Proceedings of Machine Learning Research vol 99:1–22, 2019 32nd Annual Conference on Learning Theory

Lipschitz Adaptivity with Multiple Learning Rates in Online Learning

Zakaria Mhammedi ZAK.MHAMMEDI@ANU.EDU.AU
School of Engineering and Computer Science
The Australian National University and Data61
Wouter M. Koolen WMKOOLEN@CWI.NL
Centrum Wiskunde & Informatica
Amsterdam, the Netherlands
Tim van Erven TIM@TIMVANERVEN.NL
Statistics Department
Leiden University, the Netherlands

Editors: Alina Beygelzimer and Daniel Hsu

Abstract

We aim to design adaptive online learning algorithms that take advantage of any special structure that might be present in the learning task at hand, with as little manual tuning by the user as possible. A fundamental obstacle that comes up in the design of such adaptive algorithms is to calibrate a so-called step-size or learning rate hyperparameter depending on variance, gradient norms, etc. A recent technique promises to overcome this difficulty by maintaining multiple learning rates in parallel. This technique has been applied in the MetaGrad algorithm for online convex optimization and the Squint algorithm for prediction with expert advice. However, in both cases the user still has to provide in advance a Lipschitz hyperparameter that bounds the norm of the gradients. Although this hyperparameter is typically not available in advance, tuning it correctly is crucial: if it is set too small, the methods may fail completely; but if it is taken too large, performance deteriorates significantly. In the present work we remove this Lipschitz hyperparameter by designing new versions of MetaGrad and Squint that adapt to its optimal value automatically. We achieve this by dynamically updating the set of active learning rates. For MetaGrad, we further improve the computational efficiency of handling constraints on the domain of prediction, and we remove the need to specify the number of rounds in advance.

1. Introduction

We consider *online convex optimization* (OCO) of a sequence of convex functions $\ell_1, \dots, \ell_T$ over a given bounded convex domain, which become available one by one over the course of T rounds (Shalev-Shwartz, 2011; Hazan, 2016). Typically $\ell_t(\boldsymbol{u}) = \textsc{loss}(\boldsymbol{u}, \boldsymbol{x}_t, y_t)$ represents the *loss* of predicting with parameters $\boldsymbol{u}$ on the t-th data point $(\boldsymbol{x}_t, y_t)$ in a machine learning task. At the start of each round t, a learner has to predict the best parameters $\widehat{\boldsymbol{u}}_t$ for the function ℓ_t before finding out what ℓ_t is, and the goal is to minimize the *regret*, which is the difference in the sum of function values between the learner's predictions $\widehat{\boldsymbol{u}}_1, \dots, \widehat{\boldsymbol{u}}_T$ and the best fixed oracle parameters $\boldsymbol{u}$ that could have been chosen if all the functions had been given in advance. A special case of OCO is prediction with expert advice (Cesa-Bianchi and Lugosi, 2006), where the functions $\ell_t(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{l}_t \rangle$ are convex combinations of the losses $\boldsymbol{l}_t = (l_{t,1}, \dots, l_{t,K})$ of K expert predictors and the domain is the probability simplex.

Central results in these settings show that it is possible to control the regret with virtually no prior knowledge about the functions. For instance, knowing only a $\|\cdot\|_2$-upper-bound G on the gradients $\boldsymbol{g}_t = \nabla \ell_t(\widehat{\boldsymbol{u}}_t)$, the online gradient descent (OGD) algorithm guarantees $O(G\sqrt{T})$ regret by tuning its learning rate hyperparameter η_t proportional to $1/(G\sqrt{t})$ (Zinkevich, 2003), and in the case of prediction with expert advice the Hedge algorithm achieves regret $O(L\sqrt{T \ln K})$ knowing only an upper-bound L on the range $\max_k l_{t,k} - \min_k l_{t,k}$ of the expert losses (Freund and Schapire, 1997). Here G is the $\|\cdot\|_2$-Lipschitz constant of the learning task[1], and $L/2$ is the $\|\cdot\|_1$-Lipschitz constant over the probability simplex.

The above guarantees are tight if we make no further assumptions about the functions (ℓ_t) (Hazan, 2016; Cesa-Bianchi et al., 1997), but they can be significantly improved if the functions have additional special structure that makes the learning task easier. The literature on online learning explores multiple orthogonal dimensions in which tasks may be significantly easier in practice (see 'related work' below). Here, we focus on the following refined data-dependent regret guarantees, which are known to exploit multiple types of easiness at the same time:

$$\text{OCO:} \qquad O\left(\sqrt{V_T^{\boldsymbol{u}} d \log T}\right) \text{ for all } \boldsymbol{u}, \qquad \text{with } V_T^{\boldsymbol{u}} = \sum_{t=1}^{T} \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle^2, \tag{1}$$

$$\text{Experts:} \quad O\left(\sqrt{\mathbb{E}_{\rho(k)}[V_T^k]\,\mathrm{KL}(\rho\|\pi)}\right) \text{ for all } \rho, \qquad \text{with } V_T^k = \sum_{t=1}^{T} \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{e}_k, \boldsymbol{l}_t \rangle^2, \tag{2}$$

where d is the number of parameters and $\mathrm{KL}(\rho\|\pi) = \sum_{k=1}^{K} \rho(k) \ln \rho(k)/\pi(k)$ is the Kullback-Leibler divergence from a fixed prior distribution π over experts to any (data-dependent) comparator distribution ρ; for instance, ρ is allowed here to be a point-mass on the best expert k^* in hindsight, in which case we would have $\mathrm{KL}(\rho\|\pi) = -\ln \pi(k^*)$.

The OCO guarantee is achieved by the METAGRAD algorithm (Van Erven and Koolen, 2016), and implies regret that grows at most logarithmically in T both in case the losses are curved (exp-concave, strongly convex) and in the stochastic case whenever the losses are independent, identically distributed samples with variance controlled by a Bernstein condition (Koolen et al., 2016). The guarantee for the expert case is achieved by the SQUINT algorithm (Koolen and Van Erven, 2015; Koolen, 2015). It simultaneously exploits two types of structures: in many cases the V_T^k term is much smaller than L^2T (Gaillard et al., 2014; Koolen et al., 2016) and the so-called *quantile bound* $\mathrm{KL}(\rho\|\pi)$ is much smaller than the worst case $\ln K$ when multiple experts make good predictions (Chaudhuri et al., 2009; Chernov and Vovk, 2010). SQUINT and METAGRAD are both based on the same technique of tracking the empirical performance of *multiple learning rates* in parallel over quadratic approximations of the original losses. A computational difference though is that SQUINT is able to do this by a continuous integral that can be evaluated in closed form, whereas METAGRAD uses a discrete grid of learning rates.

Unfortunately, to achieve (1) and (2), both METAGRAD and SQUINT need knowledge of the Lipschitz constant (G or L, respectively). Overestimating G or L by a factor of $c > 1$ has the effect of reducing the effective amount of available data by the same factor c, but underestimating the Lipschitz constant is even worse since it can make the methods fail completely. In fact, the ability to adapt to G has been credited (Ward et al., 2018) as one of the main reasons for the practical

1. We slightly abuse terminology here, because the standard definition of a Lipschitz constant requires an upper-bound on the gradient norms for any parameters $\boldsymbol{u}$, not just for $\boldsymbol{u} = \widehat{\boldsymbol{u}}_t$, and may therefore be larger.

success of the AdaGrad algorithm (Duchi et al., 2011; McMahan and Streeter, 2010). Thus getting the Lipschitz constant right makes the difference between having practical algorithms and having promising theoretical results.

For OCO, an important first step towards combining Lipschitz adaptivity to G with regret bounds of the form (1) was taken by Cutkosky and Boahen (2017b), who aimed for (1) but had to settle for a weaker result with $G\sum_{t=1}^{T}\|\boldsymbol{g}_t\|_2\|\widehat{\boldsymbol{u}}_t-\boldsymbol{u}\|_2^2$ instead of $V_T^{\boldsymbol{u}}$. Although not sufficient to adapt to a Bernstein condition, they do provide a series of stochastic examples where their bound already leads to a fast $O(\ln^4 T)$ rates. For the expert setting, Wintenberger (2017) has made significant progress towards a version of (2) without the quantile bound improvement, but he is left with having to specify an initial guess L_{guess} for L that enters as $O(\ln\ln(L/L_{\text{guess}}))$ in his bound, which may yet be arbitrarily large when the initial guess is on the wrong scale.

Main Contributions. Our main contributions are that we complete the process began by Cutkosky and Boahen (2017b) and Wintenberger (2017) by showing that it is indeed possible to achieve (1) and (2) without prior knowledge of G or L. In fact, for the expert setting we are able to adapt to the tighter quantity $B\geq\max_k|\langle\widehat{\boldsymbol{u}}_t-\boldsymbol{e}_k,\boldsymbol{l}_t\rangle|$. We achieve these results by dynamically updating the set of active learning rates in METAGRAD and SQUINT depending on the observed Lipschitz constants. In both cases, we encounter a similar tuning issue as Wintenberger (2017), but we avoid the need to specify any initial guess using a new restarting scheme, which restarts the algorithm when the observed Lipschitz constant increases too much. Interestingly, the scheme and its analysis are different from the well-known doubling trick (Cesa-Bianchi and Lugosi, 2006), and the regret bound is dominated by the regret incurred over the last *two* epochs instead of just the last epoch. Adding up the regret bounds over the last two epochs leads to at most an extra $\sqrt{2}$ factor multiplying the final bound, and so this is the overhead we incur for Lipschitz adaptivity. In addition to these main results, we remove the need to specify the number of rounds T in advance for METAGRAD by adding learning rates as T gets larger, and we improve the computational efficiency of how it handles constraints on the domain of prediction: by a minor extension of the black-box reduction for projections of Cutkosky and Orabona (2018), we incur only the computational cost of projecting on the domain of interest in *Euclidean* distance. This should be contrasted with the usual projections in time-varying Mahalanobis distance for second-order methods like METAGRAD.

Related Work. We build on several lines of work that achieve subsets of Lipschitz, variance and quantile adaptivity. Lipschitz adaptivity in OCO is achieved by OGD with learning rate $\eta_t\propto 1/\sqrt{\sum_{s=1}^{t}\|\boldsymbol{g}_s\|_2^2}$, which leads to $O(\sqrt{\sum_{t=1}^{T}\|\boldsymbol{g}_t\|_2^2})=O(G\sqrt{T})$ regret. This is the approach taken by AdaGrad (for each dimension separately) (Duchi et al., 2011; McMahan and Streeter, 2010). Lipschitz adaptive methods for prediction with expert advice (sometimes called scale-free) were obtained by Cesa-Bianchi et al. (2007) and De Rooij et al. (2014). These include a data-dependent variance term (though different from V_T^k in (2)), but no quantiles.

Dropping Lipschitz adaptivity, we find that bounds with V_T^k from (2) have previously been obtained by Gaillard et al. (2014) and Wintenberger (2014) without quantile bounds. Quantile adaptivity was achieved by Chaudhuri et al. (2009) and Chernov and Vovk (2010) without variance adaptivity, and with a slightly weaker notion of variance by Luo and Schapire (2015). In OCO, the analogue of quantile adaptivity is to adapt to the norm of $\boldsymbol{u}$, which has been achieved in various different ways, see for instance (McMahan and Abernethy, 2013; Cutkosky and Orabona, 2018).

Several other important (and related) criteria of easiness are actively considered in the literature. These include curvature of the loss functions, where earlier results achieve fast rates assuming that

the degree of curvature is known (Hazan et al., 2007), measured online (Bartlett et al., 2007; Do et al., 2009) or entirely unknown (Van Erven and Koolen, 2016; Cutkosky and Orabona, 2018). Fast rates are also possible for slowly-varying linear functions and, more generally, optimistically predictable gradient sequences (Hazan and Kale, 2010; Chiang et al., 2012; Rakhlin and Sridharan, 2013).

We view our results as a step towards developing algorithms that automatically adapt to multiple relevant measures of difficulty at the same time. It is not a given that such combinations are always possible. For example, Cutkosky and Boahen (2017a) show that Lipschitz adaptivity and adapting to the comparator complexity in OCO, although both achievable independently, cannot both be realized at the same time (at least not without further assumptions). A general framework to study which notions of task difficulty do combine into achievable bounds is provided by Foster et al. (2015). Foster et al. (2017) characterize the achievability of general data-dependent regret bounds for domains that are balls in general Banach spaces.

Outline. We add Lipschitz adaptivity to SQUINT for the expert setting in Section 2. Then, in Section 3, we do the same for METAGRAD in the OCO setting. The developments are analogous at a high level but differ in the details for computational reasons. We highlight the differences along the way. Section 3 further describes how to avoid specifying T in advance for METAGRAD. Then, in Section 4, we add efficient projections for METAGRAD, and finally Section 5 concludes with a discussion of directions for future work.

Problem Setting and Notation. In OCO, a learner repeatedly chooses actions $\widehat{\boldsymbol{u}}_t$ from a closed convex set $\mathcal{U} \subseteq \mathbb{R}^d$ during rounds $t = 1, \dots, T$, and suffers losses $\ell_t(\widehat{\boldsymbol{u}}_t)$, where $\ell_t : \mathcal{U} \to \mathbb{R}$ is a convex function. The learner's goal is to achieve small regret $R_T^{\boldsymbol{u}} = \sum_{t=1}^T \ell_t(\widehat{\boldsymbol{u}}_t) - \sum_{t=1}^T \ell_t(\boldsymbol{u})$ with respect to any comparator action $\boldsymbol{u} \in \mathcal{U}$, which measures the difference between the cumulative loss of the learner and the cumulative loss they could have achieved by playing the oracle action $\boldsymbol{u}$ from the start. A special case of OCO is prediction with expert advice, where $\ell_t(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{l}_t \rangle$ for $\boldsymbol{l}_t \in \mathbb{R}^K$ and the domain $\mathcal{U}$ is the probability simplex $\triangle_K = \{(u_1, \dots, u_K) : u_i \geq 0, \sum_i u_i = 1\}$. In this context we will further write $\boldsymbol{p}$ instead of $\boldsymbol{u}$ for the parameters to emphasize that they represent a probability distribution. We further define $[K] = \{1, \dots, K\}$.

2. An Adaptive Second-order Quantile Method for Experts

In this section, we present an extension of the SQUINT algorithm that adapts automatically to the loss range in the setting of prediction with expert advice.

Throughout this section, we denote the *instantaneous regret* of expert $k \in [K]$ in round t by $r_t^k := \langle \widehat{\boldsymbol{p}}_t - \boldsymbol{e}_k, \boldsymbol{l}_t \rangle$, where $\widehat{\boldsymbol{p}}_t \in \triangle_K$ is the weight vector played by the algorithm and $\boldsymbol{l}_t \in \mathbb{R}^K$ is the observed loss vector. The cumulative regret with respect to expert k is given by $R_t^k := \sum_{s=1}^t r_s^k$. The cumulative 'variance' with respect to expert k is measured by $V_t^k := \sum_{s=1}^t v_s^k$ for $v_t^k := (r_t^k)^2$. In the next subsection, we review the SQUINT algorithm.

2.1. The SQUINT Algorithm

We first describe the original SQUINT algorithm as introduced by Koolen and Van Erven (2015). Let π and γ be prior distributions with supports on $k \in [K]$ and $\eta \in \,]0, 1/2]$, respectively. After t

rounds, SQUINT outputs predictions

$$\widehat{\boldsymbol{p}}_{t+1} \propto \mathop{\mathbb{E}}_{\pi(k)\gamma(\eta)} \left[\eta e^{-\sum_{s=1}^{t} f_s(k,\eta)} \boldsymbol{e}_k\right], \tag{3}$$

where $f_t(k,\eta)$ are quadratic *surrogate losses* defined by

$$f_t(k,\eta) \coloneqq -\eta\langle \widehat{\boldsymbol{p}}_t - \boldsymbol{e}_k, \boldsymbol{\ell}_t\rangle + \eta^2 \langle \widehat{\boldsymbol{p}}_t - \boldsymbol{e}_k, \boldsymbol{\ell}_t\rangle^2.$$

Koolen and Van Erven (2015) propose to use the *improper prior* $\gamma(\eta) = 1/\eta$ which does not integrate to a finite value over its domain, but because of the weighting by η in (3) the predictions $\widehat{\boldsymbol{p}}_{t+1}$ are still well-defined. The benefit of the improper prior is that it allows calculating $\widehat{\boldsymbol{p}}_{t+1}$ in closed form (Koolen and Van Erven, 2015). It is also the natural candidate for Lipschitz adaptivity, as it is scale-invariant: the density of an interval only depends on the ratio of its endpoints, not on their location. For any distribution $\rho \in \triangle_K$, SQUINT achieves the following bound:

$$R_T^\rho = O\left(\sqrt{V_T^\rho \left(\mathrm{KL}(\rho||\pi) + \ln\ln T\right)}\right),$$

where $R_T^\rho = \mathbb{E}_{\rho(k)}\left[R_T^k\right]$ and $V_T^\rho = \mathbb{E}_{\rho(k)}\left[V_T^k\right]$. This version of SQUINT assumes the loss range $\max_k \ell_{t,k} - \min_k \ell_{t,k}$ is at most 1, and can fail otherwise. In the next subsection, we present an extension of SQUINT which does not need to know the Lipschitz constant.

2.2. Lipschitz Adaptive SQUINT

We first design a version of SQUINT, called SQUINT+C, that still requires an initial estimate B of the Lipschitz constant. We then present SQUINT+L which tunes this parameter online. For now, we consider a fixed $B > 0$. In addition to this, the algorithm takes a prior distribution $\pi \in \triangle_K$. We denote the observed Lipschitz constant in round t at the algorithm's prediction $\widehat{\boldsymbol{p}}_t$ by $b_t \coloneqq \max_k |r_t^k| = \max_k |\langle \widehat{\boldsymbol{p}}_t - \boldsymbol{e}_k, \boldsymbol{\ell}_t\rangle|$, and denote its running maximum by $B_t \coloneqq B \vee \max_{s\le t} b_s$, with the convention that $B_0 = B$. We will also require a *clipped* version of the loss vector $\bar{\boldsymbol{\ell}}_t = \boldsymbol{\ell}_t \cdot B_{t-1}/B_t$, and denote by $\bar{r}_t^k = \langle \widehat{\boldsymbol{p}}_t - \boldsymbol{e}_k, \bar{\boldsymbol{\ell}}_t\rangle$ the *clipped instantaneous regret*; we will use that $|\bar{r}_t^k| \le B_{t-1}$. Following Cutkosky (2019), it suffices to control the regret for the clipped loss, because the cumulative difference is of the order of one round (*i.e.* a negligible lower-order constant):

$$R_T^k - \bar{R}_T^k \;\coloneqq\; \sum_{t=1}^{T}\left(r_t^k - \bar{r}_t^k\right) \;=\; \sum_{t=1}^{T}(B_t - B_{t-1})\frac{r_t^k}{B_t} \;\le\; B_T - B_0. \tag{4}$$

This means we can focus on the regret for $\bar{\boldsymbol{\ell}}_t$, for which the range bound $|\bar{r}_t^k| \le B_{t-1}$ is available *ahead* of each round t. To motivate SQUINT+C, we define the potential function after T rounds by

$$\Phi_T \coloneqq \sum_k \pi_k \int_0^{\frac{1}{2B_{T-1}}} \frac{e^{\eta\bar{R}_T^k - \eta^2\bar{V}_T^k} - 1}{\eta}\,\mathrm{d}\eta \quad \text{where} \quad \bar{R}_T^k \coloneqq \sum_{t=1}^{T}\bar{r}_t^k \;\text{ and }\; \bar{V}_T^k \coloneqq \sum_{t=1}^{T}(\bar{r}_t^k)^2. \tag{5}$$

We also define $\Phi_0 = 0$ (due to the integrand being zero), even though it involves the meaningless B_{-1} in the upper limit. The algorithm is now derived from the desire of keeping this potential under control. As we will see in the analysis, this requirement uniquely forces the choice of weights

$$\widehat{p}_{T+1}^k \;\propto\; \pi_k \int_0^{\frac{1}{2B_T}} e^{\eta\bar{R}_T^k - \eta^2\bar{V}_T^k}\,\mathrm{d}\eta. \tag{6}$$

Algorithm 1 Restarts to make SQUINT+C or METAGRAD+C scale-free.

Require: ALG is either SQUINT+C or METAGRAD+C, taking as input parameter an initial scale B;

1: Play $\mathbf{0}$ for OCO or π for experts until the first time $t = \tau_1$ that $b_t \neq 0$;
2: Run ALG with input $B = B_{\tau_1}$ until the first time $t = \tau_2$ that $\frac{B_t}{B_{\tau_1}} > \sum_{s=1}^{t} \frac{b_s}{B_s}$;
3: Set $\tau_1 = \tau_2$ and goto line 2;

The predictions $\widehat{\boldsymbol{p}}_{t+1}$ take the same functional form as the original SQUINT, and can hence be evaluated in closed form (*i.e.* in terms of the Gaussian CDF). The regret analysis consists of two parts. First, we show that the algorithm keeps the potential small:

Lemma 1 *Given parameter $B > 0$,* SQUINT+C *ensures* $\Phi_T \leq \ln \frac{B_{T-1}}{B}$.

The next step of the argument is to show that a small potential Φ_T is useful. The argument here follows from (Koolen and Van Erven, 2015), specifically the version by Koolen (2015). We have:

Lemma 2 *For any comparator distribution $\rho \in \triangle_K$ the regret of* SQUINT+C *is at most*

$$\bar{R}_T^\rho \;\leq\; \sqrt{2\bar{V}_T^\rho}\left(1+\sqrt{2C_T^\rho}\right) + 5B_{T-1}\left(C_T^\rho + \ln 2\right), \quad \textit{where}$$
$$C_T^\rho \;\coloneqq\; \mathrm{KL}\left(\rho\|\pi\right) + \ln\left(\Phi_T + \frac{1}{2} + \ln\left(2 + \sum_{t=1}^{T-1}\frac{b_t}{B_t}\right)\right).$$

Keeping only the dominant terms, this reads $\bar{R}_T^\rho = O\left(\sqrt{\bar{V}_T^\rho\left(\mathrm{KL}\left(\rho\|\pi\right) + \ln\left(\Phi_T + \ln T\right)\right)}\right)$. Combining with (4), and Lemmas 1 and 2, we obtain a bound of the form

$$R_T^\rho \;=\; O\left(\sqrt{V_T^\rho\left(\mathrm{KL}\left(\rho\|\pi\right) + \ln\ln\frac{TB_{T-1}}{B}\right)} + 5B_T\left(\mathrm{KL}\left(\rho\|\pi\right) + \ln\ln\frac{TB_{T-1}}{B}\right)\right). \tag{7}$$

However, there does not seem to be any safe a-priori way to tune $B = B_0$. If we set it too small, the factor $\ln\ln(B_{T-1}/B)$ explodes. If we set it too large, with B much larger than the effective range of the data, then $B_T = B$ and the term outside the square-root on the RHS of (7) blows up. It does not appear possible to bypass this tuning dilemma directly within the current construction. Instead, we solve this problem using a new type of restarts that are different from the well-known doubling trick. For this, we present Algorithm 1, which applies to both SQUINT+C and METAGRAD+C (presented in the next section). It monitors a condition on the sequences (b_t) and (B_t) to trigger restarts.

Theorem 3 *Let* SQUINT+L *be the result of applying Algorithm 1 with* SQUINT+C *as* ALG. SQUINT+L *guarantees, for any comparator $\rho \in \triangle_K$,*

$$R_T^\rho \;\leq\; 2\sqrt{V_T^\rho}\left(1+\sqrt{2\Gamma_T^\rho}\right) + 10B_T\left(\Gamma_T^\rho + \ln 2\right) + 4B_T,$$

where $\Gamma_T^\rho \coloneqq \mathrm{KL}\left(\rho\|\pi\right) + \ln\left(\ln\left(\sum_{t=1}^{T-1} b_t/B_t\right) + \ln\left(2 + \sum_{t=1}^{T-1} b_t/B_t\right)\right) + 1/2.$

Note that Γ_T^ρ in Theorem 3 is equal to $\mathrm{KL}\left(\rho\|\pi\right) + O\left(\ln\ln T\right)$. Importantly, this theorem and Algorithm 1 do not depend on any initial guess B anymore. Instead, Algorithm 1 plays the starting parameters until the first time a non-zero loss is observed, and then monitors a data-dependent criterion that measures whether the loss range has increased by more than a factor that is roughly t, to decide when to trigger a restart. For most types of data, such large increases in the loss range should be rare after a few start-up rounds, so restarts should quickly stop occurring.

3. An Adaptive Method for Online Convex Optimization

We now present an extension of the METAGRAD algorithm which adapts automatically to the gradient norm in online convex optimization — we call this Lipschitz adaptive version METAGRAD+L. Recall that in the OCO setting, at each round t, the learner predicts a vector $\widehat{\boldsymbol{u}}_t$ in a closed convex set $\mathcal{U} \subset \mathbb{R}^d$, then suffers loss $\ell_t(\widehat{\boldsymbol{u}}_t)$, where $\ell_t : \mathcal{U} \to \mathbb{R}$ is a convex function. The goal of the learner is to minimize the regret $R_T^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \ell_t(\widehat{\boldsymbol{u}}_t) - \sum_{t=1}^T \ell_t(\boldsymbol{u})$ with respect to the single best action $\boldsymbol{u} \in \mathcal{U}$ in hindsight. In this case, convexity of the losses implies that $\ell_t(\widehat{\boldsymbol{u}}_t) - \ell_t(\boldsymbol{u}) \leq \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle$, where $\boldsymbol{g}_t \coloneqq \nabla \ell_t(\widehat{\boldsymbol{u}}_t)$, and so it suffices to control the *pseudo-regret* $\tilde{R}_T^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle$. We will assume that the set $\mathcal{U}$ is bounded, and denote its diameter by

$$D \coloneqq \sup_{\boldsymbol{u},\boldsymbol{v} \in \mathcal{U}} \|\boldsymbol{u} - \boldsymbol{v}\|_2. \tag{8}$$

Without loss of generality, we will also assume that the set $\mathcal{U}$ is centered at $\mathbf{0}$. The proofs for this section are deferred to Appendix B. We now review the METAGRAD algorithm.

3.1. The METAGRAD Algorithm

The METAGRAD algorithm runs several sub-algorithms at each round: namely, a set of slave algorithms, which learn the best action in $\mathcal{U}$ given a learning rate η in some pre-defined grid $\mathcal{G}$, and a master algorithm, which learns the best learning rate. Through this, the METAGRAD algorithm controls the sum of *surrogate losses* $\sum_{t=1}^T f_t(\boldsymbol{u}, \eta)$ over all $\eta \in \mathcal{G}$ and $\boldsymbol{u} \in \mathcal{U}$ simultaneously, where

$$f_t(\boldsymbol{u}, \eta) \coloneqq -\eta \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle + \eta^2 \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle^2, \tag{9}$$

and $\widehat{\boldsymbol{u}}_t$ is the master's prediction at round $t \in [T]$. Each slave algorithm takes as input a learning rate from a finite grid $\mathcal{G}$ (with $\lceil 1/2 \log_2 T \rceil$ points) in the form of a geometric progression and within the interval $[1/(5DG\sqrt{T}), 1/(5DG)]$, where G is an upper-bound on the norms of the gradients. In this case, G must be known in advance to construct the grid; in the proof of METAGRAD's regret bound, it is crucial for the learning rates to be in the right interval in order to invoke a certain Gaussian exp-concavity result due to Van Erven and Koolen (2016) for the surrogate losses in (9). In what follows, we let $\mathbf{S}_t \coloneqq \sum_{s=1}^t \boldsymbol{g}_s \boldsymbol{g}_s^\intercal$, for $t \geq 0$.

Slaves' Predictions. Each slave $\eta \in \mathcal{G}$ starts with $\widehat{\boldsymbol{u}}_1^\eta = \mathbf{0} \in \mathcal{U}$, and at the end of round $t \geq 1$, it receives the master's prediction $\widehat{\boldsymbol{u}}_t$ and updates its own prediction in two steps:

$$\boldsymbol{u}_{t+1}^\eta \coloneqq \widehat{\boldsymbol{u}}_t^\eta - \eta \boldsymbol{\Sigma}_{t+1}^\eta \boldsymbol{g}_t \left(1 + 2\eta \left(\widehat{\boldsymbol{u}}_t^\eta - \widehat{\boldsymbol{u}}_t\right)^\intercal \boldsymbol{g}_t\right), \text{ where } \boldsymbol{\Sigma}_{t+1}^\eta \coloneqq \left(\tfrac{\mathbf{I}}{D^2} + 2\eta^2 \mathbf{S}_t\right)^{-1}, \tag{10}$$
$$\text{and } \widehat{\boldsymbol{u}}_{t+1}^\eta = \operatorname*{argmin}_{\boldsymbol{u} \in \mathcal{U}} \left(\boldsymbol{u}_{t+1}^\eta - \boldsymbol{u}\right)^\intercal \left(\boldsymbol{\Sigma}_{t+1}^\eta\right)^{-1} \left(\boldsymbol{u}_{t+1}^\eta - \boldsymbol{u}\right).$$

Master's Predictions. After receiving the slaves' predictions, $(\widehat{\boldsymbol{u}}_t^\eta)_{\eta\in\mathcal{G}}$, at round $t \geq 1$, the master algorithm aggregates them and outputs $\widehat{\boldsymbol{u}}_t \in \mathcal{U}$ according to:

$$\widehat{\boldsymbol{u}}_t \coloneqq \frac{\sum_{\eta\in\mathcal{G}} \eta w_t^\eta \widehat{\boldsymbol{u}}_t^\eta}{\sum_{\eta\in\mathcal{G}} \eta w_t^\eta}; \quad w_t^\eta \coloneqq e^{-\sum_{s=1}^{t-1} f_s(\widehat{\boldsymbol{u}}_s^\eta,\eta)}.$$

Van Erven and Koolen (2016) showed that METAGRAD has regret bounded by (1). In the next subsection, we present an extension of METAGRAD which does not require knowledge of either the horizon T or the Lipschitz constant (*i.e.* a bound on the norms of the gradients).

3.2. Lipschitz Adaptive METAGRAD

Similar to the SQUINT case, we first design a version of METAGRAD, called METAGRAD+C, which still requires an input $B > 0$ (in this case, B/D is the initial estimate of the Lipschitz bound). We then present METAGRAD+L which sets this parameter online. For now, we consider a fixed $B > 0$. We define $b_t \coloneqq D\|\nabla \ell_t(\widehat{\boldsymbol{u}}_t)\|_2 = D\|\boldsymbol{g}_t\|_2$, for $t \geq 1$, and $b_0 \coloneqq B$. We denote the running maximum of (b_t) by $B_t \coloneqq \max_{0\leq s\leq t} b_s$. We will also require a *clipped* version of the gradient vector $\bar{\boldsymbol{g}}_t = \boldsymbol{g}_t \cdot B_{t-1}/B_t$, and denote by $\bar{r}_t^{\boldsymbol{u}} = \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \bar{\boldsymbol{g}}_t\rangle$ the *clipped instantaneous pseudo-regret* with respect to $\boldsymbol{u} \in \mathcal{U}$. In addition, it will be useful to define

$$\bar{f}_t(\boldsymbol{u},\eta) \coloneqq -\eta\bar{r}_t^{\boldsymbol{u}} + (\eta\bar{r}_t^{\boldsymbol{u}})^2 \quad \text{and} \quad \bar{\mathbf{S}}_t \coloneqq \sum_{s=1}^{t} \bar{\boldsymbol{g}}_s\bar{\boldsymbol{g}}_s^\intercal. \tag{11}$$

Recall that in the original METAGRAD, the horizon T and the Lipschitz constant G were required to construct the grid of learning rates. We circumvent this by defining an infinite grid $\mathcal{G}$ in which, at any given round $t \geq 1$, only a finite number of (active) slaves — up to $\log_2 t$ many — output a non-zero prediction. Each slave η in this grid receives a prior weight $\pi(\eta) \in [0,1]$, where $\sum_{\eta\in\mathcal{G}} \pi(\eta) = 1$. Given input $B > 0$ to METAGRAD+C, the grid $\mathcal{G}$ and the prior π are defined by

$$\mathcal{G} \coloneqq \left\{\eta_i \coloneqq \frac{1}{5B2^i} : i \in \mathbb{N}\cup\{0\}\right\}; \quad \pi(\eta_i) \coloneqq \frac{1}{(i+1)(i+2)},\ i \in \mathbb{N}\cup\{0\}. \tag{12}$$

The subset of active slaves $\mathcal{A}_t$ at a round $t \geq 1$ is given by

$$\mathcal{A}_t \coloneqq \left\{\eta \in \mathcal{G} \cap \left[0, \tfrac{1}{5B_{t-1}}\right] : s_\eta < t\right\}, \text{ with } s_\eta \coloneqq \min\left\{t \geq 0 : \frac{1}{\eta} \leq D\sum_{s=1}^{t}\|\bar{\boldsymbol{g}}_s\|_2 + B_t\right\}. \tag{13}$$

We note that restricting the slaves (or learning rates) to the set $\mathcal{G}_t \coloneqq \mathcal{G} \cap [0, 1/(5B_{t-1})]$ is similar in principle to clipping the upper integral range in the SQUINT+C case.

Slaves' Predictions. A slave $\eta \in \mathcal{G} \cap [0, 1/(5B_{t-1})]$ issues predictions $\widehat{\boldsymbol{u}}_t^\eta = \mathbf{0}$ in all rounds $t \leq s_\eta + 1$. From then on (*i.e.* at the end of round $t \geq s_\eta + 1$), it receives the master's prediction $\widehat{\boldsymbol{u}}_t$ as input and updates its own prediction in two steps:

$$\boldsymbol{u}_{t+1}^\eta \coloneqq \widehat{\boldsymbol{u}}_t^\eta - \eta\boldsymbol{\Sigma}_{t+1}^\eta\bar{\boldsymbol{g}}_t\left(1 + 2\eta\left(\widehat{\boldsymbol{u}}_t^\eta - \widehat{\boldsymbol{u}}_t\right)^\intercal\bar{\boldsymbol{g}}_t\right), \text{ where } \boldsymbol{\Sigma}_{t+1}^\eta \coloneqq \left(\tfrac{\mathbf{I}}{D^2} + 2\eta^2\left(\bar{\mathbf{S}}_t - \bar{\mathbf{S}}_{s_\eta}\right)\right)^{-1},$$
$$\text{and } \widehat{\boldsymbol{u}}_{t+1}^\eta = \operatorname*{argmin}_{\boldsymbol{u}\in\mathcal{U}} \left(\boldsymbol{u}_{t+1}^\eta - \boldsymbol{u}\right)^\intercal \left(\boldsymbol{\Sigma}_{t+1}^\eta\right)^{-1}\left(\boldsymbol{u}_{t+1}^\eta - \boldsymbol{u}\right).$$

Master's Predictions. At each round $t \geq 1$, the master algorithm receives the slaves' predictions $(\widehat{\boldsymbol{u}}_t^\eta)_{t \in \mathcal{A}_t}$ and outputs

$$\widehat{\boldsymbol{u}}_t = \frac{\sum_{\eta \in \mathcal{A}_t} \eta w_t^\eta \widehat{\boldsymbol{u}}_t^\eta}{\sum_{\eta \in \mathcal{A}_t} \eta w_t^\eta}, \quad \text{where } \; w_t^\eta \coloneqq \pi(\eta) e^{-\sum_{s=s_\eta+1}^{t-1} \bar{f}_s(\widehat{\boldsymbol{u}}_s^\eta, \eta)}. \tag{14}$$

Remark 4 (Number of Active Slaves) *At any round $t \geq 1$, the number of active slaves is at most $\lceil \log_2 t \rceil$. In fact, if $\eta \in \mathcal{A}_t$, then by definition $\eta \geq 1/(D\sum_{s=1}^{s_\eta} \|\boldsymbol{g}_s\|_2 + B_{s_\eta}) \geq 1/(tB_{t-1})$ (since $s_\eta \leq t-1$), and thus $\mathcal{A}_t \subset [1/(tB_{t-1}), 1/(5B_{t-1})]$. Since $\mathcal{A}_t$ is a grid in the form of a geometric progression with common ratio 2, there are at most $\lceil \log_2 t \rceil$ slaves in $\mathcal{A}_t$.*

To motivate METAGRAD+C, we define the potential function after $t \geq 0$ rounds by

$$\Phi_t \coloneqq \pi(\mathcal{G}_t \setminus \mathcal{A}_t) + \sum_{\eta \in \mathcal{A}_t} \pi(\eta) e^{-\sum_{s=s_\eta+1}^{t} \bar{f}_s(\widehat{\boldsymbol{u}}_s^\eta, \eta)}, \quad \text{where } \mathcal{G}_t \coloneqq \mathcal{G} \cap \left[0, \tfrac{1}{5B_{t-1}}\right]. \tag{15}$$

Let $\boldsymbol{u} \in \mathcal{U}$. Recall that the pseudo-regret is defined by $\tilde{R}_T^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle$. We now defined its *clipped* version by $\bar{R}_T^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \bar{\boldsymbol{g}}_t \rangle$. For $r_t^{\boldsymbol{u}} \coloneqq \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle$, we have, similarily to (4),

$$\tilde{R}_T^{\boldsymbol{u}} - \bar{R}_T^{\boldsymbol{u}} = \sum_{t=1}^T (r_t^{\boldsymbol{u}} - \bar{r}_t^{\boldsymbol{u}}) \;=\; \sum_{t=1}^T (B_t - B_{t-1}) \frac{r_t^{\boldsymbol{u}}}{B_t} \;\leq\; B_T - B_0, \tag{16}$$

where the last inequality follows from the Cauchy-Schwarz inequality and the fact that $\mathcal{U}$ has diameter D, which together imply that $|r_t^{\boldsymbol{u}}| \leq B_t$. Using the inequality $e^{x-x^2} - 1 \leq x$, which holds for all $x \geq -1/2$, one can shown that the potential is a decreasing function of the number of rounds:

Lemma 5 METAGRAD+C *guarantees that $\Phi_T \leq \cdots \leq \Phi_0 = 1$, for all $T \in \mathbb{N}$.*

We now give an upper-bound on $\bar{R}_T^{\boldsymbol{u}}$ in terms of the clipped 'variance' $\bar{V}_T^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T (\bar{r}_t^{\boldsymbol{u}})^2$;

Theorem 6 *Given input $B > 0$, the clipped pseudo-regret for* METAGRAD+C *is bounded by*

$$\bar{R}_T^{\boldsymbol{u}} \leq 3\sqrt{\bar{V}_T^{\boldsymbol{u}} C_T} + 15 B_T C_T, \quad \textit{for any } \boldsymbol{u} \in \mathcal{U},$$

where $C_T \coloneqq d \ln\left(1 + \frac{2\sum_{t=0}^{T-1} b_t^2}{25 d B_{T-1}^2}\right) + 2\ln\left(\log_2^+ \frac{\sqrt{\sum_{t=1}^T b_t^2}}{B} + 3\right) + 2$ and $\log_2^+ = 0 \vee \log_2$.

Remark 7 *For $\boldsymbol{u} \in \mathcal{U}$, we can relate the clipped pseudo-regret to the ordinary regret via $R_T^{\boldsymbol{u}} \leq \tilde{R}_T^{\boldsymbol{u}} \leq \bar{R}_T^{\boldsymbol{u}} + B_T$ (see (16)) and on the right-hand side we can also use that $\bar{V}_T^{\boldsymbol{u}} \leq V_T^{\boldsymbol{u}}$.*

An important aspect to note from Theorem 6 is that the ratio $\sqrt{\sum_{t=1}^T b_t^2}/B$, could in principle be arbitrarily large if the input B is too small compared to the actual norms of the gradients (for SQUINT it was the ratio B_{T-1}/B which was problematic). To resolve this issue, we use the same restart approach as in the SQUINT case:

Theorem 8 *Let* METAGRAD+L *be the result of applying Algorithm 1 to* METAGRAD+C. *Then the actual and linearised regrets for* METAGRAD+L *are both bounded by*

$$R_t^{\boldsymbol{u}} \;\le\; \tilde{R}_T^{\boldsymbol{u}} \;\le\; 3\sqrt{V_T^{\boldsymbol{u}}\Gamma_T} + 15B_T\Gamma_T + 4B_T \quad \textit{for all } \boldsymbol{u} \in \mathcal{U},$$

where $\Gamma_T \coloneqq 2d \ln\left(1 + \frac{2}{25d}\sum_{t=1}^T \frac{b_t^2}{B_t^2}\right) + 4\ln\left(\log_2^+ \sqrt{\sum_{t=1}^T (\sum_{s=1}^t \frac{b_s}{B_s})^2} + 3\right) + 4 = O(d \ln T)$.

Theorem 8 replaces the ratio $\sqrt{\sum_{t=1}^T b_t^2}/B$ appearing in the (clipped) pseudo-regret bound of METAGRAD+C by $\sigma_T \coloneqq \sqrt{\sum_{t=1}^T (\sum_{s=1}^t b_s/B_s)^2}$. The latter is independent of the input B and is always smaller than $T^{3/2}$; this is perfectly affordable since σ_T appears inside a $\ln\ln$. Our reason for including the linearised regret $\tilde{R}_T^{\boldsymbol{u}}$ in Theorem 8 is that a bound on it in terms of $V_T^{\boldsymbol{u}}$ is the precondition for fast rate results in individual-sequence settings based on curvature (Van Erven and Koolen, 2016) and in statistical settings under certain (Bernstein type) conditions (Koolen et al., 2016).

4. Efficient Implementation Through a Reduction to the Ball

Using METAGRAD (+C or +L), the computation of each slave prediction $\widehat{\boldsymbol{u}}_t^\eta$ requires a projection onto an arbitrary convex set $\mathcal{U}$ in *Mahalanobis distance*. Numerically, this typically requires $O(d^p)$ floating point operations (flops), for some $p \in \mathbb{N}$ which depends on the geometry of the set $\mathcal{U}$. Since p can be large in many applications, evaluating $\widehat{\boldsymbol{u}}_t^\eta$ for each slave η can become computationally prohibitive, especially when the number of slaves grows with T; for the METAGRAD versions discussed in this paper, there can be up to $\lceil \log_2 T \rceil$ slaves at round $T \ge 1$ (see Remark 4).

The goal of this section is to streamline these computations, which we will do in two steps. In Section 4.1, we will describe an efficient implementation of METAGRAD on the ball. The main idea here is that the Mahalanobis projections onto the ball, which are performed by the slaves, can reuse a common matrix decomposition. In Section 4.2, we will then obtain an algorithm for any bounded convex set $\mathcal{U}$ by applying the black-box reduction of Cutkosky and Orabona (2018) to METAGRAD on the ball enclosing $\mathcal{U}$. We show (Theorem 10) that the reduction also transports variance bounds. The techniques discussed here also apply to the versions of METAGRAD presented in the previous section. However, to simplify the presentation, we will only focus on the original METAGRAD. The proofs for this section are deferred to Appendix C.

4.1. Efficient Implementation of METAGRAD on the Ball

Suppose that $\mathcal{U}$ is the ball of diameter D: $\mathcal{U} = \mathcal{B}_D \coloneqq \{\boldsymbol{u} \in \mathbb{R}^d \colon \|\boldsymbol{u}\|_2 \le D/2\}$. To compute the slave's prediction $\widehat{\boldsymbol{u}}_{t+1}^\eta$, the following quadratic program needs to be solved for each η:

$$\widehat{\boldsymbol{u}}_{t+1}^\eta = \operatorname*{argmin}_{\boldsymbol{u} \in \mathcal{U}} \left(\boldsymbol{u}_{t+1}^\eta - \boldsymbol{u}\right)^\intercal \left(\boldsymbol{\Sigma}_{t+1}^\eta\right)^{-1} \left(\boldsymbol{u}_{t+1}^\eta - \boldsymbol{u}\right), \tag{17}$$

where $\boldsymbol{u}_{t+1}^\eta$ (the unprojected prediction) and $\boldsymbol{\Sigma}_{t+1}^\eta = (\mathbf{I}/D^2 + 2\eta^2 \mathbf{S}_t)^{-1}$ (the co-variance matrix) are defined in (10). Since $\mathcal{U}$ is a ball and $\boldsymbol{\Sigma}_{t+1}^\eta$ is symmetric positive-definite, (17) can be solved in $O(d^3)$ by performing a singular value decomposition of $\boldsymbol{\Sigma}_{t+1}^\eta$. Instead of doing this singular value decomposition separately for each η, we can be a little more efficient by doing a singular value decomposition of $\mathbf{S}_t$ once and then using the following lemma:

Algorithm 2 Reducing an OCO problem on $\mathcal{U} \subset R^d$ to one on a ball.

Require: A bounded convex set $\mathcal{U} \subset \mathbb{R}^d$ with diameter $D > 0$, a Lipschitz bound $G > 0$.
We write METAGRAD(D) for METAGRAD applied to the ball $\mathcal{B}_D$ enclosing $\mathcal{U}$.
for $t = 1$ **to** T **do**
Get $\widehat{\boldsymbol{u}}_t$ from METAGRAD(D) ; `//The initial input to METAGRAD is` $B = DG$.
Predict $\widehat{\boldsymbol{w}}_t = \Pi_{\mathcal{U}}(\widehat{\boldsymbol{u}}_t)$ and receive $\mathring{\boldsymbol{g}}_t = \nabla \ell_t(\widehat{\boldsymbol{w}}_t)$;
Set $\boldsymbol{g}_t \in \frac{1}{2}\left(\mathring{\boldsymbol{g}}_t + \|\mathring{\boldsymbol{g}}_t\| \partial \, \mathrm{d}_{\mathcal{U}}(\widehat{\boldsymbol{u}}_t)\right)$;
Send $\boldsymbol{g}_t$ to METAGRAD(D) ;
end for

Lemma 9 *Let $\boldsymbol{\Lambda}_t \coloneqq \operatorname{diag}((\lambda_t^i)_{i\in[d]})$ and $\mathbf{Q}_t$ be the matrices of eigenvalues and eigenvectors of $\mathbf{S}_t$, respectively, such that $\mathbf{Q}_t \mathbf{S}_t \mathbf{Q}_t^\intercal = \boldsymbol{\Lambda}_t$ and $\mathbf{Q}_t \mathbf{Q}_t^\intercal = \mathbf{I}$.*[2] *Then the solution of (17) is*

$$\widehat{\boldsymbol{u}}_{t+1}^{\eta} = \begin{cases} \boldsymbol{u}_{t+1}^{\eta}, & \text{if } \boldsymbol{u}_{t+1}^{\eta} \in \mathcal{U}, \\ \mathbf{Q}_t^\intercal (x_t^{\eta} \mathbf{I} + 2\eta^2 \boldsymbol{\Lambda}_t)^{-1} \mathbf{Q}_t \boldsymbol{v}_{t+1}^{\eta}, & \text{otherwise,} \end{cases}$$

where $\boldsymbol{v}_{t+1}^{\eta} \coloneqq \left(\mathbf{I}/D^2 + 2\eta^2 \mathbf{S}_t\right) \boldsymbol{u}_{t+1}^{\eta}$ and the scalar x_t^{η} is the unique solution of

$$\rho_t^{\eta}(x) \coloneqq \sum_{i=1}^{d} \frac{\langle \boldsymbol{e}_i, \mathbf{Q}_t \boldsymbol{v}_{t+1}^{\eta} \rangle^2}{(x + 2\eta^2 \lambda_t^i)^2} = \frac{D^2}{4}. \tag{18}$$

Since ρ_t^{η} in (18) is strictly convex and decreasing, $\rho_t^{\eta}(x) = D^2/4$ can be solved using Newton's method in linear time.

A further improvement leverages the rank-one update $\mathbf{S}_t = \mathbf{S}_{t-1} + \boldsymbol{g}_t \boldsymbol{g}_t^\intercal$ to update $\boldsymbol{\Lambda}_{t-1}$ and $\mathbf{Q}_{t-1}$. It is possible to compute the new matrices $\boldsymbol{\Lambda}_t$ and $\mathbf{Q}_t$ in, respectively, $O(d^2)$ and $O(d^3)$ flops, where the latter cost for computing $\mathbf{Q}_t$ is only due to matrix multiplication (rather than a full singular value decomposition) (Bunch et al., 1978), and thus admits an efficient parallel implementation.

4.2. A Reduction to the Ball

In this subsection, we extend the black-box technique of Cutkosky and Orabona (2018) to reduce an OCO problem on an arbitrary bounded convex set $\mathcal{U} \subset \mathbb{R}^d$ to one on a ball, where the implementation of METAGRAD from the previous subsection can be applied.

Let D be the diameter of a closed bounded convex set $\mathcal{U} \subset \mathbb{R}^d$ as in (8), so that the ball $\mathcal{B}_D$ of radius $D/2$ encloses $\mathcal{U}$. As in the previous section, we again assume, without loss of generality, that $\mathcal{U}$ is centered at $\mathbf{0}$. For $\boldsymbol{u} \in \mathcal{U}$, we denote $\mathrm{d}_{\mathcal{U}}(\boldsymbol{u}) = \min_{\boldsymbol{w}\in\mathcal{U}} \|\boldsymbol{u} - \boldsymbol{w}\|_2$ the *distance function* from the set $\mathcal{U}$, and we define $\Pi_{\mathcal{U}}(\boldsymbol{u}) \coloneqq \{\boldsymbol{w} \in \mathcal{U} : \|\boldsymbol{w} - \boldsymbol{u}\|_2 = \mathrm{d}_{\mathcal{U}}(\boldsymbol{u})\}$. Algorithm 2 reduces the OCO problem on the set $\mathcal{U}$ to one on the ball $\mathcal{B}_D$, where the METAGRAD algorithm is used as a black-box to solve it. We note that Algorithm 2 (including its METAGRAD subroutine) only performs a single projection (applied to the output of METAGRAD) onto the set $\mathcal{U}$ in *Euclidean distance* — as opposed the *time-varying Mahalanobis distance* (17); the METAGRAD subroutine only performs projections onto the ball $\mathcal{B}_D$, which can be done efficiently as described in the previous subsection.

2. The existence of such a $\mathbf{Q}_t$ and $\boldsymbol{\Lambda}_t$ is guaranteed due to $\mathbf{S}_t$ being symmetric positive-definite.

In the next theorem, we assume that a Lipschitz bound $G > 0$ is known in advance[3], and we let $\mathring{R}_T^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \langle \widehat{\boldsymbol{w}}_t - \boldsymbol{u}, \mathring{\boldsymbol{g}}_t \rangle$ and $\mathring{V}_T^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \langle \widehat{\boldsymbol{w}}_t - \boldsymbol{u}, \mathring{\boldsymbol{g}}_t \rangle^2$ be the pseudo-regret and 'variance' corresponding to Algorithm 2. We now show that the (pseudo) regret guarantee of METAGRAD readily transfers to Algorithm 2 with almost no overhead:

Theorem 10 *Let $D > 0$, and suppose that the* METAGRAD*(D) subroutine of Algorithm 2 achieves a pseudo-regret bound of the form*

$$\tilde{R}_T^{\boldsymbol{u}} \leq \sqrt{V_T^{\boldsymbol{u}} \Gamma_T} + B\Gamma_T, \text{ for all } \boldsymbol{u} \in \mathcal{B}_D,$$

where $\tilde{R}_t^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle$, $V_t^{\boldsymbol{u}} \coloneqq \sum_{t=1}^T \langle \widehat{\boldsymbol{u}}_t - \boldsymbol{u}, \boldsymbol{g}_t \rangle^2$, and $\Gamma_T = O(d \ln(T/d))$. Then, Algorithm 2 guarantees:

$$\sum_{t=1}^T (\ell_t(\widehat{\boldsymbol{w}}_t) - \ell_t(\boldsymbol{u})) \leq \mathring{R}_T^{\boldsymbol{u}} \leq \sqrt{\mathring{V}_T^{\boldsymbol{u}} \Gamma_T} + 4B\Gamma_T, \text{ for all } \boldsymbol{u} \in \mathcal{U}.$$

From the standard black-box reduction of Cutkosky and Orabona (2018), we would obtain an unsatisfactory result in which $\mathring{V}_T^{\boldsymbol{u}}$ would be measured in terms of the fake gradients $\boldsymbol{g}_t$ that are supplied internally to METAGRAD(D) instead of the actual gradients $\mathring{\boldsymbol{g}}_t$. As this would not be sufficient to adapt to the easiness conditions described in the introduction, the proof of Theorem 10 involves an extra step to relate the variance term back to the actual gradients.

5. Conclusion

We present algorithms that adapt to the Lipschitz constant of the loss for OCO and experts, with hardly any overhead in terms of regret or computation compared to their previous counterparts that had to know the Lipschitz constant up-front. This fits into a larger picture of understanding which types of adaptivity are possible at which price in terms of additional regret and additional run time.

One surprising conclusion from our work is the following observation: for OCO, Cutkosky and Boahen (2017a) show that in general it is not possible to be adaptive to both the Lipschitz constant and the norm of the comparator $\|\boldsymbol{u}\|$ at the same time. Since the analogue of $\|\boldsymbol{u}\|$ in the expert setting is the complexity measure $\mathrm{KL}(\rho\|\pi)$, we might therefore conjecture that Lipschitz adaptivity would also be incompatible with a quantile regret bound in terms of $\mathrm{KL}(\rho\|\pi)$. However, our results show this conjecture to be false: for experts there is no conflict. This holds even in cases where the prior π is not uniform, and our results can easily be extended to a countably infinite number of experts where $\mathrm{KL}(\rho\|\pi)$ cannot even be uniformly bounded.

A final and very interesting question is when is it possible to exploit scenarios with large Lipschitz constants or loss ranges that occur only very infrequently. An example of this is found in statistical learning with heavy-tailed loss distributions. For such scenarios, martingale methods (related to our potential functions) suggest that it may be necessary to replace in $f_t(\boldsymbol{u}, \eta)$ the 'surrogate' negative quadratic term that our algorithms include in the exponent by another function appropriate for the specific distribution (Howard et al., 2018, Table 3). It is not currently clear what individual sequence analogues can be obtained.

3. If one uses METAGRAD+C or METAGRAD+L as the subroutine in Algorithm 2 instead of METAGRAD, then a Lipschitz bound need not be known in adavance; a version of Theorem 10 with different constants would still hold in this case.

Acknowledgments

We thank the anonymous reviewers for feedback that improved the presentation. Part of this work was performed while Zakaria Mhammedi was conducting an internship at the Centrum Wiskunde & Informatica (CWI). This work was also supported by the Australian Research Council and Data61.

References

Peter L. Bartlett, Elad Hazan, and Alexander Rakhlin. Adaptive online gradient descent. In *Advances in Neural Information Processing Systems 20 (NIPS)*, pages 65–72, 2007.

James R. Bunch, Christopher P. Nielsen, and Danny C. Sorensen. Rank-one modification of the symmetric eigenproblem. *Numerische Mathematik*, 31(1):31–48, 1978.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

Nicolò Cesa-Bianchi, Yoav Freund, David Haussler, David P. Helmbold, Robert E. Schapire, and Manfred K. Warmuth. How to use expert advice. *J. ACM*, 44(3):427–485, 1997.

Nicolò Cesa-Bianchi, Yishay Mansour, and Gilles Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2-3):321–352, 2007.

Kamalika Chaudhuri, Yoav Freund, and Daniel Hsu. A parameter-free hedging algorithm. In *Advances in Neural Information Processing Systems 22 (NIPS 2009)*, pages 297–305, 2009.

Alexey V. Chernov and Vladimir Vovk. Prediction with advice of unknown number of experts. In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence*, pages 117–125, 2010.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Le, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Proc. of the 25th Annual Conference on Learning Theory (COLT)*, pages 6.1–6.20, 2012.

Ashok Cutkosky. Artificial constraints and lipschitz hints for unconstrained online learning. *arXiv preprint arXiv:1902.09013*, 2019.

Ashok Cutkosky and Kwabena Boahen. Online learning without prior information. In *Proceedings of the 30th Annual Conference on Learning Theory (COLT)*, pages 643–677, 2017a.

Ashok Cutkosky and Kwabena A. Boahen. Stochastic and adversarial online learning without hyperparameters. In *Advances in Neural Information Processing Systems*, pages 5059–5067, 2017b.

Ashok Cutkosky and Francesco Orabona. Black-box reductions for parameter-free online learning in Banach spaces. In *Conference On Learning Theory, COLT 2018, Stockholm, Sweden, 6-9 July 2018.*, pages 1493–1529, 2018.

Chuong B. Do, Quoc V. Le, and Chuan-Sheng Foo. Proximal regularization for online and batch learning. In *Proc. of the 26th Annual International Conference on Machine Learning (ICML)*, pages 257–264, 2009.

John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12:2121–2159, 2011.

Tim van Erven and Wouter M. Koolen. Metagrad: Multiple learning rates in online learning. In *Advances in Neural Information Processing Systems 29 (NIPS)*, pages 3666–3674, 2016.

Dylan J Foster, Alexander Rakhlin, and Karthik Sridharan. Adaptive online learning. In *Advances in Neural Information Processing Systems 28 (NIPS)*, pages 3375–3383, 2015.

Dylan J. Foster, Alexander Rakhlin, and Karthik Sridharan. Zigzag: A new approach to adaptive online learning. In *Proc. of the 2017 Annual Conference on Learning Theory (COLT)*, pages 876–924, 2017.

Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 1997.

Pierre Gaillard, Gilles Stoltz, and Tim van Erven. A second-order bound with excess losses. In *Proc. of the 27th Annual Conference on Learning Theory (COLT)*, pages 176–196, 2014.

Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2 (3–4):157–325, 2016.

Elad Hazan and Satyen Kale. Extracting certainty from uncertainty: Regret bounded by variation in costs. *Machine learning*, 80(2-3):165–188, 2010.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

Steve Howard, Aaditya Ramdas, John McAuliffe, and Jasjeet Sekhon. Exponential line-crossing inequalities. *ArXiv e-prints*, August 2018.

Wouter M. Koolen. The relative entropy bound for Squint, August 2015. Adversarial Intelligence blog.

Wouter M. Koolen and Tim van Erven. Second-order quantile methods for experts and combinatorial games. In *Conference on Learning Theory*, pages 1155–1175, 2015.

Wouter M. Koolen, Peter D. Grünwald, and Tim van Erven. Combining adversarial guarantees and stochastic fast rates in online learning. In *Advances in Neural Information Processing Systems*, pages 4457–4465, 2016.

Haipeng Luo and Robert E. Schapire. Achieving all with no parameters: Adaptive NormalHedge. In *Proc. of The 28th Annual Conference on Learning Theory (COLT)*, pages 1286–1304, 2015.

Brendan McMahan and Jacob Abernethy. Minimax optimal algorithms for unconstrained linear optimization. In *Advances in Neural Information Processing Systems 26 (NIPS)*, pages 2724–2732, 2013.

H. Brendan McMahan and Matthew J. Streeter. Adaptive bound optimization for online convex optimization. In *Proceedings of the 23rd Annual Conference on Learning Theory (COLT)*, pages 244–256, 2010.

Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *Proc. of the 26th Annual Conference on Learning Theory (COLT)*, pages 993–1019, 2013.

Steven de Rooij, Tim van Erven, Peter D. Grünwald, and Wouter M. Koolen. Follow the leader if you can, hedge if you must. *Journal of Machine Learning Research*, 15(1):1281–1316, 2014.

Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011.

Rachel Ward, Xiaoxia Wu, and Léon Bottou. Adagrad stepsizes: Sharp convergence over nonconvex landscapes, from any initialization. *ArXiv:1806.01811 preprint*, 2018.

Olivier Wintenberger. Optimal learning with Bernstein online aggregation. *ArXiv:1404.1356v2 preprint*, 2014.

Olivier Wintenberger. Optimal learning with Bernstein online aggregation. *Machine Learning*, 106 (1):119–141, 2017.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th Annual International Conference on Machine Learning (ICML)*, pages 928–936, 2003.

Appendix A. Proofs of Section 2

Proof of Lemma 1 We proceed by induction on T. By definition $\Phi_0 = 0$. For $T \geq 0$, the definition (5) gives

$$\Phi_{T+1} \;=\; \underbrace{\sum_k \pi_k \int_0^{\frac{1}{2B_T}} \frac{e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k} \left(e^{\eta \bar{r}_{T+1}^k - \eta^2 (\bar{r}_{T+1}^k)^2} - 1 \right)}{\eta} \,\mathrm{d}\eta}_{=:Q_1} + \underbrace{\sum_k \pi_k \int_0^{\frac{1}{2B_T}} \frac{e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k} - 1}{\eta} \,\mathrm{d}\eta}_{=:Q_2}.$$

To control the first term Q_1, we apply the so-called 'prod bound' $e^{x-x^2} \leq 1 + x$ for $x \geq -1/2$ (Cesa-Bianchi et al., 2007) to $x = \eta \bar{r}_{T+1}^k$, which we may do as $\eta \bar{r}_{T+1} \geq -\frac{1}{2B_T} B_T$. Linearity and the definition of the weights (6), yield the following upper-bound on the term Q_1

$$\sum_k \pi_k \int_0^{\frac{1}{2B_T}} \frac{e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k} \eta \bar{r}_{T+1}^k}{\eta} \,\mathrm{d}\eta \;=\; \left\langle \sum_k \pi_k \int_0^{\frac{1}{2B_T}} e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k} \left(\widehat{\boldsymbol{p}}_{T+1} - \boldsymbol{e}_k \right) \mathrm{d}\eta, \bar{\boldsymbol{l}}_{T+1} \right\rangle \;=\; 0.$$

To control the second term Q_2, we extend the range of the integral to find

$$Q_2 \;\leq\; \sum_k \pi_k \int_0^{\frac{1}{2B_{T-1}}} \frac{e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k} - 1}{\eta} \,\mathrm{d}\eta + \ln \frac{B_T}{B_{T-1}} \;=\; \Phi_T + \ln \frac{B_T}{B_{T-1}}.$$

■

Proof of Lemma 2 For any $\epsilon \in [0, 1/(2B_{T-1})]$, we may split the potential (5) as follows

$$\Phi_T \;=\; \underbrace{\sum_k \pi_k \int_0^{\epsilon} \frac{e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k} - 1}{\eta} \,\mathrm{d}\eta}_{=:Q_1} + \underbrace{\sum_k \pi_k \int_{\epsilon}^{\frac{1}{2B_{T-1}}} \frac{e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k} - 1}{\eta} \,\mathrm{d}\eta}_{=:Q_2}.$$

For convenience, let us introduce $\bar{b}_t \coloneqq \max_k |\bar{r}_t^k| = b_t \cdot B_{t-1}/B_t$ and abbreviate $\bar{S}_T \coloneqq \sum_{t=1}^T \bar{b}_t$. To bound the left term Q_1 from below, we use $e^x - 1 \geq x$. Then combined with $\bar{R}_T^k \geq -\bar{S}_T$ and $\bar{V}_T^k \leq \sum_{t=1}^{T-1} \bar{b}_t^2 \leq B_{T-1}\bar{S}_T$ we find

$$Q_1 \;\geq\; \sum_k \pi_k \int_0^{\epsilon} \bar{R}_T^k - \eta \bar{V}_T^k \,\mathrm{d}\eta \;\geq\; -\left(\epsilon + \frac{\epsilon^2}{2} B_{T-1}\right) \bar{S}_T.$$

For the right term Q_2, we use KL duality to find

$$\begin{aligned} Q_2 \;&=\; \sum_k \pi_k \int_{\epsilon}^{\frac{1}{2B_{T-1}}} \frac{e^{\eta \bar{R}_T^k - \eta^2 \bar{V}_T^k}}{\eta} \,\mathrm{d}\eta + \ln\left(2B_{T-1}\epsilon\right), \\ &\geq\; e^{-\mathrm{KL}(\rho\|\pi)} \int_{\epsilon}^{\frac{1}{2B_{T-1}}} \frac{e^{\eta \bar{R}_T^\rho - \eta^2 \bar{V}_T^\rho}}{\eta} \,\mathrm{d}\eta + \ln\left(2B_{T-1}\epsilon\right). \end{aligned}$$

Way pick the admissible $\epsilon = 1/(2(\bar{S}_T + B_{T-1}))$ for which $\left(\epsilon + B_{T-1} \cdot \epsilon^2/2\right) \bar{S}_T \leq 1/2$ (as it is increasing in $\bar{S}_T \geq 0$ and decreasing in $B_{T-1} \geq 0$), and find

$$\Phi_T \;\geq\; e^{-\mathrm{KL}(\rho\|\pi)} \int_{\epsilon}^{\frac{1}{2B_{T-1}}} \frac{e^{\eta \bar{R}_T^\rho - \eta^2 \bar{V}_T^\rho}}{\eta} \,\mathrm{d}\eta - \frac{1}{2} - \ln\left(1 + \frac{\bar{S}_T}{B_{T-1}}\right),$$

which we may reorganise to

$$Q_3 \coloneqq \ln \int_{\frac{1}{2(\bar{S}_T + B_{T-1})}}^{\frac{1}{2B_{T-1}}} \frac{e^{\eta \bar{R}_T^\rho - \eta^2 \bar{V}_T^\rho}}{\eta} \,\mathrm{d}\eta \;\leq\; \mathrm{KL}\left(\rho\|\pi\right) + \ln\left(\Phi_T + \frac{1}{2} + \ln\left(1 + \frac{\bar{S}_T}{B_{T-1}}\right)\right).$$

The argument to bound the integral in Q_3 splits in 3 cases. Let us abbreviate $R \equiv \bar{R}_T^\rho$ and $V \equiv \bar{V}_T^\rho$. Let $\hat{\eta} = \frac{R}{2V}$ be the maximiser of $\eta \to \eta R - \eta^2 V$.

1. First the important case, where $[\hat{\eta} - 1/\sqrt{2V}, \hat{\eta}] \subseteq [1/(2(\bar{S}_T + B_{T+1})), 1/(2B_{T-1})]$. Then

$$\begin{aligned} Q_3 \;&\geq\; \ln \int_{\hat{\eta} - \frac{1}{\sqrt{2V}}}^{\hat{\eta}} \frac{e^{\eta R - \eta^2 V}}{\eta} \,\mathrm{d}\eta \;\geq\; \ln \int_{\hat{\eta} - \frac{1}{\sqrt{2V}}}^{\hat{\eta}} \frac{e^{\left(\hat{\eta} - \frac{1}{\sqrt{2V}}\right)R - \left(\hat{\eta} - \frac{1}{\sqrt{2V}}\right)^2 V}}{\eta} \,\mathrm{d}\eta \\ &=\; \left(\hat{\eta} - \frac{1}{\sqrt{2V}}\right) R - \left(\hat{\eta} - \frac{1}{\sqrt{2V}}\right)^2 V + \ln\ln \frac{\hat{\eta}}{\hat{\eta} - \frac{1}{\sqrt{2V}}} \\ &=\; \frac{R^2}{4V} - \frac{1}{2} + \ln\ln \frac{1}{1 - \frac{\sqrt{2V}}{R}} \;\geq\; \frac{1}{2}\left(\frac{R}{\sqrt{2V}} - 1\right)^2 \end{aligned}$$

where the last inequality uses $\ln\ln(x/(x-1)) \geq 1-x$ for $x \geq 1$, which can be easily verified by a one-dimensional plot. We conclude

$$R \;\leq\; \sqrt{2V}\left(1+\sqrt{2\,\mathrm{KL}\left(\rho\|\pi\right)+2\ln\left(\Phi_T+\frac{1}{2}+\ln\left(1+\frac{\bar{S}_T}{B_{T-1}}\right)\right)}\right).$$

2. Then in the case where $\hat{\eta}-1/\sqrt{2V} < 1/\bar{S}_T$, we have

$$R \;<\; \sqrt{2V}+\frac{2V}{\bar{S}_T} \;\leq\; \sqrt{2V}+2B_{T-1},$$

and we are done again.

3. We come to the final case where $\hat{\eta} > 1/(2B_{T-1})$, meaning that $R > V/B_{T-1}$. Here we use that for any $u \in [1/(2(\bar{S}_T+B_{T-1})), 1/(2B_{T-1})]$

$$Q_3 \;\geq\; \ln\int_u^{\frac{1}{2B_{T-1}}}\frac{e^{uR-u^2V}}{\eta}\,\mathrm{d}\eta \;\geq\; uR(1-uB_{T-1})+\ln\ln\frac{1}{2uB_{T-1}},$$

and hence

$$R \leq \frac{Q_3-\ln\ln\frac{1}{2uB_{T-1}}}{u(1-uB_{T-1})}.$$

Picking the feasible $u=(5-\sqrt{5})/(10B_{T-1})$ and using $-\ln\ln(5/(5-\sqrt{5})) \leq \ln 2$ yields

$$R \;\leq\; 5B_{T-1}\left(\mathrm{KL}\left(\rho\|\pi\right)+\ln\left(\Phi_T+\frac{1}{2}+\ln\left(1+\frac{\bar{S}_T}{B_{T-1}}\right)\right)+\ln 2\right).$$

Finally, using the fact that

$$\frac{\bar{S}_T}{B_{T-1}} = \frac{1}{B_{T-1}}\sum_{t=1}^{T}\frac{B_{t-1}}{B_t}b_t \leq 1+\sum_{t=1}^{T-1}\frac{b_t}{B_t}$$

concludes the proof.

■

Proof of Theorem 3 The idea of the proof is to analyse the rounds in three parts, as shown in Figure 1.

For comparator $\rho \in \triangle_K$, $B > 0$ and $\tau_1, \tau_2 \in \mathbb{N}$ such that $\tau_1 < \tau_2$, we define the regret $R^{\rho}_{(\tau_1,\tau_2]}$ and variance $V^{\rho}_{(\tau_1,\tau_2]}$ of SQUINT+C started at round τ_1+1 (with input B_{τ_1}) and terminated after round τ_2 by

$$R^{\rho}_{(\tau_1,\tau_2]} \coloneqq \sum_{t=\tau_1+1}^{\tau_2}\mathbb{E}_{\rho(k)}\left[r_t^k\right], \quad V^{\rho}_{(\tau_1,\tau_2]} \coloneqq \sum_{t=\tau_1+1}^{\tau_2}\mathbb{E}_{\rho(k)}\left[(r_t^k)^2\right].$$

We also define

$$\Gamma^{\rho}_{(\tau_1,\tau_2]} \;\coloneqq\; \mathrm{KL}\left(\rho\|\pi\right)+\ln\left(\ln\sum_{t=1}^{\tau_2-1}\frac{b_t}{B_t}+\frac{1}{2}+\ln\left(2+\sum_{t=\tau_1+1}^{\tau_2-1}\frac{b_t}{B_t}\right)\right).$$

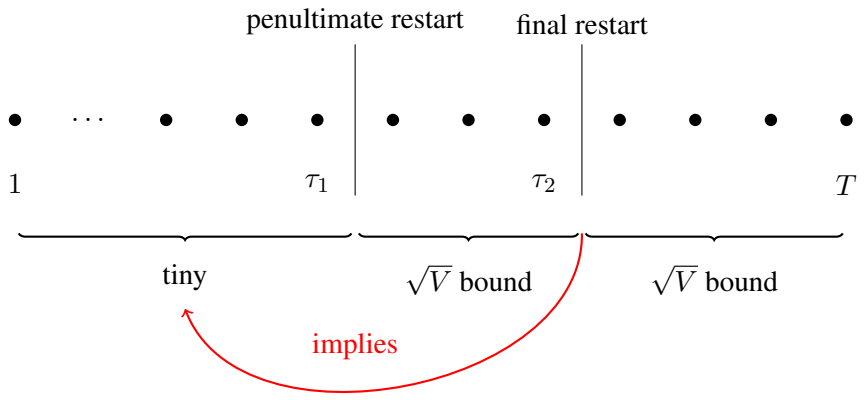

Figure 1: Regret bounding strategy; most general case

Lemma 11 *Let $\rho \in \triangle_K$ and $\tau_1, \tau_2 \in \mathbb{N}$ be such that $\tau_1 < \tau_2$. Suppose that $B_{\tau_2-1}/B_{\tau_1} \leq \sum_{t=1}^{\tau_2-1} b_t/B_t$ (this corresponds to the case where the restart condition in line 2 of Algorithm 1 is not triggered at the end of round $\tau_2 - 1$). Then, the regret $R^\rho_{(\tau_1,\tau_2]}$ of* SQUINT+C *satisfies:*

$$R^\rho_{(\tau_1,\tau_2]} \leq \sqrt{2V^\rho_{(\tau_1,\tau_2]}}\left(1+\sqrt{2\Gamma^\rho_{(\tau_1,\tau_2]}}\right) + 5B_{\tau_2}\left(\Gamma^\rho_{(\tau_1,\tau_2]} + \ln 2\right) + B_{\tau_2}. \tag{19}$$

Proof of Lemma 11 By the assumption that $B_{\tau_2-1}/B_{\tau_1} \leq \ln \sum_{t=1}^{\tau_2-1} b_t/B_t$ and Lemma 1, the potential function Φ_{τ_2} can be upper-bounded by

$$\Phi_{\tau_2} \leq \ln \frac{B_{\tau_2-1}}{B_{\tau_1}} \leq \ln \sum_{t=1}^{\tau_2-1} \frac{b_t}{B_t}.$$

Using this, together with Lemma 2 and (4), we get (19). ■

Assume without loss of generality that $b_1 \neq 0$. Then the regret of SQUINT+L in round $t = 1$ is bounded by $B_1 \leq B_T$, and SQUINT+C is started for the first time in round $t = 2$ with input $B = B_1$.

Now suppose first that the restart condition in line 2 of Algorithm 1 is never triggered, which means that $B_t/B_1 \leq \sum_{s=1}^t b_s/B_s$ for all rounds $t = 2, \ldots, T$. Then for any comparator distribution $\rho \in \triangle_K$, the result follows from Lemma 2 and the facts that $V^\rho_{(1:T]} \leq V^\rho_T$ and $\Gamma^\rho_{(1:T]} \leq \Gamma^\rho_T$.

Alternatively, suppose there is at least one restart. Then let $1 \leq \tau_1 < \tau_2 < T$ be such that $(\tau_1, \tau_2]$ and $(\tau_2, T]$ are the two intervals over which the last two runs of SQUINT+C occurred. We invoke Lemma 2 separately for both these intervals and use Lemma 11 to bound

$$\begin{aligned}
R^\rho_{(\tau_1,T]} &\leq \sqrt{2V^\rho_{(\tau_1,\tau_2]}}\left(1+\sqrt{2\Gamma^\rho_{(\tau_1,\tau_2]}}\right) + 5B_{\tau_2}\left(\Gamma^\rho_{(\tau_1,\tau_2]} + \ln 2\right) + B_{\tau_2} \\
&\quad + \sqrt{2V^\rho_{(\tau_2,T]}}\left(1+\sqrt{2\Gamma^\rho_{(\tau_2,T]}}\right) + 5B_T\left(\Gamma^\rho_{(\tau_2,T]} + \ln 2\right) + B_T, \\
&\leq 2\sqrt{V^\rho_{(\tau_1,T]}}\left(1+\sqrt{2\Gamma^\rho_{(\tau_1,T]}}\right) + 10B_T\left(\Gamma^\rho_{(\tau_1,T]} + \ln 2\right) + 2B_T, \qquad (20) \\
&\leq 2\sqrt{V^\rho_T}\left(1+\sqrt{2\Gamma^\rho_T}\right) + 10B_T\left(\Gamma^\rho_T + \ln 2\right) + 2B_T. \qquad (21)
\end{aligned}$$

where in (20) we used the fact that $\sqrt{x}+\sqrt{y} \le \sqrt{2x+2y}$. If there is exactly one restart, then (21) implies the desired result. If there are multiple restarts, then the proof is completed by bounding the contribution to the regret of all rounds $2, \dots, \tau_1$ by

$$R^{\boldsymbol{u}}_{(1,\tau_1]} \le \sum_{t=2}^{\tau_1} b_t \le B_{\tau_1} \sum_{t=1}^{\tau_1} \frac{b_t}{B_t} \le B_{\tau_1} \sum_{t=1}^{\tau_2} \frac{b_t}{B_t} < B_{\tau_2} \le B_T,$$

where the second to last inequality holds because there was a restart at the end of round $t = \tau_2$. Finally, by bounding the instantaneous regret from the first round by B_T, we obtain the desired result. ■

Appendix B. Proofs of Section 3

Proof of Lemma 5 Let $t \ge 1$. To simplify notation, we denote $\bar{r}^\eta_s := \langle \widehat{\boldsymbol{u}}_s - \widehat{\boldsymbol{u}}^\eta_s, \bar{\boldsymbol{g}}_s \rangle$, for $\boldsymbol{u} \in \mathcal{U}$ and $s \in \mathbb{N}$. By appealing to the prod-bound (*i.e.* $e^{x-x^2} - 1 \le x$, for $x \ge -1/2$), we have

$$\begin{aligned}
\Phi_{t+1} &= \pi(\mathcal{G}_{t+1} \setminus \mathcal{A}_{t+1}) + \sum_{\eta \in \mathcal{A}_{t+1}} w^\eta_{t+1} \left(e^{\eta \bar{r}^\eta_{t+1} - \eta (\bar{r}^\eta_{t+1})^2} - 1 \right) + \sum_{\eta \in \mathcal{A}_{t+1}} w^\eta_{t+1}, \\
&\le \pi(\mathcal{G}_{t+1} \setminus \mathcal{A}_{t+1}) + \sum_{\eta \in \mathcal{A}_{t+1}} w^\eta_{t+1} \eta \bar{r}^\eta_{t+1} + \sum_{\eta \in \mathcal{A}_{t+1}} w^\eta_{t+1}.
\end{aligned}$$

Now by (14)

$$\sum_{\eta \in \mathcal{A}_{t+1}} w^\eta_{t+1} \eta \bar{r}^\eta_{t+1} = \sum_{\eta \in \mathcal{A}_{t+1}} \eta w^\eta_{t+1} (\widehat{\boldsymbol{u}}_{t+1} - \widehat{\boldsymbol{u}}^\eta_{t+1})^\intercal \bar{\boldsymbol{g}}_t = 0.$$

Moreover, by definition of $\mathcal{G}_t$ and $\mathcal{A}_t$,

$$\begin{aligned}
&\pi(\mathcal{G}_{t+1} \setminus \mathcal{A}_{t+1}) + \sum_{\eta \in \mathcal{A}_{t+1}} w^\eta_{t+1} = \pi(\{\eta \in \mathcal{G}_{t+1} : s_\eta > t\}) + \sum_{\eta \in \mathcal{G}_{t+1} : s_\eta \le t} w^\eta_{t+1}, \\
&\le \pi(\{\eta \in \mathcal{G}_t : s_\eta > t\}) + \sum_{\eta \in \mathcal{G}_t : s_\eta \le t} w^\eta_{t+1} = \pi(\{\eta \in \mathcal{G}_t : s_\eta \ge t\}) + \sum_{\eta \in \mathcal{G}_t : s_\eta < t} w^\eta_{t+1}, \\
&= \pi(\mathcal{G}_t \setminus \mathcal{A}_t) + \sum_{\eta \in \mathcal{A}_t} w^\eta_{t+1} = \Phi_t.
\end{aligned}$$

Where we used that $w^\eta_{s_\eta+1} = \pi(\eta)$. Finally, as $\mathcal{A}_0 = \emptyset$ and $\mathcal{G}_0 = \mathcal{G}$, we find $\Phi_0 = \pi(\mathcal{G}) = 1$. ■

Proof of Theorem 6 Throughout this proof we will deal with slaves $\eta \in \mathcal{G}_T \setminus \mathcal{A}_T$ that are provisioned but not active yet by time T, and we will interpret their $s_\eta = T$ for uniform treatment, even though technically all we know from (13) is that $s_\eta \ge T$.

First due to Lemma 5, we have $\Phi_T \le 1$, where Φ_T is the potential defined in (15). Taking logarithms and rearranging, we find

$$\forall \eta \in \mathcal{G}_T, \quad - \sum_{t=s_\eta+1}^{T} \bar{f}_t(\widehat{\boldsymbol{u}}^\eta_t, \eta) \le -\ln \pi(\eta). \tag{22}$$

Moreover, every slave $\eta \in \mathcal{G}_T$ guarantees the following regret for the rounds $t = s_\eta + 1, \dots, T$ (see Van Erven and Koolen 2016, Lemma 5):

$$\sum_{t=s_\eta+1}^{T} \left(\bar{f}_t(\widehat{\boldsymbol{u}}_t^\eta, \eta) - \bar{f}_t(\boldsymbol{u}, \eta)\right) \le \ln\det\left(\mathbf{I} + 2\eta^2 D^2(\bar{\mathbf{S}}_T - \bar{\mathbf{S}}_{s_\eta})\right) + \tfrac{\|\boldsymbol{u}\|^2}{2D^2},$$
$$\le d\ln\left(1 + \tfrac{2D^2}{25dB_{T-1}^2}\operatorname{tr}\bar{\mathbf{S}}_T\right) + \tfrac{\|\boldsymbol{u}\|^2}{2D^2}, \tag{23}$$

where in (23) we used concavity of $\ln\det$, $\bar{\mathbf{S}}_{s_\eta} \succeq \mathbf{0}$, and the fact that $\eta \in \mathcal{G}_T \subset [0, 1/(5B_{T-1})]$. We then invert the 'wake up condition' (13) at time $s_\eta - 1$ to infer

$$-\sum_{t=1}^{s_\eta} \bar{f}_t(\boldsymbol{u}, \eta) \le \eta \sum_{t=1}^{s_\eta} \bar{r}_t^{\boldsymbol{u}} \le \frac{\sum_{t=1}^{s_\eta-1} \bar{r}_t^{\boldsymbol{u}} + \bar{r}_{s_\eta}^{\boldsymbol{u}}}{D\sum_{t=1}^{s_\eta-1}\|\bar{\boldsymbol{g}}_t\|_2 + B_{s_\eta-1}} \le 1. \tag{24}$$

Combining the bounds (22), (23), and (24), then dividing through by η, gives:

$$\forall \eta \in \mathcal{G}_T, \quad \bar{R}_T^{\boldsymbol{u}} \le \eta \bar{V}_T^{\boldsymbol{u}} + \tfrac{1}{\eta} C_T(\eta), \tag{25}$$

where $C_T(\eta) \coloneqq d\ln\left(1 + \tfrac{2D^2}{25dB_{T-1}^2}\operatorname{tr}\bar{\mathbf{S}}_T\right) - \ln\pi(\eta) + 2$.

Let C_T be as in the theorem statement and η_* be the estimator defined by $\eta_* \coloneqq \sqrt{C_T/\bar{V}_T^{\boldsymbol{u}}}$. Suppose that $\eta_* \le 1/(5B_{T-1})$. By construction of the grid $\mathcal{G}_T$, there exists $i \in \mathbb{N}$ such that

$$\hat{\eta} \coloneqq 2^{-i}/(5B_0) \in \mathcal{G}_T \;\text{ and }\; \hat{\eta} \in [\eta_*/2, \eta_*]\,. \tag{26}$$

Since $C_T \ge 1$, the estimator η_* can be lower-bounded by $1/\sqrt{\bar{V}_T^{\boldsymbol{u}}}$, and thus due to (26) we have $2^{-i}/(5B_0) \ge 1/\sqrt{4\bar{V}_T^{\boldsymbol{u}}}$. This implies that the prior weight on $\hat{\eta}$ satisfies

$$\frac{1}{\pi(\hat{\eta})} = (i+1)(i+2) \le \left(\log_2 \tfrac{2\sqrt{\bar{V}_T^{\boldsymbol{u}}}}{5B_0} + 1\right)\left(\log_2 \tfrac{2\sqrt{\bar{V}_T^{\boldsymbol{u}}}}{5B_0} + 2\right) \le \left(\log_2 \tfrac{\sqrt{\bar{V}_T^{\boldsymbol{u}}}}{B_0} + 3\right)^2. \tag{27}$$

Now from the fact that $1/\sqrt{\bar{V}_T^{\boldsymbol{u}}} \le \eta_* \le 1/(5B_{T-1}) \le 1/(5B_0)$, we have $\sqrt{\bar{V}_T^{\boldsymbol{u}}}/B_0 \ge 2$. This, combined with (27), implies that $C_T(\hat{\eta}) \le C_T$, where C_T is as in the theorem statement. Plugging $\eta = \hat{\eta}$ into (25) and using the fact that $\hat{\eta} \in [\eta_*/2, \eta_*]$, gives

$$\bar{R}_T^{\boldsymbol{u}} \le \hat{\eta}\bar{V}_T^{\boldsymbol{u}} + \tfrac{1}{\hat{\eta}} C_T(\hat{\eta}) \le \eta_*\bar{V}_T^{\boldsymbol{u}} + \tfrac{2}{\eta_*} C_T = 3\sqrt{\bar{V}_T^{\boldsymbol{u}} C_T}. \tag{28}$$

Now suppose that $\eta_* > 1/(5B_{T-1})$, and let $\hat{\eta} \coloneqq \max \mathcal{G}_T \ge 1/(10B_{T-1})$, where the last inequality follows by construction of $\mathcal{G}_T$. Note that in this case $\frac{1}{\pi(\hat{\eta})} \le (\log_2 \frac{2B_{T-1}}{B_0} + 1)(\log_2 \frac{2B_{T-1}}{B_0} + 2)$, and the inequality $C_T(\hat{\eta}) \le C_T$ still holds. Plugging $\eta = \hat{\eta}$ into (25) and using the assumption on η_*, *i.e.* $\eta_* > 1/(5B_{T-1})$, we obtain

$$\bar{R}_T^{\boldsymbol{u}} \le \hat{\eta}\bar{V}_T^{\boldsymbol{u}} + \tfrac{1}{\hat{\eta}} C_T(\hat{\eta}) \le \hat{\eta}\bar{V}_T^{\boldsymbol{u}} + \tfrac{1}{\hat{\eta}} C_T \le 15 B_T C_T. \tag{29}$$

By combining (28) and (29), we get the desired result. ∎

Proof of Theorem 8 Assume without loss of generality that $b_1 \neq 0$. Then the regret of METAGRAD+L in round one is bounded by $B_1 \leq B_T$, and METAGRAD+C is started for the first time in round $t = 2$ with parameter $B = B_1$.

Let $V^{\boldsymbol{u}}_{(1:T]}$ and $C_{(1:T]}$ represent the quantities denoted by $V^{\boldsymbol{u}}_T$ and C_T in Theorem 6 but measured on rounds $2, \dots, T$. Now suppose first that the restart condition in line 2 of Algorithm 1 is never triggered, which means that

$$\frac{B_t}{B_1} \leq \sum_{s=1}^{t} \frac{b_s}{B_s}, \qquad \text{for all rounds } t = 2, \dots, T. \tag{30}$$

Then the result follows from Theorem 6, $V^{\boldsymbol{u}}_{(1:T]} \leq V^{\boldsymbol{u}}_T$, for all $\boldsymbol{u} \in \mathcal{U}$, and

$$\begin{aligned} C_{(1:T]} &= d \ln \left(1 + \frac{2}{25d} \frac{\sum_{t=1}^{T-1} b_t^2}{B_{T-1}^2} \right) + 2 \ln \left(\log_2^+ \frac{\sqrt{\sum_{t=2}^{T} b_t^2}}{B_1} + 3 \right) + 2, \\ &\leq d \ln \left(1 + \frac{2}{25d} \frac{\sum_{t=1}^{T-1} b_t^2}{B_{T-1}^2} \right) + 2 \ln \left(\log_2^+ \sqrt{\sum_{t=2}^{T} \left(\sum_{s=1}^{t} \frac{b_s}{B_s} \right)^2} + 3 \right) + 2, \qquad (31) \\ &\leq \Gamma_T, \end{aligned}$$

where in (31), we used (30). Alternatively, suppose there is at least one restart. Then let $1 \leq \tau_1 < \tau_2 < T$ be such that $(\tau_1, \tau_2]$ and $(\tau_2, T]$ are the two intervals over which the last two runs of METAGRAD+C occurred. We invoke Theorem 6 separately for both these intervals to bound

$$\begin{aligned} R^{\boldsymbol{u}}_{(\tau_1,T]} &\leq 3\sqrt{V^{\boldsymbol{u}}_{(\tau_1,\tau_2]} C_{(\tau_1,\tau_2]}} + 15 B_T C_{(\tau_1,\tau_2]} + B_{\tau_2} \\ &\quad + 3\sqrt{V^{\boldsymbol{u}}_{(\tau_2,T]} C_{(\tau_2,T]}} + 15 B_T C_{(\tau_2,T]} + B_T, \\ &\leq 3\sqrt{V^{\boldsymbol{u}}_{(\tau_1,\tau_2]} \Gamma_T / 2} + 3\sqrt{V^{\boldsymbol{u}}_{(\tau_2,T]} \Gamma_T / 2} + 15 B_T \Gamma_T + 2B_T, \\ &\leq 3\sqrt{V^{\boldsymbol{u}}_{(\tau_1,T]} \Gamma_T} + 15 B_T \Gamma_T + 2B_T, \qquad (32) \end{aligned}$$

where a subscript $(\tau_1, \tau_2]$ indicates a quantity measured only on rounds $\tau_1 + 1, \dots, \tau_2$ and the last inequality uses $\sqrt{x} + \sqrt{y} \leq \sqrt{2x + 2y}$. If there is exactly one restart, then (32) implies the desired result. If there are multiple restarts, then the proof is completed by bounding the contribution to the regret of all rounds $2, \dots, \tau_1$ by

$$R^{\boldsymbol{u}}_{(1,\tau_1]} \leq \sum_{t=2}^{\tau_1} b_t \leq B_{\tau_1} \sum_{t=1}^{\tau_1} \frac{b_t}{B_t} \leq B_{\tau_1} \sum_{t=1}^{\tau_2} \frac{b_t}{B_t} < B_{\tau_2} \leq B_T,$$

where the second to last inequality holds because there was a restart at $t = \tau_2$. Finally, by bounding the instantaneous regret from the first round by B_T, we obtain the desired result. ∎

Appendix C. Proofs of Section 4

Proof of Lemma 9 We use the Lagrangian multiplier to solve (17). For this, let

$$\mathcal{L}(\boldsymbol{u},\mu) \coloneqq \left(\boldsymbol{u}_{t+1}^{\eta}-\boldsymbol{u}\right)^{\intercal}\left(\boldsymbol{\Sigma}_{t+1}^{\eta}\right)^{-1}\left(\boldsymbol{u}_{t+1}^{\eta}-\boldsymbol{u}\right)+\mu(\boldsymbol{u}^{\intercal}\boldsymbol{u}-D^2).$$

Setting $\frac{\partial\mathcal{L}}{\partial\boldsymbol{u}}=0$ implies that $2\left(\boldsymbol{\Sigma}_{t+1}^{\eta}\right)^{-1}\left(\boldsymbol{u}-\boldsymbol{u}_{t+1}^{\eta}\right)+2\mu\boldsymbol{u}=0$. After rearranging, this becomes

$$\begin{aligned}\boldsymbol{u} &= \left(\left(\mu+\tfrac{1}{D^2}\right)\mathbf{I}+2\eta^2\mathbf{S}_t\right)^{-1}\left(\boldsymbol{\Sigma}_{t+1}^{\eta}\right)^{-1}\boldsymbol{u}_t^{\eta},\\ &= \mathbf{Q}_t^{\intercal}\left(x\mathbf{I}+2\eta^2\boldsymbol{\Lambda}_t\right)^{-1}\mathbf{Q}_t\boldsymbol{v}_{t+1}^{\eta},\end{aligned}$$

where we set $x\coloneqq\mu+1/D^2$. The result follows after observing that $\boldsymbol{u}^{\intercal}\boldsymbol{u}=D^2/4 \iff \rho_t^{\eta}(x)=D^2/4$. ■

Proof of Theorem 10 Let $\mathring{R}_T^{\boldsymbol{u}}\coloneqq\sum_{t=1}^T\langle\widehat{\boldsymbol{w}}_t-\boldsymbol{u},\mathring{\boldsymbol{g}}_t\rangle$ and $\mathring{V}_T^{\boldsymbol{u}}\coloneqq\sum_{t=1}^T\langle\widehat{\boldsymbol{w}}_t-\boldsymbol{u},\mathring{\boldsymbol{g}}_t\rangle^2$ be the pseudo-regret and 'variance' of Algorithm 2. From our assumption on the pseudo-regret $\tilde{R}_T^{\boldsymbol{u}}$ of METAGRAD and the fact that $2\sqrt{x}=\inf_{\eta>0}\{\eta x+1/\eta\}$, we have

$$\forall\boldsymbol{u}\in\mathcal{U}\subset\mathcal{B}_D,\forall\eta>0,\quad \eta\tilde{R}_T^{\boldsymbol{u}}-\tfrac{\eta^2}{2}V_T^{\boldsymbol{u}}\le\tfrac{1}{2}\Gamma_T+\eta B\Gamma_T. \tag{33}$$

Now, as in the proof of (Cutkosky and Orabona, 2018, Theorem 3), we have

$$\langle\widehat{\boldsymbol{w}}_t-\boldsymbol{u},\mathring{\boldsymbol{g}}_t\rangle\le 2\mathring{\ell}_t(\widehat{\boldsymbol{u}}_t)-2\mathring{\ell}_t(\boldsymbol{u}), \tag{34}$$

where $\widehat{\boldsymbol{w}}_t=\Pi_{\mathcal{U}}(\widehat{\boldsymbol{u}}_t)$ is the prediction of Algorithm 2 at round t and $\mathring{\ell}_t$ is the function defined by $\mathring{\ell}_t(\boldsymbol{u})\coloneqq\frac{1}{2}\left(\langle\mathring{\boldsymbol{g}}_t,\boldsymbol{u}\rangle+\|\mathring{\boldsymbol{g}}_t\|\,\mathrm{d}_{\mathcal{U}}(\boldsymbol{u})\right)$. By convexity of $\mathring{\ell}_t$ and the fact that $\boldsymbol{g}_t\in\partial\mathring{\ell}_t(\widehat{\boldsymbol{u}}_t)$, we have

$$\langle\widehat{\boldsymbol{u}}_t-\boldsymbol{u},\boldsymbol{g}_t\rangle\ge\mathring{\ell}_t(\widehat{\boldsymbol{u}}_t)-\mathring{\ell}_t(\boldsymbol{u})\ge\tfrac{1}{2}\langle\widehat{\boldsymbol{w}}_t-\boldsymbol{u},\mathring{\boldsymbol{g}}_t\rangle,\quad\text{for }\boldsymbol{u}\in\mathcal{U}, \tag{35}$$

where the right-most inequality follows from (34). Since the function $x\mapsto x-x^2/2$ is strictly increasing on the interval $]-\infty,1]$, (35) implies that for all $\eta\in]0,1/B]=]0,1/(DG)]$,

$$\tfrac{\eta}{2}\langle\widehat{\boldsymbol{w}}_t-\boldsymbol{u},\mathring{\boldsymbol{g}}_t\rangle-\tfrac{\eta^2}{8}\langle\widehat{\boldsymbol{w}}_t-\boldsymbol{u},\mathring{\boldsymbol{g}}_t\rangle^2\le\eta\langle\widehat{\boldsymbol{u}}_t-\boldsymbol{u},\boldsymbol{g}_t\rangle-\tfrac{\eta^2}{2}\langle\widehat{\boldsymbol{u}}_t-\boldsymbol{u},\boldsymbol{g}_t\rangle^2,\quad\text{for }\boldsymbol{u}\in\mathcal{U}.$$

Summing this over $t=1,\dots,T$ and using (33), we get for all $\eta\in]0,1/B]$ and $\boldsymbol{u}\in\mathcal{U}$,

$$\begin{aligned}&\tfrac{1}{2}\mathring{R}_T^{\boldsymbol{u}}-\tfrac{\eta}{8}\mathring{V}_T^{\boldsymbol{u}}\le\tilde{R}_T^{\boldsymbol{u}}-\tfrac{\eta}{2}V_T^{\boldsymbol{u}}\le\tfrac{1}{2\eta}\Gamma_T+B\Gamma_T,\quad\text{and so}\\ &\mathring{R}_T^{\boldsymbol{u}}\le\tfrac{\eta}{4}\mathring{V}_T^{\boldsymbol{u}}+\tfrac{1}{\eta}\Gamma_T+2B\Gamma_T.\end{aligned} \tag{36}$$

The 'unconstrained' $\eta\in[0,+\infty]$ which minimizes the RHS of (36) is given by $\eta_*\coloneqq2\sqrt{\Gamma_T/\mathring{V}_T^{\boldsymbol{u}}}$. We consider two cases: suppose first that $\eta_*\le1/B$. For $\eta=\eta_*$, we have

$$\tfrac{\eta}{4}\mathring{V}_T^{\boldsymbol{u}}+\tfrac{1}{\eta}\Gamma_T=\sqrt{\mathring{V}_T^{\boldsymbol{u}}\Gamma_T}. \tag{37}$$

Now suppose that $\eta_*>1/B$. For $\eta=1/B$, we have

$$\tfrac{\eta}{4}\mathring{V}_T^{\boldsymbol{u}}+\tfrac{1}{\eta}\Gamma_T\le2B\Gamma_T. \tag{38}$$

Combining (36)–(38) yields the desired bound. ■