Penalized FTRL with Time-Varying Constraints

Leith, Douglas J.; Iosifidis, George

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Penalized FTRL with Time-Varying Constraints

Douglas J. Leith[1(✉)] and George Iosifidis[2]

[1] Trinity College Dublin, Dublin, Republic of Ireland
`doug.leith@tcd.ie`
[2] Delft University of Technology, Delft, The Netherlands

Abstract. In this paper we extend the classical Follow-The-Regularized-Leader (FTRL) algorithm to encompass time-varying constraints, through adaptive penalization. We establish sufficient conditions for the proposed Penalized FTRL algorithm to achieve $\mathcal{O}(\sqrt{t})$ regret and violation with respect to a strong benchmark $\hat{X}_t^{max}$. Lacking prior knowledge of the constraints, this is probably the largest benchmark set that we can reasonably hope for. Our sufficient conditions are necessary in the sense that when they are violated there exist examples where $\mathcal{O}(\sqrt{t})$ regret and violation is not achieved. Compared to the best existing primal-dual algorithms, Penalized FTRL substantially extends the class of problems for which $\mathcal{O}(\sqrt{t})$ regret and violation performance is achievable.

Keywords: FTRL · Online convex optimization · Constrained optimization

1 Introduction

The introduction of online convex optimization (OCO) [15] offered an effective way to tackle online learning and dynamic decision problems, with applications that range from portfolio selection, to routing optimization and ad placement, see [3]. One of the seminal OCO algorithms is the Follow-The-Regularized-Leader (FTRL), which includes online gradient descent and mixture of experts as special cases. Indeed, FTRL is widely used in different contexts, e.g., with linear or non-linear objective functions, composite objectives [8], or budget constraints [1].

The general form of the FTRL update is:

$$x_{\tau+1} \in \arg\min_{x \in X} \left\{ R_\tau(x) + \sum_{i=1}^{\tau} F_i(x) \right\} \tag{1}$$

where action set $X \subset \mathbb{R}^n$ is bounded, function $F_i : X \to \mathbb{R}$ and regularizer $R_\tau : X \to \mathbb{R}$ are strongly convex. When the sum-loss $\sum_{i=1}^{\tau} F_i(x)$ is convex and $F_i(x)$ and $\left(R_i(x) - R_{i-1}(x)\right)$ are uniformly Lipschitz, the FTRL-generated sequence $\{x_\tau\}_{\tau=1}^{t}$ induces regret $\sum_{i=1}^{t} \left(F_i(x_i) - F_i(x)\right) \leq \mathcal{O}(\sqrt{t})$, $\forall x \in X$, cf. [8]. Importantly, the set X of admissible actions must be fixed and this is intrinsic to the method of proof, i.e., it is not a minor or incidental assumption.

The focus of this paper is to extend the FTRL algorithm in order to accommodate time-varying action sets, i.e., cases where at each time τ the fixed set action X is replaced by set X_τ which may vary over time. We refer to this extension to FTRL as *Penalized FTRL* or `Pen-FTRL`.

In general, it is too much to expect to be able to simultaneously achieve $\mathcal{O}(\sqrt{t})$ regret and strict feasibility $x_\tau \in X_\tau$, $\tau = 1, \ldots, t$. We therefore allow limited violation of the action sets $\{X_\tau\}$ and instead aim to simultaneously achieve $\mathcal{O}(\sqrt{t})$ regret and $\mathcal{O}(\sqrt{t})$ constraint violation. That is, defining loss function $f_\tau : D \to \mathbb{R}$ on domain $D \subset \mathbb{R}^n$ and constraint functions $g_\tau^{(j)} : D \to \mathbb{R}$ such that $X_\tau = \left\{ x \in D : g_\tau^{(j)}(x) \le 0, j = 1, \ldots, m \right\}$ then we aim to simultaneously achieve regret and violation:

$$\mathcal{R}_t = \sum_{i=1}^{t} \left(f_i(x_i) - f_i(x) \right) \le \mathcal{O}(\sqrt{t}), \ \mathcal{V}_t = \sum_{j=1}^{m} \max\left\{ 0, \sum_{i=1}^{t} g_i^{(j)}(x_i) \right\} \le \mathcal{O}(\sqrt{t})$$

for all $x \in X_t^{max} := \left\{ x \in D : \sum_{i=1}^{t} g_i^{(j)}(x) \le 0, j = 1, \ldots, m \right\}$.

Importance of Using A Strong Benchmark. We know from [7] that $\mathcal{O}(\sqrt{t})$ regret and violation with respect to benchmark set X_t^{max} is not achievable for all possible sequences of constraints $\{g_i^{(j)}\}$. It is therefore necessary to: *(i)* change the benchmark set X_t^{max} to something more restrictive; or *(ii)* restrict the admissible set of constraint sequences $\{g_i^{(j)}\}$; or *(iii)* both. In the literature, it is common to adopt the weaker benchmark:

$$X_t^{min} := \left\{ x \in D : g_i^{(j)}(x) \le 0, i = 1, \ldots, t, j = 1, \ldots, m \right\} \subset X_t^{max}$$

i.e., to focus on actions x which *simultaneously satisfy every constraint at every time.* But this weak benchmark is in fact so restrictive and easy for a learning algorithm to outperform, where the achieved regret $\mathcal{R}_t$ is often negative.

One of our primary interests, therefore, is in retaining a benchmark that is close to X_t^{max}. To this end, we consider the following benchmark:

$$\hat{X}_t^{max} := \left\{ x \in D : \sum_{i=1}^{\tau} g_i^{(j)}(x) \le 0, \forall j = 1, \ldots, m \tau = 1, \ldots, t \right\}.$$

We can see immediately that $X_t^{min} \subset \hat{X}_t^{max}$. The set $\hat{X}_t^{max}$ requires $\sum_{i=1}^{\tau} g_i^{(j)}(x) \le 0$ to hold at every time $\tau \le t$ rather than just at the end of the horizon t, and so is still smaller than X_t^{max}. Lacking, however, predictions or prior knowledge of the constraints $g_i^{(j)}$, it is probably the best we can reasonably hope for. To illustrate the difference between $\hat{X}_t^{max}$ and X_t^{min}, suppose the time-varying constraint is $x \le 1/\sqrt{t}$. Then $X_t^{min} = [0, 1/\sqrt{t}]$ which tends to set $\{0\}$ for t large, $X_t^{max} = D = [0, 1]$ for $t \ge 1$, and $\hat{X}_t^{max} = D = [0, 1]$.

Role of `Pen-FTRL`. Almost all of the literature focuses on using primal-dual algorithms to accommodate time-varying constraints (see Sect. 2). In contrast, here we use a direct penalty-based approach [12], which we refer to as *Penalized*

FTRL or `Pen-FTRL`. This has the important advantages of *(i)* conceptually separating the issue of multiplier selection (i.e. λ in the above primal-dual update) from the issue of sum-constraint violation and so facilitating analysis; and *(ii)* maintaining a direct link with the well-established FTRL algorithm.

Contributions. In summary, the main contributions of the present paper are: *(i)* introduction of the `Pen-FTRL` extension to FTRL; and *(ii)* establishing sufficient conditions for `Pen-FTRL` to achieve $\mathcal{O}(\sqrt{t})$ regret and violation with respect to benchmark $\hat{X}_t^{max}$. Lacking prior knowledge of the constraints, $\hat{X}_t^{max}$ is probably the largest benchmark set that we can reasonably hope for. Our sufficient conditions are necessary in the sense that when they are violated there exist examples where $\mathcal{O}(\sqrt{t})$ regret and violation is not achieved. Compared to the best existing primal-dual algorithms, `Pen-FTRL` substantially extends the class of problems for which $\mathcal{O}(\sqrt{t})$ regret and violation performance is achievable.

2   Related Work

The literature on online learning with time-varying constraints focuses on primal-dual algorithms (see update (7) in the sequel), and largely fails to obtain $\mathcal{O}(\sqrt{t})$ regret and violation simultaneously even w.r.t. the weak X_t^{min} benchmark. The standard problem setup consists of a sequence of convex cost functions $f_t : D \rightarrow \mathbb{R}$ and constraints $g_t^{(j)} : X \rightarrow \mathbb{R}$, $j = 1, \ldots, m$, where actions $x \in D \subset \mathbb{R}^n$. The canonical algorithm performs a primal-dual gradient descent iteration, namely:

$$x_{t+1} = \Pi_D \left(x_t - \eta_t (\partial f_t(x_t) + \lambda_t^T \partial g_t(x_t)) \right), \ \lambda_{t+1} = \left[(1 - \theta_t)\lambda_t + \mu_t g_t(x_{t+1}) \right]^+ \quad (2)$$

with step-size parameters η_t, μ_t and regularization parameter θ_t; while $\Pi_D(\alpha)$ denotes the project of α onto D. Commonly, the parameter $\theta_t \equiv 0$, with exceptions being [4,6], and [10] that employ non-zero θ_t. [14] approximate $g_t(x_{t+1})$ in the λ_{t+1} update by $g_t(x_t) + \partial g_t(x_t)(x_{t+1} - x_t)$.

The $\mathcal{R}_t$ is commonly measured w.r.t. the baseline action set X_t^{min}, with the exception of [11] where a slightly larger set is considered; [14] that considers stochastic constraints and the baseline action set is $\{x \in D : E[g_i^{(j)}(x)] \leq 0, j = 1, \ldots, m\}$; and [5] which considers a K-slot moving window for the sum-constraint satisfaction. The original work on this topic restricted attention to time-invariant constraints $g_i^{(j)}(x) = g^{(j)}(x)$. With this restriction, the work in [4] achieves $\mathcal{R}_t \leq \mathcal{O}(\max\{t^\beta, t^{1-\beta}\})$ and $\mathcal{V}_t \leq \mathcal{O}(t^{1-\beta/2})$ constraint violation, which yields $\mathcal{R}_t, \mathcal{V}_t \leq \mathcal{O}(t^{2/3})$ with $\beta = 2/3$. Similar bounds are derived in [6]. It is worth noting that these results are primarily of interest for their analysis of the primal-dual algorithm rather than the performance bounds per se, since classical algorithms such as FTRL are already known to achieve $\mathcal{O}(\sqrt{t})$ regret and no constraint violation for constant constraints.

For general time-varying cost and constraint functions, [10] achieves $\mathcal{O}(\sqrt{t})$ regret and $\mathcal{O}(t^{3/4})$ constraint violation; [5] achieve $\mathcal{R}_t = \mathcal{O}(\sqrt{t} + Kt/V)$ and $\mathcal{V}_t = \mathcal{O}(\sqrt{Vt})$, with $K = 1$ corresponding to baseline set X_t^{min} and V a design parameter. Selecting $V = \sqrt{t}$ gives $\mathcal{O}(\sqrt{t})$ regret and $\mathcal{O}(t^{3/4})$ constraint violation,
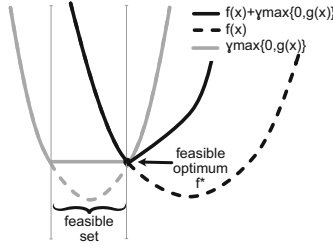
Fig. 1. Illustrating use of a penalty to convert constrained optimization $\min_{x:g(x)\leq 0} f(x)$ into unconstrained optimization $\min_x f(x) + \gamma \max\{0, g(x)\}$, $\gamma > 0$. Within the feasible set $g(x) \leq 0$ and $\gamma \max\{0, g(x)\} = 0$. Outwith this set $\gamma \max\{0, g(x)\} = \gamma g(x) > 0$. The idea is that γ is selected large enough that outwith the feasible set $f(x) + \gamma \max\{0, g(x)\} > f^*$, the min value of f inside the feasible set.

similarly to [10]. By restricting the constraints, [11] improves this to $\mathcal{O}(\sqrt{t})$ regret and constraint violation. As already noted, this requires restricting the constraints to be $g_i^{(j)}(x) = g^{(j)}(x) - b_{i,j}$ with $b_{i,j} \in \mathbb{R}$ i.e. the constraints are $g^{(j)}(x) \leq b_i^{(j)}$ with time-variation confined to threshold $b_i^{(j)}$. Yu et al [14] also achieve $\mathcal{O}(\sqrt{t})$ regret and *expected* constraint violation (i.e. $E[\sum_{i=1}^{t} g_i^{(j)}(x_t)] \leq \mathcal{O}(\sqrt{t})$), this time by restricting the constraints to be i.i.d. stochastic. Yi et al [13] obtain $\mathcal{R}_t, \mathcal{V}_t = \mathcal{O}(t^{2/3})$ by restricting the cost and constraint functions to be separable, while [2] focuses on a form of dynamic regret that upper bounds the static regret and show $o(t)$ regret and $\mathcal{O}(t^{2/3})$ constraint violation under a slow variation condition on the constraints and dynamic baseline action.

3   Preliminaries

3.1   Exact Penalties

We begin by recalling a classical result of Zangwill [12]. Consider the convex optimization problem P:

$$\min_{x \in D} f(x) \qquad \text{s.t.} \qquad g^{(j)}(x) \leq 0, \ j = 1, \cdots, m$$

where $D \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ and $g^{(j)} : \mathbb{R}^n \to \mathbb{R}$, $j = 1, \cdots, m$ are convex. Let $X := \{x : x \in D, g^{(j)}(x) \leq 0, j = 1, \cdots, m\}$ denote the feasible set and $X^* \subset X$ the set of optimal points. Define:

$$F(x) := f(x) + \gamma \sum_{j=1}^{m} \max\left\{0, g^{(j)}(x)\right\}, \quad \gamma \in \mathbb{R}. \tag{3}$$

$F(x)$ is convex since $f(\cdot), g^{(j)}(\cdot)$ are convex and $\max\{\cdot\}$ preserves convexity.

The key idea is that the penalty (second term in (3)) is zero for $x \in X$, but large when $x \notin X$. Provided γ is selected large enough, the penalty forces the minimum of $F(x)$ to *(i)* lie in X and *(ii)* match $\min_{x \in X} f(x)$; see example in Fig. 1. The next lemma, proved in the Appendix, corresponds to [12, Lemma 2].

Lemma 1 (Exact Penalty). *Assume that a Slater point exists i.e. a feasible point $z \in D$ such that $g^{(j)}(z) < 0$, $j = 1, \cdots, m$. Let $f^* := \inf_{x \in X} f(x)$ (the solution to optimization P) . Then there exists a finite threshold $\gamma_0 \geq 0$ such that $F(x) \geq f^*$ for all $x \in D$, $\gamma \geq \gamma_0$, with equality only when $x \in X^*$. It is sufficient to choose $\gamma_0 = \frac{f^* - f(z) - 1}{\max_{j \in \{1, \cdots, m\}} \{g^{(j)}(z)\}}$.*

3.2 FTRL Results

We also recall the following standard FTRL results (for proofs see, e.g., [9]).

Lemma 2 (Be-The-Leader). *Let $F_i, i = 1, \ldots, t$ be a sequence of (possibly non-convex) functions $F_i : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$. Assume that $\arg\min_{x \in D} \sum_{i=1}^{\tau} F_i(x)$ is not empty for $\tau = 1, \ldots, t$. Selecting sequence $w_{i+1}, i = 1, \ldots, t$ according to the Follow The Leader (FTL) update $w_{\tau+1} \in \arg\min_{x \in D} \sum_{i=1}^{\tau} F_i(x)$, ensures $\sum_{i=1}^{t} F_i(w_{i+1}) \leq \sum_{i=1}^{t} F_i(y)$ for every $y \in D$.*

Condition 1 (FTRL) *(i) Domain D is bounded (potentially non-convex), (ii) $\sum_{i=1}^{\tau} F_i(x)$ is convex (the individual F_i's need not be convex), (iii) $F_i(x)$ is uniformly L_f-Lipschitz on D i.e. $|F_i(x) - F_i(y)| \leq L_f \|x - y\|$ for all $x, y \in D$ and where L_f does not depend on i, and (iv) $R_\tau(x)$ is $\sqrt{\tau}$-strongly convex and $(R_i(x) - R_{i-1}(x))$ is uniformly Lipschitz, e.g. $\sqrt{\tau}\|x\|_2^2$.*

Lemma 3 (Regret of FTRL). *When Condition 1 holds, the sequence $\{x_\tau\}_{\tau=1}^{t}$ generated by the FTRL update $x_{\tau+1} \in \arg\min_{x \in D} R_\tau(x) + \sum_{i=1}^{\tau} F_i(x)$ has regret $\mathcal{R}_t = \sum_{i=1}^{t} F_i(x_i) - F_i(x) \leq \mathcal{O}(\sqrt{t})$ for all $x \in D$.*

Lemma 4 (σ_τ-Strongly Convex Regularizer). *When $\sum_{i=1}^{\tau} F_i(x)$ is σ_τ-strongly convex, $F_i(x)$ uniformly L_f-Lipschitz over D and $w_{\tau+1} \in \arg\min_{x \in D} \sum_{i=1}^{\tau} F_i(x)$, it holds $\|w_{\tau+1} - w_\tau\| \leq 2L_f/(\sigma_\tau + \sigma_{\tau-1})$*

4 Penalized FTRL

4.1 Exact Penalties for Time-Invariant Constraints

We begin by demonstrating the application of Lemma 1 to FTRL update (1) with time-invariant action set X. Selecting $F_i(x) = f_i(x) + \gamma h(x)$ with $h(x) = \sum_{j=1}^{m} \max\{0, g^{(j)}(x)\}$ and defining the bounded domain D with $X \subset D$, then by standard analysis, cf. [8], the `Pen-FTRL` update[1]:

$$x_{\tau+1} \in \arg\min_{x \in D} \left\{ R_\tau(x) + \sum_{i=1}^{\tau} F_i(x) \right\} \tag{4}$$

ensures regret $\sum_{i=1}^{t}(F_i(x_i) - F_i(x)) \leq \mathcal{O}(\sqrt{t})$ for all $x \in D$, and since $X \subset D$ for all $x \in X$. Of course this says nothing about whether the actions x_i lie in set

---

[1] Note the subtle yet crucial difference w.r.t. non-`Pen-FTRL` update (1).

X nor anything much about the regret of $f_i(x_i)$, but when set X has a Slater point and γ is selected large enough then by Lemma 1 we have that $x_{\tau+1} \in X$ for all τ. It follows that $F_i(x_i) = f_i(x_i)$ (since $h(x_i) = 0$ when $x_i \in X$) and so regret $\sum_{i=1}^{t}(F_i(x_i) - F_i(x)) = \sum_{i=1}^{t}(f_i(x_i) - f_i(x)) \leq \mathcal{O}(\sqrt{t})$ for all $x \in X$.

4.2   Penalties for Time-Varying Constraints

We now extend consideration to FTRL with time-varying constraints. Our aim is to define a penalty which is zero on a set $\hat{X}_\tau^{max} \approx X_{max}$, and large enough outside this set to force the minimum of $\sum_{i=1}^{\tau} F_i(x)$ to lie in $\hat{X}^{max}$.

Penalties Which Are Zero When $x \in \hat{X}_\tau^{max}$. Consider extending the penalty-based FTRL (4) to time-varying constraints. We might try selecting:

$$F_i(x) = f_i(x) + \gamma h_i(x), \quad \text{with} \quad h_i(x) = \sum_{j=1}^{m} \max\left\{0, g_i^{(j)}(x)\right\},$$

but we immediately run into the following difficulty. We have that $\sum_{i=1}^{\tau} F_i(x) = \sum_{i=1}^{\tau} f_i(x) + \gamma \sum_{i=1}^{\tau} \sum_{j=1}^{m} \max\{0, g_i^{(j)}(x)\}$ and so to make the second term zero requires $g_i^{(j)}(x) \leq 0$ for all $i \leq \tau$ and $j \leq m$, i.e. requires every constraint over all time to simultaneously be satisfied. This penalty choice $h_i(\cdot)$ therefore corresponds to benchmark X_t^{min}, whereas our interest is in set X_t^{max}. It is perhaps worth noting that this corresponds to the penalty used in the primal-dual literature, so it is unsurprising that those results are confined to X^{min}.

With this in mind, consider instead selecting

$$h_\tau(x) = \sum_{j=1}^{m} \max\left\{0, \sum_{i=1}^{\tau} g_i^{(j)}(x)\right\} - \sum_{j=1}^{m} \max\left\{0, \sum_{i=1}^{\tau-1} g_i^{(j)}(x)\right\}$$

with $h_1(x) = \sum_{j=1}^{m} \max\{0, g_i^{(j)}(x)\}$. Then,

$$\sum_{i=1}^{\tau} F_i(x) = \sum_{i=1}^{\tau} f_i(x) + \gamma \sum_{j=1}^{m} \max\left\{0, \sum_{i=1}^{\tau} g_i^{(j)}(x)\right\}.$$

We now have a sum-constraint in the second term, as desired. Unfortunately, this choice of $h_i(\cdot)$ violates the conditions needed for FTRL to achieve $\mathcal{O}(\sqrt{t})$ regret. Namely, it is required that $F_i(\cdot)$ is uniformly Lipschitz but $h_i(\cdot)$ does not satisfy this condition, and so neither does $F_i(\cdot)$. To see this, observe that when $g_i^{(j)}(\cdot)$ is uniformly Lipschitz with constant L_g, then $\sum_{i=1}^{\tau} g_i^{(j)}(x)$ has a Lipschitz constant τL_g that scales with τ, and so there exists no uniform upper bound. The max operator in $h_i(\cdot)$ does not change the Lipschitz constant (see Lemma 5); thus $h_i(\cdot)$ is τL_g Lipschitz, which prevents FTRL achieving $\mathcal{R}_t \leq \mathcal{O}(\sqrt{t})$.

These considerations lead us to the following penalty,

$$h_\tau(x) = \sum_{j=1}^{m} \max\left\{0, \frac{1}{\tau}\sum_{i=1}^{\tau} g_i^{(j)}(x)\right\}. \tag{5}$$

When $g_i^{(j)}(\cdot)$ is uniformly Lipschitz with constant L_g then so is $h_i(\cdot)$ due to the $1/\tau$ prefactor added to the sum and the following Lemma which just states that when a function $h(x)$ is L-Lipschitz then $\max\{0, h(x)\}$ is also L-Lipschitz:

Lemma 5. *When $|h(x)-h(y)| \leq L\|x-y\|$ then $|\max\{0, h(x)\}-\max\{0, h(y)\}| \leq L\|x-y\|$.*

Proof. Observe that $2\max\{0, h(x)\} = h(x)+|h(x)|$. Therefore, $2|\max\{0, h(x)\}-\max\{0, h(y)\}| = |h(x)-h(y)+|h(x)|-|h(y)|| \leq |h(x)-h(y)|+||h(x)|-|h(y)|| \leq |h(x)-h(y)|+|h(x)-h(y)| \leq 2L\|x-y\|$.

With this choice, we can write:

$$\sum_{i=1}^{\tau} F_i(x) = \sum_{i=1}^{\tau} f_i(x) + \gamma\sum_{j=1}^{m}\sum_{i=1}^{\tau}\max\left\{0, \frac{1}{i}\sum_{k=1}^{i} g_k^{(j)}(x)\right\}.$$

The second term is zero when

$$x_\tau \in \hat{X}_\tau^{max} := \left\{x \in D : \sum_{i=1}^{\tau} h_i(x) \leq 0\right\} = \left\{x : \sum_{k=1}^{i} g_k^{(j)}(x) \leq 0, j \leq m, i \leq \tau\right\}$$

Penalties Which Are Large When $x \notin \hat{X}_\tau^{max}$. In addition to requiring the penalty for time-varying constraints to be zero for $x \in \hat{X}_\tau^{max}$, we also require the penalty to be large enough when $x \notin \hat{X}_\tau^{max}$ so as to force the minimum of $\sum_{i=1}^{\tau} F_i(x)$ to lie in set $\hat{X}_\tau^{max}$, or at least to only result in $\mathcal{O}(\sqrt{\tau})$ violation.

As already noted, to use FTRL we need $F_i(\cdot)$ to be uniformly Lipschitz, which requires $f_i(\cdot)$ to be uniformly Lipschitz. When $f_i(\cdot)$ is L_f-Lipschitz then $|\sum_{i=1}^{\tau} f_i(x)|$ may grow linearly with τ at rate τL_f. We therefore require the penalty $\sum_{i=1}^{\tau} h_i(x)$ to also grow at least linearly with τ since otherwise for all τ large enough $|\sum_{i=1}^{\tau} f_i(x)| \gg \sum_{i=1}^{\tau} h_i(x)$ and the penalty may become ineffective i.e. we can have $x_\tau \notin \hat{X}_\tau$ for all τ large enough and so end up with $\mathcal{O}(t)$ constraint violation, which is no good.

We formalize the requirement the sum-penalty $\sum_{i=1}^{\tau} h_i(x)$ in (5) needs to grow quickly enough as follows. Let $\partial\hat{X}_\tau^{max}$ denote the boundary of $\hat{X}_\tau^{max}$, and

$$k_\tau := \min_{x \in \partial\hat{X}_\tau^{max}} \left|\left\{(i, j) : \frac{1}{i}\sum_{k=1}^{i} g_k^{(j)}(x) \geq 0, i = 1, \ldots, \tau, j = 1, \ldots, m\right\}\right|.$$

That is, k_τ is the minimum number of constraints active at the boundary of $\hat{X}_\tau^{max}$. Observe that $1 \leq k_\tau \leq \tau$ with, for example, $k_\tau = \tau$ when $g_i^{(j)}(x) = g^{(j)}(x)$ does not depend on i.

Condition 2 (Penalty Growth). *Let $z \in D$ be a common Slater point such that $\frac{1}{\tau} \sum_{i=1}^{\tau} g_i^{(j)}(z) < -\eta < 0$ for $j = 1, \ldots, m$ and $\tau > t_\epsilon$ (the same z must work for all τ and j). We require that $k_\tau \geq \frac{\beta}{\eta}\tau$ for all $\tau > t_\epsilon$, where $\beta > 0$ and the same β must work for all $\tau = 1, \ldots, t$.*

Time-Varying Exact Penalties. We are now in a position to extend the penalty approach to time-varying constraints. We begin by applying Lemma 1 to optimization problem P': $\min_{x \in D} f(x)$ s.t. $\frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(x) \leq 0$, $i = 1, \cdots, t$, $j = 1, \cdots, m$ where $f(\cdot)$ and $g_i^{(j)}(\cdot)$, $i = 1, \ldots, t$, $j = 1, \cdots, m$ are convex and $D \subset \mathbb{R}^n$ is convex and bounded. Let $C^* = \arg\min_{x \in \hat{X}_t^{max}} f(x)$. Define

$$H(x) := f(x) + \gamma \sum_{i=1}^{t} \sum_{j=1}^{m} \max\left\{0, \frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(x)\right\}$$

where $\gamma \in \mathbb{R}$. Note that $H(\cdot)$ is convex since $f(\cdot)$, $g_i^{(j)}(\cdot)$ are convex and composition with max preserves convexity.

Lemma 6. *Assume a Slater point exists, i.e. a $z \in D$ such that $\frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(z) < -\eta < 0$, $i = 1, \ldots, t$, $j = 1, \ldots, m$. Let $f^* := \min_{x \in \hat{X}_t^{max}} f(x)$. Then there exists a finite threshold $\gamma_0 \geq 0$ such that $H(x) \geq f^*$ for all $x \in D$, $\gamma \geq \gamma_0$, with equality only when $x \in \hat{X}_t^{max}$. It is sufficient to choose $\gamma_0 \geq \frac{f^* - f(z) - 1}{-k_t \eta}$.*

Proof. Setting the expression for γ_0 to one side for now, the result follows from applying Lemma 1 to P'. Turning now to expression $\gamma_0 \geq \frac{f^* - f(z) - 1}{k_t \eta}$, comparing this with the expression in Lemma 1, observe that the only change is in the denominator, which applying Lemma 1 to P' is $\max_{i \leq t, j \leq m}\{\frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(z)\} = -\eta$. Referring to (8) in the proof of Lemma 1, it is sufficient the denominator G of γ_0 is such that $\frac{\sum_{(i,j) \in A} \frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(z)}{G} \geq 1$, where $A \subset \{1, \ldots, t\} \times \{1, \ldots, m\}$. By assumption $g_k^{(j)}(z) \leq -\eta$ and so $\sum_{j \in A} \frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(z) \leq -|A|\eta$ with $|A| \geq 1$. Now $k_t \in [1, |A|]$, thus suffices to see setting $G = -k_t \eta$ also meets this requirement.

Theorem 1 (Time-Varying Exact Penalty). *The sequence x_τ, $\tau = 1, \ldots, t$ generated by the FTRL update (4) with $F_i(x) = f_i(x) + \gamma h_i(x)$ and $h_i(x) := \sum_{j=1}^{m} \max\{0, \frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(x)\}$ satisfies $x_{\tau+1} \in \hat{X}_\tau^{max}$ for $\tau > t_\epsilon$ when Condition 2 holds and parameter $\gamma > \frac{E + L + 1}{\beta}$ where $E \geq \max_{y \in D, i \in \{1, \ldots, t\}}(R_i(y) - R_i(z))/i$, $L \geq \max_{y \in D, i \in \{1, \ldots, t\}} f_i(y) - f_i(z)$ with $z \in D$ a Slater point.*

Proof. The result follows by application of Lemma 6 at times $\tau > t_\epsilon$ with $h(x) = R_\tau(x) + \sum_{i=1}^{\tau} f_i(x)$. We have that $h(x) - h(z) = R_\tau(x) - R_\tau(z) + \sum_{i=1}^{\tau}(f_i(x) - f_i(z)) \leq E\tau + L\tau$. Hence for $x_{\tau+1} \in \hat{X}_\tau^{max}$ it is sufficient to choose:

$$\gamma > \gamma_0 = \frac{(E + L)\tau - 1}{k_\tau \eta} \leq \frac{E + L + 1/\tau}{\beta} \leq \frac{E + L + 1}{\beta}.$$

When Condition 2 holds, $\beta > 0$.

Theorem 1 states a lower bound on γ in terms of constants E, L and β. For a quadratic regularizer $R_\tau(x) = \sqrt{\tau}\|x\|_2^2$ we can choose $E = \max_{y,z \in D}(\|y\|_2^2 - \|z\|_2^2)$. Since functions f_i are uniformly Lipschitz then $|f_i(z) - f_i(y)| \leq L_f\|z - y\| \leq L_f\|D\|$ and so we can choose $L = L_f\|D\|$. A value for β may be unknown but to apply Theorem 1 in practice we just need to select γ large enough, so a pragmatic approach is simply to make γ grow with time and then freeze it when it is large enough i.e. when the constraint violations are observed to cease.

4.3   Main Result: Penalized FTRL $\mathcal{O}(\sqrt{t})$ Regret & Violation

Our main result extends the standard FTRL analysis to time-varying constraints:

Theorem 2 (Penalized FTRL). *Assume Conditions 1 and 2 hold for $F_i(x) = f_i(x) + \gamma h_i(x)$ with $h_i(x) = \sum_{j=1}^m \max\{0, \frac{1}{i}\sum_{k=1}^i g_k^{(j)}(x)\}$, and the constraints $g_i^{(j)}$ are uniformly Lipschitz. Let the sequence of actions $\{x_\tau\}_{\tau=1}^t$ be generated by the* `Pen-FTRL` *update:*

$$x_{\tau+1} \in \arg\min_{x \in D} R_\tau(x) + \sum_{i=1}^\tau F_i(x) \tag{6}$$

Then, if γ is sufficiently large, the regret and constraint violation satisfy:

$$\mathcal{R}_t := \sum_{i=1}^t f_i(x_i) - f_i(y) \leq \mathcal{O}(\sqrt{t}), \qquad \mathcal{V}_t := \sum_{i=1}^t h_i(x_i) \leq \mathcal{O}(\sqrt{t}), \quad \forall y \in \hat{X}_t^{max}$$

$$\hat{X}_t^{max} = \Big\{ x \in D : \sum_{k=1}^i g_k^{(j)}(x) \leq 0, \forall i \leq t, j \leq m \Big\} = \Big\{ x \in D : \sum_{k=1}^i h_k(x) = 0, \forall i \leq t \Big\}$$

Proof. *Regret*: Applying Lemma 3 then $\sum_{i=1}^t F_i(x_i) - F_i(y) \leq \mathcal{O}(\sqrt{t})$ for all $y \in D$. This holds in particular for all $y \in \hat{X}_t^{max}$ and for these points $\sum_{i=1}^t F_i(y) = \sum_{i=1}^t f_i(y)$. Therefore, $\sum_{i=1}^t F_i(x_i) - f_i(y) \leq \mathcal{O}(\sqrt{t})$ i.e. $\mathcal{R}_t = \sum_{i=1}^t f_i(x_i) - f_i(y) \leq \mathcal{O}(\sqrt{t}) - \gamma\sum_{i=1}^t h_i(x_i) \leq \mathcal{O}(\sqrt{t})$ since $h_i(x_i) \geq 0$.

Constraint Violation: By Theorem 1, $x_{\tau+1} \in \hat{X}_\tau^{max}$ for $\tau > t_\epsilon$. Our interest is in bounding the violation of $\hat{X}_{\tau+1}^{max}$ by $x_{\tau+1}$. We can ignore the finite interval from 1 to t_ϵ since it will incur at most a finite constraint violation and so not affect an $\mathcal{O}(\sqrt{t})$ bound i.e. when obtaining the $\mathcal{O}(\sqrt{t})$ bound we can take $t_\epsilon = 0$. We follow a "Be-The-Leader" type of approach and apply Lemma 2 with $F_i(x) = h_i(x)$. We have that $h_i(x) \geq 0$ and by Condition 2, there exists a Slater point $z \in D$ such that $h_i(z) = 0$, $i = 1, \ldots, t$. Hence, $\min_{x \in D}\sum_{i=1}^\tau F_i(x) = 0$ and $\arg\min_{x \in D}\sum_{i=1}^\tau F_i(x)$ is not empty. Now, $x_{\tau+1} \in \hat{X}_\tau^{max} = \big\{x \in D : \sum_{i=1}^\tau h_i(x) = 0\big\} = \arg\min_{x \in D}\sum_{i=1}^\tau h_i(x)$ i.e. $x_{\tau+1}$ is a Follow-The-Leader update with respect to $\sum_{i=1}^\tau h_i(x)$. Hence, by Lemma 2, it is $\sum_{i=1}^t h_i(y) \geq \sum_{i=1}^t h_i(x_{i+1})$, $\forall y \in D$. Multiplying both sides of this inequality by -1 and adding $\sum_{i=1}^t h_i(x_i)$, it follows that:

$$\sum_{i=1}^t \Big(h_i(x_i) - h_i(y) \Big) \leq \sum_{i=1}^t \Big(h_i(x_i) - h_i(x_{i+1}) \Big) \quad \forall y \in D.$$

In particular, for $y \in \hat{X}_t^{max}$ then $\sum_{i=1}^{t} h_i(y) = 0$ and so

$$\mathcal{V}_t = \sum_{i=1}^{t} h_i(x_i) \le \sum_{i=1}^{t} \Big(h_i(x_i) - h_i(x_{i+1}) \Big).$$

Since $g_i^{(j)}$ is uniformly Lipschitz then by Lemma 5, we get that h_i is uniformly Lipschitz, i.e. $|h_i(x_i) - h_i(x_{i+1})| \le L_g \|x_i - x_{i+1}\|$ and $\mathcal{V}_t \le L_g \sum_{i=1}^{t} \|x_i - x_{i+1}\|$, where L_g is the Lipschitz constant. Since the regularizer $R_\tau(x)$ in the Pen-FTRL update is $\sqrt{\tau}$-strongly convex, by Lemma 4 we get that $\|x_i - x_{i+1}\|$ is $\mathcal{O}(1/\sqrt{i})$ and so $\sum_{i=1}^{t} \|x_i - x_{i+1}\|$ is $\mathcal{O}(\sqrt{t})$. Hence, $\mathcal{V}_t \le \mathcal{O}(\sqrt{t})$ as claimed.

We can immediately generalize Theorem 2 by observing that a sequence of constraints $\{g_i^{(j)}\}$ which are active at no more than $\mathcal{O}(\sqrt{t})$ time steps can be violated while still maintaining $\mathcal{O}(\sqrt{t})$ overall sum-violation.

Corollary 1 (Relaxation). *Define the sets*

$$P_- = \{j : \sum_{i=1}^{t} \max\{0, \frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(x)\} \le \mathcal{O}(\sqrt{t})\}, \ and \ P_+ = \{1, \ldots, m\} \setminus P_-.$$

In Theorem 2 relax Condition 2 so that it only holds for the subset P_+ of constraints. Then the Pen-FTRL update still ensures $\mathcal{O}(\sqrt{t})$ regret and constraint violation with respect to:

$$\hat{X}_t^{max} = \Big\{ x \in D : \sum_{k=1}^{i} g_k^{(j)}(x) \le 0, i = 1, \ldots, t, j \in P_+ \Big\}.$$

In effect, Corollary 1 says that we only need Condition 2 to hold for a *subset* of the constraints (i.e. subset P_+). The effect will be to increase the sum-violation, but only by $\mathcal{O}(\sqrt{t})$. This is the key advantage of the penalty-based approach, namely it allows a soft trade-off between sum-constraint satisfaction/violation, Condition 2 and benchmark set $\hat{X}_t^{max}$. Importantly, note that the Pen-FTRL update itself remains unchanged and does not require knowledge of the partitioning of constraints into sets P_+ and P_-.

With this in mind, it is worth noting that we also have the flexibility to partition the constraints in other ways. For example:

Corollary 2. *Consider the setup in Theorem 2 but using penalty*

$$h_i(x) = \sum_{j=1}^{m} \max \Big\{ 0, \frac{1}{i} \sum_{k=1}^{i} g_k^{(j)}(x) \Big\} + \delta_i^{(j)}(x)$$

Then the Pen-FTRL update ensures regret and violation

$$\mathcal{R}_t := \sum_{i=1}^{t} \Big(f_i(x_i) - f_i(y) \Big) \le \mathcal{O}(\sqrt{t}) - \sum_{i=1}^{t} \sum_{j=1}^{m} \Big(\delta_i^{(j)}(x_i) - \delta_i^{(j)}(y) \Big)$$

$$\mathcal{V}_t := \sum_{i=1}^{t} h_i(x_i) \le \mathcal{O}(\sqrt{t}) + \sum_{i=1}^{t} \sum_{j=1}^{m} \delta_i^{(j)}(x)$$

for all $y \in \hat{X}_t^{max} = \Big\{ x \in D : \sum_{k=1}^{i} g_k^{(j)}(x) \le 0, i = 1, \ldots, t, j = 1, \ldots, m \Big\}.$

When $\delta_i^{(j)} \leq \mathcal{O}(1/\sqrt{t})$ then Corollary 2 shows that the `Pen-FTRL` update achieves $\mathcal{O}(\sqrt{t})$ regret and violation, this Corollary will prove useful in the next section. Other variations of this sort are also possible.

4.4   Necessity of Penalty Growth Condition

Condition 2 is necessary for Theorems 1 and 2 to hold in the sense that when the condition is violated then there exist examples where these theorems fail.

Returning again to the example from the Introduction, selecting $h_i(x)$ according to (5) then $h_i(x) = \max\{0, -0.01\} + \max\{0, \frac{n_{2,i}}{i}x\} = \max\{0, \frac{n_{2,i}}{i}x\}$. Hence, the penalty is $\sum_{i=1}^{\tau} h_i(x) \leq \sum_{i=1}^{\tau} \frac{n_{2,i}}{i}x$. When $n_{2,i} < \mathcal{O}(i)$ then $\sum_{i=1}^{\tau} h_i(x) < \mathcal{O}(\tau)$ (since $\sum_{i=1}^{\tau} \frac{1}{i^c} \leq \int_0^{\tau} \frac{1}{i^c}di = \frac{\tau^{1-c}}{1-c}$ for $0 \leq c \leq 1$) and Condition 2 is violated (since $k_\tau \leq n_{2,\tau} < \mathcal{O}(\tau)$ and so there does not exist any $\beta > 0$ such that $k_\tau \geq \frac{\beta}{\eta}\tau$). For τ large enough the penalty $\sum_{i=1}^{\tau} h_i(x)$ therefore inevitably becomes small relative to $\sum_{i=1}^{\tau} f_i(x) = -2\tau x$, which leads to persistent violation of constraint $x \leq 0$ i.e. Theorem 1 fails. This is what we see in Fig. 2(a).

When $n_{2,i} \leq \mathcal{O}(\sqrt{i})$ then $\frac{n_{2,i}}{i} \leq \mathcal{O}(1/\sqrt{i})$ and the constraint sum-violation $\sum_{i=1}^{\tau} h_i(x) \leq \mathcal{O}(\sqrt{i})$. Hence, Corollary 1 still works even though Theorem 1 fails. However, when $n_{2,i}$ greater than $\mathcal{O}(\sqrt{i})$ but less than $\mathcal{O}(i)$ then the constraint violation is greater than $\mathcal{O}(\sqrt{i})$ and so Corollary 1 also fails.

Note that while we might consider gaining penalty growth by scaling γ with t, this is inadmissible because Condition 1 requires $F_t(x) = f_t(x) + \gamma h_t(x)$ to be uniformly Lipschitz, i.e., the same Lipschitz constant to apply at all times t.

4.5   Constraints Satisfying Penalty Growth Condition

A natural question to ask is which classes of time-varying constraints satisfy Condition 2. In this section we present some useful examples. In particular, we consider the classes of constraints considered by [11] and [14], since these are the only previous works for time-varying constraints that report $\mathcal{R}_t, \mathcal{V}_t = \mathcal{O}(\sqrt{t})$.

Perturbed Constraints. In [11] the considered constraints are of the form:

$$g_i^{(j)}(x) = g^{(j)}(x) + b_i^{(j)}$$

with common Slater point and $b_i^{(j)}$ upper bounded by some value, i.e., $b_i^{(j)} \leq \bar{b}^{(j)}, \forall i$. For this class of constraints we have that:

$$h_i(x) = \sum_{j=1}^{m} \max\left\{0, \frac{1}{i}\sum_{k=1}^{i}(g^{(j)}(x) + b_k^{(j)})\right\} = \sum_{j=1}^{m} \max\left\{0, g^{(j)}(x) + \frac{1}{i}\sum_{k=1}^{i} b_k^{(j)}\right\}$$

Defining $\underline{b}_t^{(j)} = \frac{1}{t} \sum_{k=1}^{t} b_k^{(j)}$ and $\Delta_i^{(j)}(x) = \frac{1}{i} \sum_{k=1}^{i} (b_k^{(j)} - \underline{b}_t)$, then we can rewrite the penalty equivalently as

$$h_i(x) = \sum_{j=1}^{m} \max \left\{ 0, g^{(j)}(x) + \underline{b}_t^{(j)} \right\} + \delta_i^{(j)}(x)$$

with $\delta_i^{(j)}(x) = \max \left\{ 0, g^{(j)}(x) + \underline{b}_t^{(j)} + \Delta_i^{(j)}(x) \right\} - \max \left\{ 0, g^{(j)}(x) + \underline{b}_t^{(j)} \right\}$. When $|\Delta_i^{(j)}(x)|$ is $\mathcal{O}(1/\sqrt{i})$ then, by Lemma 5, so is $|\delta_i^{(j)}(x)|$. Hence, when $|\Delta_i^{(j)}(x)|$ is $\mathcal{O}(1/\sqrt{i})$ then we can use the fact that Condition 2 holds for constraints $g^{(j)}(x) + \underline{b}_t^{(j)} \leq 0$ to show, by Corollary 2, that the Pen-FTRL update achieves $\mathcal{O}(\sqrt{t})$ regret and violation with respect to benchmark set $\hat{X}_t^{max} = \{x : g^{(j)}(x) + \underline{b}_t^{(j)} \leq 0\}$. This corresponds to one extreme of [11]'s benchmark but Theorem 1 provides more general conditions under which it is applicable, while [11] only considers constraints that are either time-invariant or i.i.d.

Alternatively, defining $\Delta_i^{(j)}(x) = \frac{1}{i} \sum_{k=1}^{i} (b_k^{(j)} - \bar{b}^{(j)})$, we can rewrite the penalty equivalently as:

$$h_i(x) = \sum_{j=1}^{m} \max \left\{ 0, g^{(j)}(x) + \bar{b}^{(j)} \right\} + \delta_i^{(j)}(x)$$

with $\delta_i^{(j)}(x) = \max\{0, g^{(j)}(x) + \bar{b}^{(j)} + \Delta_i^{(j)}(x)\} - \max\{0, g^{(j)}(x) + \bar{b}^{(j)}\}$. Observe that $\delta_i^{(j)}(x) \leq 0$ since $\Delta_i^{(j)}(x) \leq 0$. Hence, $\delta_i^{(j)}(x)$ does not add to the upper bound on the sum-constraint violation and so, by Corollary 2, the Pen-FTRL update achieves $\mathcal{O}(\sqrt{t})$ regret and violation with respect to benchmark set $\hat{X}_t^{max} = \{x : g^{(j)}(x) + \bar{b}^{(j)} \leq 0\}$. This corresponds to the other extreme of [11]'s benchmark, and in fact corresponds to the weak benchmark X_t^{min} and so is perhaps less interesting.

Families of Constraints. Suppose the time-varying constraint functions $g_i^{(j)}$ are selected from some family. That is, let $A^{(j)} = \{a_1^{(j)}, \ldots, a_{n_j}^{(j)}\}$ be a family of functions indexed by $k = 1, \ldots, n_j$ with $a_k^{(j)} : D \rightarrow \mathbb{R}$ being L_g-Lipschitz and $|a_k^{(j)}(x)| \leq a_{max}$ for all $x \in D$. At time i, constraint $g_i^{(j)} = a_k^{(j)}$ for some $k \in \{1, \ldots, n_j\}$, i.e. at each time step the constraint $g_i^{(j)}$ is selected from family $A^{(j)}$. Let $n_{k,\tau}^{(j)}$ denote the number of times that function $a_k^{(j)}$ is visited up to time τ and $p_{k,\tau}^{(j)} = n_{k,\tau}^{(j)}/\tau$ the fraction of times that $a_k^{(j)}$ is visited. With this setup the penalty is:

$$h_i(x) = \sum_{j=1}^{m} \max \left\{ 0, \frac{1}{i} \sum_{k=1}^{i} g_i^{(j)} \right\} = \sum_{j=1}^{m} \max \left\{ 0, \sum_{k=1}^{n_j} p_{k,i}^{(j)} a_k^{(j)}(x) \right\}.$$

We proceed by rewriting the penalty equivalently as

$$h_i(x) = \sum_{j=1}^{m} \max \left\{ 0, \sum_{k=1}^{n_j} p_k^{(j)} a_k^{(j)}(x) \right\} + \delta_i^{(j)}(x)$$

with $\delta_i^{(j)}(x) = \max\{0, \sum_{k=1}^{n_j} p_{k,i}^{(j)} a_k^{(j)}(x)\} - \max\{0, \sum_{k=1}^{n_j} p_k^{(j)} a_k^{(j)}(x)\}$ By Lemma 5, $|\delta_i^{(j)}(x)| \leq |\sum_{k=1}^{n_j} (p_{k,i}^{(j)} - p_k^{(j)}) a_k^{(j)}(x)|$. Assume the following condition holds:

Condition 3 ($1/\sqrt{t}$-Convergence). *For $\epsilon > 0$ there exists $t_0 > 0$ and $0 \leq p_k^{(j)} \leq 1$, $\sum_{j=1}^m \sum_{k=1}^{n_j} p_k^{(j)} = 1$ such that $|p_{k,\tau}^{(j)} - p_k^{(j)}| \leq \epsilon/\sqrt{\tau}$ for all $\tau > t_0$.*

Then for all $\tau > t_0$, $|\delta_i^{(j)}(x)| \leq n_j \frac{\epsilon}{\sqrt{\tau}} a_{max} \leq \bar{n} \frac{\epsilon}{\sqrt{\tau}} a_{max}$ with $\bar{n} := \max_j n_j$. By Corollary 2 it now follows that Pen-FTRL achieves $\mathcal{O}(\sqrt{t})$ regret and violation with respect to benchmark $\hat{X}_t^{max} = \{x : \sum_{k=1}^{n_j} p_k^{(j)} a_k^{(j)}(x) \leq 0\}$. Observe that in this case $\hat{X}_t^{max} = X_\infty^{max}$, i.e., we obtain $\mathcal{O}(\sqrt{t})$ regret and violation with respect to the strong benchmark, which is very appealing. Note that we don't need to know the relative frequencies in advance for this analysis to work.

Example. Suppose $D = [-10, 10]$, loss function $f_\tau(x) = -2x$ and constraint $g_\tau(x)$ alternates between $a_1(x) = -0.01$ and $a_2(x) = x$, equaling $a_2(x)$ at time τ with probability[2] $0.1c/\tau^{1-c}$. Figure 2(a) shows the performance vs c of the Pen-FTRL update with quadratic regularizer $R_\tau(x) = \sqrt{\tau} x^2$ and $F_\tau(x) = f_\tau(x) + \gamma \max\{0, p_{1,\tau} a_1(x) + p_{2,\tau} a_2(x)\}$ with parameter $\gamma = 25$. It can be seen that for $c = 1$ and $c = 0.5$ the constraint violation is well-behaved, staying close to zero, but for $c \in (0.5, 1)$ the constraint violation grows with time.

What is happening here is that when $c = 1$ then $p_{1,\tau} \to 0.9$, $p_{2,\tau} \to 0.1$ and the penalty term $\gamma \max\{0, p_{1,\tau} a_1(x) + p_{2,\tau} a_2(x)\}$ in $F_\tau(x)$ ensures the violation $\sum_{i=1}^t g_i(x) = t(p_{1,t} a_1(x) + p_{2,t} a_2(x))$ stays small. When $c = 0.5$, then $p_{1,\tau} \to 1$, $p_{2,\tau} \to 0$ and the penalty term ensures $t p_{1,t} a_1(x)$ stays small while $t p_{2,t} a_2(x)$ is $\mathcal{O}(\sqrt{t})$, thus $\sum_{i=1}^t g_i(x)$ is $\mathcal{O}(\sqrt{t})$. When $c \in (0.5, 1)$ then again $p_{1,\tau} \to 1$, $p_{2,\tau} \to 0$ and the penalty term ensures $t p_{1,t} a_1(x)$ stays small but now $t p_{2,t} a_2(x)$ is larger than $\mathcal{O}(\sqrt{t})$ and so $\sum_{i=1}^t g_i(x)$ is also larger than $\mathcal{O}(\sqrt{t})$.

We claim that $1/\sqrt{t}$-convergence is sufficient for Pen-FTRL to achieve $\mathcal{O}(\sqrt{t})$ regret and violation with respect to X^*, but it remains an open question whether or not it is also a necessary condition. Nevertheless, in simulations we observe that when $1/\sqrt{t}$-convergence does not hold then performance is often poor and that this is not specific to the FTRL algorithm, e.g. Figure 2(b) illustrates the performance of the canonical online primal-dual update (e.g. see [11]),

$$x_{t+1} = \Pi_D\left(x_t - \alpha_t(\partial f_t(x_t) + \lambda_t \partial g_t(x_t))\right), \quad \lambda_{t+1} = \left[\lambda_t + \alpha_t g_t(x_{t+1})\right]^+ \quad (7)$$

where Π_D denotes projection onto set D and step size $\alpha_t = 5/\sqrt{t}$.

I.i.d Stochastic Constraints. In [14] i.i.d. constraint functions drawn from a family are considered and a primal-dual algorithm is presented that achieves

---

[2] Recall that $c \sum_{\tau=0}^t \frac{1}{\tau^{1-c}} \approx c \int_0^t \frac{1}{\tau^{1-c}} d\tau = t^c$ for $0 \leq c \leq 1$. Hence, with this choice $E[n_{2,t}] \approx 0.1 t^c$ and $E[p_{2,t}] \approx 0.1 t^{c-1}$..

(a) Penalized FTRL

(b) Primal-Dual
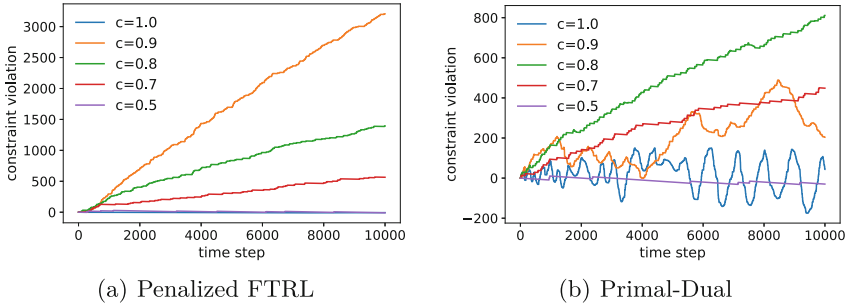
Fig. 2. Example about the role of $1/\sqrt{t}$-convergence in achieving $\mathcal{V}_t = \mathcal{O}(\sqrt{t})$.

$\mathcal{O}(\sqrt{t})$ regret and expected violation. Since with high probability the empirical mean converges at rate $1/\sqrt{t}$ with high probability, we can immediately apply the foregoing analysis to the sample paths to show that `Pen-FTRL` achieves $\mathcal{O}(\sqrt{t})$ regret and violation with respect to X_t^{max} with high probability. In more detail, let indicator random variable $I_{k,i}^j = 1$ when constraint function $a_k^{(j)}$ is selected at time i, and otherwise $I_{k,i}^{(j)} = 0$. By the law of large numbers (we can use any convenient concentration inequality, e.g. Chebyshev), with high probability the empirical mean satisfies $|\frac{1}{\tau} \sum_{i=1}^{\tau} I_{k,i}^{(j)} - p_k^j| \leq 1/\sqrt{\tau}$ with high probability. That is, Condition 3 holds with high probability and we are done.

Periodic Constraints. Let indicator $I_{k,i}^j = 1$ when constraint function $a_k^{(j)}$ is selected at time i, and otherwise $I_{k,i}^{(j)} = 0$. When the constraints are visited in a periodic fashion then

$$I_{k,i}^{(j)} = \begin{cases} 1 & i = nT_k^{(j)}, n = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

where $T_k^{(j)}$ is the period of constraint $a_k^{(j)}$. Then $|\frac{1}{\tau} \sum_{i=1}^{\tau} I_{k,i}^{(j)} - \frac{1}{T_k^{(j)}}| = \frac{1}{\tau}|\lfloor \frac{\tau}{T_k^{(j)}} \rfloor - \frac{\tau}{T_k^{(j)}}| \leq \frac{1}{\tau}$. Hence Condition 3 holds and we are done.

5   Summary and Conclusions

In this paper we extend the classical FTRL algorithm to encompass time-varying constraints by leveraging, for the first time in this context, the seminal penalty method of [12]. We establish sufficient conditions for this new `Pen-FTRL` algorithm to achieve $\mathcal{O}(\sqrt{t})$ regret and violation with respect to a strong benchmark $\hat{X}_t^{max}$ that expands significantly the previously-employed benchmarks in the literature. This result matches the performance of the best existing primal-dual algorithms in terms of regret and constraint violation growth rates, while substantially extending the class of problems covered. The key to this improvement

lies in how the time-varying constraints are incorporated into the FTRL algorithm. We conjecture that adopting a similar formulation with a primal-dual algorithm, namely using:

$$x_{t+1} = \Pi_D(x_t - \alpha_t(\partial f_t(x_t) + \lambda_t \partial h_t(x_t))), \ \lambda_{t+1} = [\lambda_t + \alpha_t h_t(x_{t+1})]^+$$

where $h_t(x) = \frac{1}{t}\sum_{i=1}^{t} g_i(x_{t+1})$, would allow similar performance to be achieved by primal-dual algorithms as by FTRL but we leave this to future work.

Appendix A: Proofs

A.1   Proof of Lemma 1

Proof. Firstly note that for feasible points $x \in X$ we have that $g^{(j)}(x) \leq 0$, $j = 1, \cdots, m$ and so $F(x) = f(x)$. By definition $f(x) \geq f^* = \inf_{x\in X} f(x)$ and so the stated result holds trivially for such points. Now consider an infeasible point $w \notin X$. Let z be an interior point satisfying $g^{(j)}(z) < 0$, $j = 1, \cdots, m$; by assumption such a point exists. Let $\gamma_0 = \frac{f^*-f(z)-1}{G}$. It is sufficient to show that $F(w) > f^*$ for $\gamma \geq \gamma_0$ and $G = \max_{j\in\{1,\cdots,m\}}\{g^{(j)}(z)\}$.

Let $v = \beta z + (1-\beta)w$ be a point on the chord between points w and z, with $\beta \in (0,1)$ and v on the boundary of X (that is $g^{(j)}(v) \leq 0$ for all $j = 1, \cdots, m$ and $g^{(j)}(v) = 0$ for at least one $j \in \{1, \cdots, m\}$). Such a point v exists since z lies in the interior of X and $w \notin X$. Let $A := \{j : j \in \{1, \cdots, m\}, g^{(j)}(v) = 0\}$ and $t(x) := f(x) + \gamma \sum_{j\in A} g^{(j)}(x)$. Then $t(v) = f(v) \geq f^*$. Also, by the convexity of $g^{(j)}(\cdot)$ we have that for $j \in A$ that $g^{(j)}(v) = 0 \leq \beta g^{(j)}(z) + (1 - \beta)g^{(j)}(w)$. Since $g^{(j)}(z) < 0$, it follows that $g^{(j)}(w) > 0$. Hence, $\sum_{j\in A} g^{(j)}(w) = \sum_{j\in A} \max\{0, g^{(j)}(w)\} \leq \sum_{j=1}^{m} \max\{0, g^{(j)}(w)\}$ and so $t(w) \leq F(w, \gamma)$. Now, observe that $t(z) = f(z) + \gamma \sum_{j\in A} g^{(j)}(z) \leq f(z) + \gamma_0 \sum_{j\in A} g^{(j)}(z)$ since $g^{(j)}(z) < 0$ and $\gamma \geq \gamma_0$. Hence,

$$t(z) \leq f(z) + (f^* - f(z) - 1)\frac{\sum_{j\in A} g^{(j)}(z)}{G} \tag{8}$$

Selecting G such that $\frac{\sum_{j\in A} g^{(j)}(z)}{G} \geq 1$ then $t(z) \leq f^* - 1 \leq t(v) - 1$. So we have established that $f^* \leq t(v)$, $t(z) \leq t(v) - 1$ and $t(w) \leq F(w)$. Finally, by the convexity of $t(\cdot)$, $t(v) \leq \beta t(z) + (1 - \beta)t(w)$. Since $t(z) \leq t(v) - 1$ it follows that $t(v) \leq \beta(t(v) - 1) + (1 - \beta)t(w)$ i.e. $t(v) \leq -\frac{\beta}{1-\beta} + t(w)$. Therefore $f^* \leq -\frac{\beta}{1-\beta} + F(w) < F(w)$ as claimed.

References

1. Anderson, D., Iosifidis, G., Leith, D.J.: Lazy lagrangians for optimistic learning With budget constraints. IEEE ACM Trans. Netw. (2023). IEEE
2. Chen, T., Ling, Q., Giannakis, G.B.: An online convex optimization approach to proactive network resource allocation. IEEE Trans. Signal Process. **65**(24), 6350–6364 (2017)
3. Hazan, E.: Introduction to online convex optimization. Found. Trends Optim. **2**, 157–325 (2016)
4. Jenatton, R., Huang, J.C., Archambeau, C.: Adaptive algorithms for online convex optimization with long-term constraints. In: Proceedings of ICML, pp. 402–411 (2016)
5. Liakopoulos, N., Destounis, A., Paschos, G., Spyropoulos, T., Mertikopoulos, P.: Cautious regret minimization: online optimization with long-term budget constraints. In: Proceedings of ICML, pp. 3944–3952 (2019)
6. Mahdavi, M., Jin, R., Yang, T.: Trading regret for efficiency: online convex optimization with long term constraints. J. Mach. Learn. Res. **13**(81), 2503–2528 (2012)
7. Mannor, S., Tsitsiklis, J.N., Yu, J.Y.: Online learning with sample path constraints. J. Mach. Learn. Res. **10**(20), 569–590 (2009)
8. McMahan, H.B.: A survey of algorithms and analysis for adaptive online learning. J. Mach. Learn. Res. **18**, 1–50 (2017)
9. Shalev-Shwartz, S.: Online learning and online convex optimization. Found. Trends Optim. **4**, 107–194 (2011)
10. Sun, W., Dey, D., Kapoor, A.: Safety-aware algorithms for adversarial contextual bandit. In: Proceedings of ICML, pp. 3280–3288 (2017)
11. Valls, V., Iosifidis, G., Leith, D., Tassiulas, L.: Online convex optimization with perturbed constraints: optimal rates against stronger benchmarks. In: Proceedings of AISTATS, pp. 2885–2895 (2020)
12. Zangwill, W.J.: Nonlinear programming via penalty functions. Manag. Sci. **13**(5), 344–358 (1967)
13. Yi, X., Li, X., Xie, L., Johansson, K.H.: Distributed online convex optimization with time-varying coupled inequality constraints. IEEE Trans. Signal Process. **68**, 731–746 (2020)
14. Yu, H., Nelly, M., Wei, X.: Online convex optimization with stochastic constraints. In: Proceedings of NIPS (2017)
15. Zinkevich, M.: Online convex programming and generalized infinitesimal gradient ascent. In: Proceedings of ICML (2003)