded:  $\mathbb{E} \|\hat{z}_\alpha(\mathbf{x}) - z_\alpha(\mathbf{x})\|_q^2 \leq \sigma^2$ .

We define a premetric  $\rho$ :

$$\rho(f, g) := \inf_{\mathbf{x} \in \mathbb{S}} f(\mathbf{x}) + g(\mathbf{x}) - f(\mathbf{x}_f^*) - g(\mathbf{x}_g^*)$$

which is 0 if and only if  $\mathbf{x}_f^* = \mathbf{x}_g^*$  (assuming  $f$  and  $g$  have a unique minimizer). Define  $\psi(\delta) := \min_{\alpha \neq \beta \in \mathcal{V}} \rho(g_\alpha, g_\beta)$ . We study  $\rho(g_\alpha, g_\beta)$  where  $\alpha, \beta \in \mathcal{V}$  such that  $\alpha \neq \beta$ . By examining the function carefully, one can show that  $\rho(g_\alpha, g_\beta) \geq \frac{c}{d} (\sum_i \mathbf{1}\{\alpha_i \neq \beta_i\}) 4\delta r \sigma$ . Since  $\sum_i \mathbf{1}\{\alpha_i \neq \beta_i\} \geq \frac{d}{4}, \forall \alpha \neq \beta \in \mathcal{V}$ , we have

$$\psi(\delta) = \min_{\alpha \neq \beta} \rho(g_\alpha, g_\beta) \geq c\delta r \sigma. \quad (13)$$

Now, the main argument is as follows. If  $\epsilon^* \geq \frac{cr}{18}$ , then we have the half of the theorem statement. Therefore, it suffices to consider the regime  $\epsilon^* < \frac{cr}{18}$ .

In this regime, we consider the function class  $\mathcal{G}(\delta)$  with  $\delta = \frac{9}{cr\sigma} \epsilon^*$ . This implies that (i)  $\delta \leq \frac{9}{cr\sigma} \frac{cr}{18} = \frac{1}{2\sigma} \leq \frac{1}{4}$  and that (ii) there exists a method  $\mathcal{M}^*$  such that  $\sup_{f \in \mathcal{G}(\delta)} \mathbb{E}[f(\mathbf{x}_T) - f(\mathbf{x}_f^*)] \leq \epsilon^* = \frac{c\delta r \sigma}{9} \leq \psi(\delta)/9$  by the definition of  $\epsilon^*$  and (13).

By Agarwal et al. (2012, Lemma 2), these two conditions,  $\delta \leq 1/4$  and  $\sup_{f \in \mathcal{G}(\delta)} \mathbb{E}[f(\mathbf{x}_T) - f(\mathbf{x}_f^*)] \leq \psi(\delta)/9$ , imply the following: For any  $\alpha^* \in \mathcal{V}$ , facing to solve the optimization problem with the function  $g_{\alpha^*}$ , one can invoke  $\mathcal{M}^*$  to construct an estimator  $\hat{\alpha} \in \mathcal{V}$  of the true  $\alpha^*$ :

$$\forall \alpha^* \in \mathcal{V}, \mathbb{P}(\hat{\alpha} \neq \alpha^*) \leq 1/3.$$

On the other hand, Agarwal et al. (2012, Lemma 3) use Fano's inequality to show that

$$\mathbb{P}(\hat{\alpha} \neq \alpha^*) \geq 1 - 2 \frac{16T\delta^2 + \ln 2}{d \ln(2/\sqrt{e})}.$$

Combining these two results, we have  $1 - 2 \frac{16T\delta^2 + \ln 2}{d \ln(2/\sqrt{e})} \leq \frac{1}{3}$ . Using  $\delta = \frac{9}{cr\sigma} \epsilon^*$ , one can show that, for  $d \geq 11$ ,

$$\epsilon^* = \Omega\left(c \frac{\sigma r \sqrt{d}}{\sqrt{T}}\right).$$

For  $d \leq 10$ , simply consider a reduction to  $d = 1$  case and use the Le Cam's bound (Agarwal et al., 2012, Lemma 4). This completes the first part of the proof.

**Case 2:**  $q \geq 2$ .

For the second part, we consider a different oracle that chooses  $d$  independent coin flips  $b_i \sim \text{Bernoulli}(\frac{1}{2} + \alpha_i \delta)$ ,  $i \in \{1, \dots, d\}$ , and return the function value and the subgradient of

$$\hat{g}_\alpha(\mathbf{x}) = \frac{c}{d} \sum_i (b_i f_i^+(\mathbf{x}) + (1 - b_i) f_i^-(\mathbf{x})).$$

This provides unbiased function values and subgradients, and corresponds to revealing one coin outcome for each dimension. While this provides more information for the coin tossing (easier problem), but it allows steeper per-coordinate subgradients than the oracle A (harder problem), given the same Lipschitz constants.

The difference of the proof is just on the subgradient norms and how we set  $c$ . Recall that  $\|z_\alpha(\mathbf{x})\|_q^2 \leq \max\{\frac{c^2}{d^2}, \frac{4c^2\delta^2\sigma^2}{d^2}\} \|\mathbf{1}\|_q^2 = c^2 d^{(2/q)-2}$ . One can see that  $\mathbb{E}\|\hat{z}_\alpha(\mathbf{x})\|_q^2 = \frac{c^2}{d^2} \sigma^2 \|\mathbf{1}\|_q^2 = c^2 \sigma^2 d^{\frac{2}{q}-2}$ . Then, the subgradient noise variance is bounded:

$$\mathbb{E}\|\hat{z}_\alpha(\mathbf{x}) - z_\alpha(\mathbf{x})\|_q^2 \leq 2\mathbb{E}[\|\hat{z}_\alpha(\mathbf{x})\|_q^2] + 2\|z_\alpha(\mathbf{x})\|_q^2 \leq 2(c^2 \sigma^2 d^{\frac{2}{q}-2} + c^2 d^{\frac{2}{q}-2}) \leq 4c^2 \sigma^2 d^{\frac{2}{q}-2}.$$

By setting  $c = \frac{1}{2} d^{-(\frac{1}{q}-1)}$ , we satisfy 1-Lipschitz ( $\|z_\alpha(\mathbf{x})\|_q \leq 1$ ) and the noise level controlled:  $\mathbb{E}\|\hat{z}_\alpha(\mathbf{x}) - z_\alpha(\mathbf{x})\|_q^2 \leq \sigma^2$ .

Again, the oracle here is equivalent to discovering all the  $d$  coin outcomes in each iteration rather than one. By Agarwal et al. (2012, Lemma 3) with  $\ell = d$ , we have that  $\mathbb{P}(\hat{\alpha} \neq \alpha^*) \geq 1 - 2 \frac{16Td\delta^2 + \ln 2}{d \ln(2/\sqrt{e})}$ . With the same logic, we have  $1 - 2 \frac{16Td\delta^2 + \ln 2}{d \ln(2/\sqrt{e})} \leq \frac{1}{3}$ . Again, by  $\delta = \frac{9}{c\sigma} \epsilon^*$ , one can show that, for  $d \geq 11$ ,

$$\epsilon^* = \Omega\left(c \frac{\sigma r}{\sqrt{T}}\right) = \Omega\left(d^{1-1/q} \frac{\sigma r}{\sqrt{T}}\right).$$

For  $d \leq 10$ , the same argument as the case 1 can be made.  $\blacksquare$

To prove Theorem 6, it suffices to notice that the largest  $r$  such that  $\mathbb{B}_\infty(r) \subseteq \mathbb{S}$  with  $\mathbb{S}$  being the  $\ell_p$ -norm ball of radius  $U$  is  $r = Ud^{-1/p}$ .

## Appendix F. Proof of Lemma 8

**Proof** It suffices to consider  $\sigma_{\text{ID}} = 1$  since the result for  $\sigma_{\text{ID}} \neq 1$  can be obtained by replacing  $S$  below with  $\sum_{s=1}^t \hat{g}_s / \sigma_{\text{ID}}$ . Let  $S = \sum_{s=1}^t \hat{g}_s$ . Define  $\beta^* = S/t$  and  $u = \beta^* - \frac{1}{\sqrt{t}}$ . Then,  $\exp(\beta S - \beta^2 t/2)$  is maximized at  $\beta = \beta^*$  and increasing in  $[u, \beta^*]$ . Recall that  $F_t(\sum_{s=1}^t \hat{g}_s) = \int_{-\infty}^{\infty} \pi(\beta) \exp\left(\beta S - \frac{\beta^2 t}{2}\right) d\beta$ . To evaluate the integral, it suffices to assume  $S \geq 0$  since the integrand is symmetric. Using the fact that the prior is nonincreasing in  $(0, \infty)$ ,

$$\begin{aligned} F_t\left(\sum_{s=1}^t \hat{g}_s\right) &\geq \frac{1}{2\pi} \int_u^{\beta^*} \frac{1}{\beta^* (\ln^2 \beta^* + 1)} \exp(uS - u^2 t/2) d\beta \\ &= \frac{1}{2\pi} \frac{1/\sqrt{t}}{\beta^* (\ln^2 \beta^* + 1)} \cdot \exp(uS - u^2 t/2) \\ &= \frac{1}{2\pi} \frac{1}{\frac{S}{\sqrt{t}} (\ln^2 \frac{S}{t} + 1)} \cdot \exp\left(\frac{S^2}{2t} - \frac{1}{2}\right). \end{aligned}$$

By (11),

$$\mathbb{P}\left(\max_t \frac{1}{2\pi\sqrt{e}} \frac{1}{\frac{S}{\sqrt{t}} (\ln^2 \frac{S}{t} + 1)} \cdot \exp\left(\frac{S^2}{2t}\right) \geq \frac{1}{\delta}\right) \leq \mathbb{P}\left(\max_t F_t(S) \geq \frac{1}{\delta}\right) \leq \delta.$$

Rearranging the inequality in the LHS above, we have

$$\max_t S^2 \geq 2t \ln\left(\frac{2\pi\sqrt{e}}{\delta} \cdot \frac{S}{\sqrt{t}} \left(\ln^2 \frac{S}{t} + 1\right)\right).$$

To complete the proof, it suffices to find a tighter and simpler inequality. This is equivalent to assuming  $S^2 \leq [\text{the RHS above}]$  and deriving an upper bound on  $S^2$ , then inverting it. Therefore, it suffices to show

$$S^2 < 2t \ln\left(\frac{2\pi\sqrt{e}}{\delta} \cdot \frac{S}{\sqrt{t}} \left(\ln^2 \frac{S}{t} + 1\right)\right) \implies S^2 < 2t \ln\left(\left(\frac{6\pi\sqrt{e}}{\delta}\right)^{3/2} \cdot (\ln^2(\sqrt{t}) + 1)\right). \quad (14)$$

Let  $A = 2\pi\sqrt{e}/\delta$ . Using  $\ln^2(x) + 1 \leq x, \forall x \geq 1$ , and  $x \leq (1/2) \ln x, \forall x > 0$ ,

$$S^2 < 2t \ln\left(A \cdot \frac{S}{\sqrt{t}} \left(\ln^2 \frac{S}{t} + 1\right)\right)$$

$$\begin{aligned}
 &\leq 2t \ln \left( A \cdot \frac{S^2}{t^{3/2}} \right) = 4t \ln \left( \sqrt{A} \cdot \frac{S}{t^{3/4}} \right) \\
 &\leq 2t \cdot \sqrt{A} \cdot \frac{S}{t^{3/4}} \\
 \implies &S \leq 2t^{1/4} \sqrt{A} \\
 \stackrel{(a)}{\implies} &S^2 < 2t \ln \left( 2 \frac{A^{3/2}}{t^{1/4}} \left( \ln^2 \frac{S}{t} + 1 \right) \right),
 \end{aligned}$$

where (a) is by the first inequality.

It suffices to assume the regime  $S^2 > t$  since  $S^2 \leq t$  trivially implies the RHS of (14). Since  $\ln^2 x$  is decreasing up to 1 and then increasing, we perform a case by case analysis.

**Case 1:**  $S \leq t$ .

Since  $\ln^2(S/t) = \ln^2(t/S)$  and  $t/S \geq 1$ , we need to upper-bound  $t/S$ . Using  $S^2 > t$ , we have  $\ln^2(t/S) \leq \ln^2(\sqrt{t})$ , which implies the RHS of (14).

**Case 2:**  $S > t$ .

With a similar derivation as above, we have  $S^2 < 6t \ln(A^{1/3} S^{2/3} t^{-1/2}) \leq 3t^{1/2} A^{1/3} S^{2/3}$ , which implies  $S < 3^{3/4} t^{3/8} A^{1/4}$ . Then,

$$S^2 < 2t \ln \left( A \cdot \frac{S^2}{t^{3/2}} \right) \leq 2t \ln \left( A \cdot \frac{3^{3/2} t^{3/4} A^{1/2}}{t^{3/2}} \right) \leq 2t \ln \left( \frac{(3A)^{3/2}}{t^{3/4}} \right),$$

which implies the RHS of (14). ■

## Appendix G. Proof of Lemma 7

The Laplace mechanism noise can be obtained by multiplying independent random variables  $\mathbf{z}$  and  $m$ , where  $\mathbf{z}$  is drawn uniformly over the L2 ball, and  $m$  is an Erlang distribution with shape equal to  $d$  and rate  $\frac{\epsilon}{2}$  (Wu et al., 2017). This implies that

$$\mathbb{E}_{\xi}[\exp(\beta \langle \xi, \mathbf{a} \rangle)] = \mathbb{E}_{\mathbf{z}, m}[\exp(\beta m \langle \mathbf{z}, \mathbf{a} \rangle)] = \mathbb{E}_{\alpha, m}[\exp(\beta m \alpha)],$$

where  $\alpha$  is a random variable that model the cosine of the angles between  $\mathbf{z}$  and  $\mathbf{a}$ . In the one-dimensional case, it is easy to see that  $\alpha$  is a Rademacher variable. Hence, we have

$$\mathbb{E}_{\xi}[\exp(\beta \langle \xi, \mathbf{a} \rangle)] = \frac{1}{2} \mathbb{E}_m[\exp(\beta m) + \exp(-\beta m)].$$

Instead, for  $d \geq 2$ , the calculation is more involved, but we show that we still get the same result. In particular, observing that  $\langle \mathbf{z}, \mathbf{a} \rangle$  is the cosine of random angles distributed uniformly between  $-\pi$  and  $\pi$ , we have that  $\alpha$  is drawn from the distribution  $\rho_{\alpha}(x) = \frac{1}{\pi \sqrt{1-x^2}}$ . The expectation  $\mathbb{E}_{\alpha}[\exp(\beta m \alpha)]$  can be computed in a closed form, being equal to modified Bessel function of the first kind  $I_0(\beta m)$ . From Luke (1972, Formula 6.25), we use the inequality

$$\Gamma(\nu + 1) \left( \frac{2}{x} \right)^{\nu} I_{\nu}(x) < \frac{1}{2} (\exp(x) + \exp(-x)), \quad \forall x > 0, \nu > -\frac{1}{2},$$

that implies

$$\mathbb{E}_{\alpha, m}[\exp(\beta m \alpha)] < \frac{1}{2} \mathbb{E}_m[\exp(\beta m) + \exp(-\beta m)],$$

as in the one-dimensional case.

Hence, taking the expectation with respect to  $m$  and using the formula for the moment generating function of the Erlang distribution, we get

$$\begin{aligned}
 \mathbb{E}_{\alpha, m}[\exp(\beta m \alpha)] &< \frac{1}{2} \left[ \left(1 - \frac{2\beta}{\epsilon}\right)^{-d} + \left(1 + \frac{2\beta}{\epsilon}\right)^{-d} \right] \\
 &= \frac{1}{2} \left[ \exp\left(d \ln \frac{1}{1 - \frac{2\beta}{\epsilon}}\right) + \exp\left(d \ln \frac{1}{1 + \frac{2\beta}{\epsilon}}\right) \right] \\
 &= \frac{1}{2} \left[ \exp\left(d \ln \left(1 + \frac{2\beta}{\epsilon - 2\beta}\right)\right) + \exp\left(d \ln \left(1 - \frac{2\beta}{\epsilon + 2\beta}\right)\right) \right] \\
 &\leq \frac{1}{2} \left[ \exp\left(d \frac{2\beta}{\epsilon - 2\beta}\right) + \exp\left(d \frac{2\beta}{\epsilon + 2\beta}\right) \right],
 \end{aligned}$$

where in the last inequality we used the elementary  $\ln(1+x) \leq x$ ,  $\forall x > -1$ . We now observe that

$$\begin{aligned}
 \frac{1}{2} \left[ \exp\left(d \frac{2\beta}{\epsilon - 2\beta}\right) + \exp\left(d \frac{2\beta}{\epsilon + 2\beta}\right) \right] &= \frac{1}{2} \exp \frac{d \left(\frac{2\beta}{\epsilon}\right)^2}{1 - \left(\frac{2\beta}{\epsilon}\right)^2} \left[ \exp \frac{d \frac{2\beta}{\epsilon}}{1 - \left(\frac{2\beta}{\epsilon}\right)^2} + \exp \frac{-d \frac{2\beta}{\epsilon}}{1 - \left(\frac{2\beta}{\epsilon}\right)^2} \right] \\
 &\leq \exp \frac{d \left(\frac{2\beta}{\epsilon}\right)^2}{1 - \left(\frac{2\beta}{\epsilon}\right)^2} \exp \frac{d^2 \frac{2\beta^2}{\epsilon^2}}{\left(1 - \left(\frac{2\beta}{\epsilon}\right)^2\right)^2},
 \end{aligned}$$

where we used the elementary inequality  $\exp(x) + \exp(-x) \leq 2 \exp(x^2/2)$ ,  $\forall x$ . Overapproximating and using the assumption on  $\beta$ , we have the stated bound.