Penalised FTRL: FTRL With Time-Varying Constraints

Douglas J. Leith, George Iosifidis

1Trinity College Dublin, Ireland 2TU Delft, The Netherlands

Abstract. In this paper we extend the classical Follow-The-Regularised-Leader (FTRL) algorithm to encompass time-varying constraints. We establish sufficient conditions for the new Penalised FTRL algorithm to achieve $O(\sqrt{t})$ regret and violation with respect to a strong benchmark. This matches the performance of the best existing primal-dual algorithms, while substantially extending the class of problems covered.

Keywords: FTRL · online convex optimisation · constrained optimisation

1 Introduction

Follow-The-Regularised-Leader (FTRL), which includes online gradient descent and mixture of experts as special cases, is one of the standard algorithmic frameworks for online convex optimisation. The general form of the FTRL update is:

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

where action set $X \subset \mathbb{R}^n$ is bounded, function $F_i : X \rightarrow \mathbb{R}$ and regulariser $R_{\tau} : X \rightarrow \mathbb{R}$ is strongly convex.

When the sum-loss $\sum_{i=1}^{\tau} F_i(x)$ is convex and $F_i(x)$ and $R_{\tau}(x) - R_{\tau-1}(x)$ are uniformly Lipschitz then the sequence $x_{\tau}, \tau = 1, \ldots, t$ generated by the FTRL update has regret $\sum_{i=1}^{t}(F_i(x_i) - F_i(x)) \leq O(\sqrt{t})$ for all $x \in X$. The proof is standard. 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.

In this paper we extend the FTRL algorithm to accommodate time-varying action sets i.e. where at each time $\tau$ the fixed set action $X$ is replaced by set $X_{\tau}$ that may vary over time. We refer to this extension to FTRL as Penalised FTRL.

In general, it is too much to expect to be able to simultaneously achieve $O(\sqrt{t})$ 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 $O(\sqrt{t})$ regret and $O(\sqrt{t})$ constraint violation. That is, defining loss function $f_{\tau} : D \rightarrow \mathbb{R}$ on domain $D \subset \mathbb{R}^n$ and constraint functions $g_{\tau}^{(j)} : D \rightarrow \mathbb{R}$ such that $X_{\tau} = \{x \in D : g_{\tau}^{(j)}(x) \leq 0, j = 1, \ldots, m\}$ then we aim to simultaneously achieve
regret $\sum_{i=1}^{t}(f_i(x_i) - f_i(x)) \leq O(\sqrt{t})$ and violation $\sum_{i=1}^{t}(g_i^{(j)}(x_i) - g_i^{(j)}(x)) \leq O(\sqrt{t})$ for all $x \in X_t^{max} := \{x \in \mathbb{R}^n : \sum_{i=1}^{t} g_i^{(j)}(x) \leq 0, j = 1, \ldots, m\}$.

**Importance of $X_t^{max}$ Strong Benchmark.** We know from Mannor et al [5] that $O(\sqrt{t})$ regret and violation with respect to $X_t^{max}$ is not achievable for all sequences of constraint functions $\{g_i^{(j)}\}$. It is therefore necessary to (i) change the benchmark set $X_t^{max}$ to something more restrictive, (ii) restrict the admissible set of constraint function sequences $\{g_i^{(j)}\}$ or (iii) both. In the literature it is common to restrict the benchmark set to the weaker benchmark $X_t^{min} := \{x \in D : g_i^{(j)}(x) \leq 0, i = 1, \ldots, t, j = 1, \ldots, m\} \subset X_t^{max}$ i.e. to actions $x$ which simultaneously satisfy every constraint at every time. But this weak benchmark is typically so restrictive and easy to beat that in practice the achieved regret $R_t$ is often negative and indeed $-R_t \leq O(t)$.

One of our primary interests here is therefore in retaining a benchmark set that is close to $X_t^{max}$. To this end we consider benchmarks $X_t^{max}$ that are within $O(\sqrt{t})$ (in a sense to be defined later) of set $\{x \in D : \sum_{i=1}^{\tau} g_i^{(j)}(x) \leq 0, j = 1, \ldots, m, \tau = 1, \ldots, t\}$. This set requires the sum-constraint $\sum_{i=1}^{\tau} g_i^{(j)}(x) \leq 0$ to hold at every time $\tau \in \{1, \ldots, t\}$ rather than just at the end time $t$ and so is still smaller than $X_t^{max}$, but lacking predictions or prior knowledge of the constraints functions $g_i^{(j)}$ it is probably the best that we can reasonably hope for.

For example, suppose the time-varying constraint is $x \leq 1/\sqrt{t}$. Then $X_t^{min} = [0, 1/\sqrt{t}]$ which tends to set $\{0\}$ for $t$ large, $X_t^{max} = [0, 2\sqrt{t} - 2] \cap D$ which is set $D = [0, 1]$ for $t$ large, and $X_t^{max} = D = [0, 1]$.

**Role of Penalised FTRL.** Almost all of the literature focuses on using primal-dual algorithms to accommodate time-varying constraints. In contrast, here we use a direct penalty-based approach which we refer to as Penalised FTRL. This has the important advantages of (i) conceptually separating the issue of multiplier selection (i.e. $\lambda$ 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 Penalised FTRL extension to FTRL and (ii) establishing sufficient conditions for the Penalised FTRL algorithm to achieve $O(\sqrt{t})$ regret and violation with respect to benchmark $X_t^{max}$. This matches the performance of the best existing primal-dual algorithms, while substantially extending the class of problems covered.

2 Related Work

In summary, the literature on online learning with time-varying constraints focuses on the use of primal-dual algorithms similar to (7) and largely fails to obtain $O(\sqrt{t})$ regret and violation simultaneously, even with respect to the weak $X_t^{min}$ benchmark.

The standard setup consists of a sequence of convex cost functions $f_i : D \rightarrow \mathbb{R}$ and constraint functions $g_i^{(j)} : X \rightarrow \mathbb{R}, j = 1, \ldots, m$ and actions $x$ are selected...
from domain $D \subset \mathbb{R}^n$. The canonical algorithm considered in most papers is the following primal-dual gradient descent algorithm:

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

with step-size parameters $\eta_t$, $\mu_t$ and regularisation parameter $\theta_t$. Commonly, the parameter $\theta_t = 0$, exceptions being Mahdavi et al [4], Jenneton et al [2], Sun et al [9] where non-zero $\theta_t$ is used. Yu et al [8] approximate $g_t(x_{t+1})$ in the $\lambda_{t+1}$ update by $g_t(x_t) + \partial g_t(x_t)(x_{t+1} - x_t)$.

Regret bounds are mostly with respect to the baseline action set $X_T^{\text{min}} = \{x \in D : g_t^{(j)}(x) \leq 0, t = 1, \ldots, T, j = 1, \ldots, m\}$, the exception being Valls et al [10] where a slightly larger set is considered and Yu et al [13] where the constraints are restricted to be stochastic and the baseline action set is $\{x \in D : E[g_t^{(j)}(x)] \leq 0, j = 1, \ldots, m\}$.

The original work on this topic restricted attention to time-invariant constraint functions i.e. $g_t^{(j)}(x) = g_j(x)$. With this restriction, Jenneton et al achieve $O(\max\{T^\beta, T^{1-\beta}\})$ regret and $O(T^{1-\beta/2})$ constraint violation. Choosing $\beta = 2/3$ this gives $O(T^{2/3})$ regret and constraint violation. Similarly for the work of Mahdavi et al [4]. It is worth noting that these results of Mahdavi et al and Jenneton et al 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 $O(T^{1/2})$ regret and no constraint violation for constant contraints, e.g. see FTRL or, better still, the adaptive approach of Mohri et al [5].

For general time-varying cost and constraint functions, Sun et al [9] achieve $O(T^{1/2})$ regret and $O(T^{3/4})$ constraint violation. Liakopoulos et al [3] achieve $O(T^{1/2} + KT/V)$ regret and $O((VT)^{1/2})$ constraint violation, with $K = 1$ corresponding to baseline set $X_T^{\text{min}}$ and $V$ a design parameter. Selecting $V = T^{1/2}$ gives $O(T^{1/2})$ regret and $O(T^{3/4})$ constraint violation, similarly to Sun et al [9].

By restricting the constraint functions, Valls et al [10] improves this to $O(T^{1/2})$ regret and constraint violation. This requires restricting the constraints to be of the form $g_t^{(j)}(x) = g^{(j)}(x) - b_{t,j}$ with $b_{t,j} \in \mathbb{R}$ i.e. the constraints are $g^{(j)}(x) \leq b_{t,j}^{(j)}$ with time-variation confined to threshold $b_{t,j}^{(j)}$. Yu et al [13] also achieve $O(T^{1/2})$ regret and expected constraint violation (i.e. $E[\sum_{t=1}^T g_t^{(j)}(x_t)] \leq O(T^{1/2})$), this time by restricting the constraints to be i.i.d. stochastic. Yi et al [12] obtain $O(T^{2/3})$ regret and constraint violation by restricting the cost and constraint functions to be separable. Chen et al [1] focus on a form of dynamic regret that upper bounds the static regret and show $o(T)$ regret and $O(T^{2/3})$ 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 [11]. Consider the convex optimisation problem $P: \min_{x \in D} f(x)$ s.t. $g^{(j)}(x) \leq 0, j = 1, \ldots, m$ where
Fig. 1. Illustrating use of a penalty function to convert constrained optimisation
\[ \min_{g(x) \leq 0} f(x) \] into unconstrained optimisation \[ \min_x f(x) + \gamma \max\{0, g(x)\} \], \( \gamma > 0 \). 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 \( \gamma \) is selected large enough that out-with the feasible set \( f(x) + \gamma \max\{0, g(x)\} > f^* \), the min value of \( f \) inside the feasible set.

\[ D \subset \mathbb{R}^n, \quad f : \mathbb{R}^n \to \mathbb{R} \] 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 \( X^* \subset X \) denote set of optimal points. Define

\[ F(x) := f(x) + \gamma \sum_{j=1}^{m} \max\{0, g^{(j)}(x)\} \]

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

The basic idea here is that the penalty \( \sum_{j=1}^{m} \max\{0, g^{(j)}(x)\} \) is zero for \( x \in X \) but large for \( x \notin X \). Provided \( \gamma \) is selected large enough then the effect is therefore to force the minimum of \( F(x) \) to (i) lie in set \( X \) and (ii) match \( \min_{x \in X} f(x) \). This is illustrated schematically in Figure 1. The following lemma corresponds to [11, 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 optimisation \( P \) ). 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 \( \gamma_0 = \frac{f^*-f(z)-1}{\max_{j \in \{1, \cdots, m\}} (g^{(j)}(z))} \).

**Proof.** We include a proof of this lemma in the Appendix since we will make use of it later.

### 3.2 FTRL

We also recall the following standard FTRL results.

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

**Condition 1 (FTRL)** (i) Domain $D$ bounded (there is no need for convexity of $D$), (ii) $\sum_{i=1}^{t} F_i(x)$ is convex (the individual $F_i$’s need not be convex, only the sum), (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) $R_{t}(x)$ is $\sqrt{\tau}$-strongly convex and $R_t(x) - R_{t-1}(x)$ is uniformly Lipschitz e.g. $\sqrt{\tau} \|x\|_{2}$.

**Lemma 3 (Regret of FTRL).** When Condition 1 holds then 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 $\sum_{i=1}^{\tau} F_i(x) - F_i(x) \leq O(\sqrt{t})$ for all $x \in D$.

**Lemma 4 ($\sigma$-Strongly Convex Regulariser).** Suppose $\sum_{i=1}^{\tau} F_i(x)$ is $\sigma_{\tau}$-strongly convex, $F_i(x)$ is uniformly $L_f$-Lipschitz over $D$ and $w_{t+1} \in \arg \min_{x \in D} \sum_{i=1}^{t} F_i(x)$. Then $\|w_{t+1} - w_t\| \leq 2L_f/(\sigma_{\tau} + \sigma_{\tau-1})$

## 4 Penalised FTRL

### 4.1 Exact Penalties For Time-Invariant Constraints

We begin by demonstrating the application Lemma 1 to the 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 bounded domain $D$ with $X \subset D$ then by standard analysis the FTRL update

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

ensures regret $\sum_{i=1}^{\tau} (F_i(x_{\tau}) - F_i(x)) \leq O(\sqrt{t})$ 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 $X$ nor anything much about the regret of $f_i(x_i)$, but when set $X$ has a Slater point and $\gamma$ is selected large enough then by Lemma 1 we have that $x_{\tau+1} \in X$ for all $\tau$. It follows that $F_i(x_i) = f_i(x_i)$ (since $h(x_i) = 0$ when $x_i \in X$) and so regret $\sum_{i=1}^{\tau} (F_i(x_{\tau}) - F_i(x)) = \sum_{i=1}^{\tau} (f_i(x_{\tau}) - f_i(x)) \leq O(\sqrt{t})$ 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 $\hat{X}_{\max} \approx X_{\max}$, and large enough outside this set to force the minimum of $\sum_{i=1}^{\tau} F_i(x)$ to lie in set $\hat{X}_{\max}$.
Penalties Which Are Zero When \( x \in \hat{X}^{\text{max}} \) Now consider extending the penalty-based FTRL update (4) to time-varying constraints. We might try selecting, \( F_i(x) = f_i(x) + \gamma h_i(x) \) with \( h_i(x) = \sum_{j=1}^{m} \max\{0, g_{ij}^{(j)}(x)\} \) but we immediately run into the following difficulty. We have that \( \sum_{i=1}^{\tau} F_i(x) = \sum_{i=1}^{\tau} f_i(x) + \gamma \sum_{i=1}^{\tau} \sum_{j=1}^{m} \max\{0, g_{ij}^{(j)}(x)\} \) and so to make the second term zero requires \( g_{ij}^{(j)}(x) \leq 0 \) for all \( i = 1, \ldots, \tau \) and \( j = 1, \ldots, m \) i.e. requires every constraint over all time to simultaneously be satisfied. This choice of penalty \( h_i(\cdot) \) therefore corresponds to benchmark set \( X_i^{\text{min}} \), whereas our interest is in set \( X_i^{\text{max}} \). 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 mainly confined to benchmark set \( X_i^{\text{min}} \).

With this in mind, consider instead selecting \( h_\tau(x) = \sum_{j=1}^{m} \max\{0, \sum_{i=1}^{\tau} g_{ij}^{(j)}(x)\} - \sum_{j=1}^{m} \max\{0, \sum_{i=1}^{\tau-1} g_{ij}^{(j)}(x)\} \) with \( h_1(x) = \sum_{j=1}^{m} \max\{0, g_{ij}^{(j)}(x)\} \). Then, \( \sum_{i=1}^{\tau} F_i(x) = \sum_{i=1}^{\tau} f_i(x) + \gamma \sum_{i=1}^{\tau} \max\{0, \sum_{j=1}^{m} g_{ij}^{(j)}(x)\} \). 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 \( O(\sqrt{T}) \) 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 \( g_{ij}^{(j)}(\cdot) \) is uniformly Lipschitz with constant \( L_g \) then \( \sum_{i=1}^{\tau} g_{ij}^{(j)}(x) \) is Lipschitz with constant \( \tau L_g \) i.e. the Lipschitz constant scales with \( \tau \) and so there exists no uniform upper bound. The max operator in \( h_i(\cdot) \) doesn’t change the Lipschitz constant (see Lemma 5) and so \( h_i(\cdot) \) is \( \tau L_g \) Lipschitz, which prevents us obtaining \( O(\sqrt{T}) \) regret with FTRL.

These considerations lead us to the following penalty,

\[
  h_\tau(x) = \sum_{j=1}^{m} \max\{0, \frac{1}{\tau} \sum_{i=1}^{\tau} g_{ij}^{(j)}(x)\} \tag{5}
\]

When \( g_{ij}^{(j)}(\cdot) \) 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, \( 2 \max\{0, h(x)\} - \max\{0, h(y)\} = |h(x) - h(y)| + |h(x) - h(y)| - |h(x) - h(y)| \leq |h(x) - h(y)| + |h(x) - h(y)| \leq |h(x) - h(y)| + |h(x) - h(y)| \leq 2L \|x - y\| \).

With this choice,

\[
  \sum_{i=1}^{\tau} F_i(x) = \sum_{i=1}^{\tau} f_i(x) + \gamma \sum_{j=1}^{m} \sum_{i=1}^{\tau} \max\{0, \frac{1}{\tau} \sum_{k=1}^{i} g_{kj}^{(j)}(x)\} \tag{6}
\]

The second term is zero when \( x_\tau \in \hat{X}^{\text{max}} := \{ x \in D : \sum_{i=1}^{\tau} h_i(x) \leq 0 \} = \{ x : \sum_{k=1}^{i} g_{kj}^{(j)}(x) \leq 0, j = 1, \ldots, m, i = 1, \ldots, \tau \} \)
Penalties Which Are Large When \( x \notin \hat{X}_r^{\max} \) In addition to requiring the penalty for time-varying constraints to be zero for \( x \in \hat{X}_r^{\max} \) we also require the penalty to be large enough when \( x \notin \hat{X}_r^{\max} \) to force the minimum of \( \sum_{i=1}^r f_i(x) \) to lie in set \( \hat{X}_r^{\max} \), or at least to only result in \( O(\sqrt{\tau}) \) violation.

As already noted, to use FTRL we need \( f_i(\cdot) \) to be uniformly Lipschitz with \( \max \) preserves convexity. We are now in a position to extend the optimisation problem to time-varying constraints. Let \( \hat{X}_r^{\max} \) denotes the boundary of \( \hat{X}_r^{\max} \). Let \( k_r := \min_{x \in \hat{X}_r^{\max}} |\{ (i, j) : \frac{1}{r} \sum_{k=1}^i g_k^{(j)}(x) \geq 0, i \in \{1, \ldots, r\}, j \in \{1, \ldots, m\} \} | \). That is, \( k_r \) is the minimum number of constraints active at the boundary of \( \hat{X}_r^{\max} \). Observe that \( 1 \leq k_r \leq \tau \) with, for example, \( k_r = \tau \) 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 \( \frac{1}{\tau} \sum_{i=1}^r g_i^{(j)}(z) < -\eta < 0 \) for \( j = 1, \ldots, m \) and \( \tau > t_\epsilon \) (the same \( z \) must work for all \( \tau \) and \( j \)). We require that \( k_r \geq \frac{\beta}{\eta} \tau \) for all \( \tau > t_\epsilon \) where \( \beta > 0 \) and the same \( \beta \) 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 optimisation problem \( P' \): \( \min_{x \in D} f(x) \) s.t. \( \frac{1}{r} \sum_{k=1}^i g_k^{(j)}(x) \leq 0, i = 1, \ldots, t, j = 1, \ldots, m \) where \( f(\cdot) \) and \( g_i^{(j)}(\cdot) \) are convex and \( D \subset \mathbb{R}^n \) is convex and bounded. Let \( C^* = \arg \min_{x \in \hat{X}_r^{\max}} f(x) \). Define

\[
H(x) := f(x) + \gamma \sum_{i=1}^t \max_{j=1}^m \{ 0, \frac{1}{r} \sum_{k=1}^i g_k^{(j)}(x) \}
\]

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

**Lemma 6.** Assume that a Slater point exists, i.e. a \( z \in D \) such that \( \frac{1}{r} \sum_{i=1}^r g_i^{(j)}(z) < -\eta < 0, i = 1, \ldots, t, j = 1, \ldots, m \). Let \( f^* := \min_{x \in \hat{X}_r^{\max}} f(x) \). 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 \( x \in \hat{X}_r^{\max} \). It is sufficient to choose \( \gamma_0 \geq \frac{r-f(z)-1}{k_r \eta} \).

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

**Theorem 1 (Time-Varying Exact Penalty).** The sequence $x_\tau$, $\tau = 1, \ldots, t$ generated by the FTRL update (4) with $F_i(x) = f_i(x) + \gamma h_i(x)$ and $h_i(x) := \sum_{j=1}^{m} \max\{0, \frac{1}{t} \sum_{k=1}^{i} g_k^{(j)}(x)\}$ satisfies $x_{\tau+1} \in \hat{X}_\tau^{\max}$ for $\tau > t_e$ 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 $\tau > t_e$ 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 $\gamma > \frac{E+L+1}{\beta}$. When Condition 2 holds, $\beta > 0$.

Theorem 1 states a lower bound on $\gamma$ in terms of constants $E$, $L$ and $\beta$. For a quadratic regulariser $R_\tau(x) = \sqrt{\tau}\|x\|_2^2$ we can choose $E = \max_{y, z \in D} (\|y\|_2^2 - \|z\|_2^2)$. Since functions $f_i$ are uniformly Lipschitz then $|f_i(z) - f_i(y)| \leq L_i \|z - y\| \leq L_i \|D\|$ and we can choose $L = L_i \|D\|$. A value for $\beta$ may be unknown but to apply Theorem 1 in practice we just need to select $\gamma$ large enough, so a pragmatic approach is simply to make $\gamma$ 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: Penalised FTRL $O(\sqrt{t})$ Regret & Violation

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

**Theorem 2 (Penalised FTRL).** When Conditions 1 and 2 hold for $F_i(x) = f_i(x) + \gamma h_i(x)$ with $h_i(x) = \sum_{j=1}^{m} \max\{0, \frac{1}{t} \sum_{k=1}^{i} g_k^{(j)}(x)\}$, and the constraint functions $g_k^{(j)}$ are uniformly Lipschitz. Let the sequence of actions $x_\tau$, $\tau = 1, \ldots, t$ be generated by the Penalised FTRL update

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

(6)

Then when $\gamma$ is selected sufficiently large the regret and constraint violation satisfy

\[ R_\tau := \sum_{i=1}^{\tau} f_i(x_i) - f_i(y) \leq O(\sqrt{t}) \quad V_\tau := \sum_{i=1}^{\tau} h_i(x_i) \leq O(\sqrt{t}) \]

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

\[ \text{for all } y \in \hat{X}_\tau^{\max} \]
Proof. Regret: Applying Lemma 3 then \(\sum_{i=1}^{t} F_i(x_i) - F_i(y) \leq O(\sqrt{t})\) 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, \(\sum_{i=1}^{t} F_i(x_i) - f_i(y) \leq O(\sqrt{t})\) i.e. \(R_t = \sum_{i=1}^{t} f_i(x_i) - f_i(y) \leq O(\sqrt{t}) - \gamma \sum_{i=1}^{t} h_i(x_i) \leq O(\sqrt{t})\) since \(h_i(x_i) \geq 0\).

Constraint Violation: By Theorem 1, \(x_{t+1} \in \hat{X}_{t}^{max}\) for \(t > t_\epsilon\). Our interest is in bounding the violation of \(\hat{X}_{t+1}^{max}\) by \(x_{t+1}\). We can ignore the finite interval from 1 to \(t_\epsilon\) since it will incur at most a finite constraint violation and so not affect an \(O(\sqrt{t})\) bound i.e. when obtaining the \(O(\sqrt{t})\) 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, \(\min_{x \in D} \sum_{i=1}^{t} f_i(x) = 0\) and \(\arg \min_{x \in D} \sum_{i=1}^{t} f_i(x)\) is not empty. Now, \(x_{t+1} \in \hat{X}_{t}^{max} = \{x \in D : \sum_{i=1}^{t} h_i(x) = 0\}\) = \(\arg \min_{x \in D} \sum_{i=1}^{t} f_i(x)\) i.e. \(x_{t+1}\) is a Follow-The-Leader update with respect to \(\sum_{i=1}^{t} h_i(x)\). Hence, by Lemma 2, \(\sum_{i=1}^{t} h_i(y) \geq \sum_{i=1}^{t} h_i(x_{i+1})\) for every \(y \in D\). Multiplying both sides of this inequality by -1 and adding \(\sum_{i=1}^{t} h_i(x_i)\) it follows that,

\[
\sum_{i=1}^{t} (h_i(x_i) - h_i(y)) \leq \sum_{i=1}^{t} (h_i(x_i) - h_i(x_{i+1}))
\]

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

\[
\mathcal{V}_t = \sum_{i=1}^{t} h_i(x_i) \leq \sum_{i=1}^{t} (h_i(x_i) - h_i(x_{i+1}))
\]

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

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

**Corollary 1 (Relaxation).** Define \(P_\epsilon = \{j : \sum_{i=1}^{t} \max\{0, \frac{1}{t} \sum_{k=1}^{t} g^{(j)}(x)\} \leq O(\sqrt{t})\}\) and \(P_+ = \{1, \ldots, m\} \setminus P_\epsilon\). In Theorem 2 relax Condition 2 so that it only holds for the subset \(P_+\) of constraints. Then the Penalised FTRL update still ensures \(O(\sqrt{t})\) regret and constraint violation with respect to \(X_{t}^{max} = \{x \in D : \sum_{k=1}^{t} g^{(j)}(x) \leq 0, i = 1, \ldots, t, j \in P_+\}\).

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 \(O(\sqrt{t})\). 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 $X_{t}^{max}$. Importantly, note that the Penalised 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

$$h_{i}(x) = \sum_{j=1}^{m} \max \{0, \frac{1}{i} \sum_{k=1}^{i} g_{k}^{(j)}(x)\} + \delta_{i}^{(j)}(x)$$

Then the Penalised FTRL update ensures regret and violation

$$R_{t} := \sum_{i=1}^{t} f_{i}(x_{i}) - f_{i}(y) \leq O(\sqrt{t}) - \sum_{i=1}^{t} \sum_{j=1}^{m} (\delta_{i}^{(j)}(x_{i}) - \delta_{i}^{(j)}(y))$$

$$V_{t} := \sum_{i=1}^{t} h_{i}(x_{i}) \leq O(\sqrt{t}) + \sum_{i=1}^{t} \sum_{j=1}^{m} \delta_{i}^{(j)}(x)$$

for all $y \in X_{t}^{max} = \{x \in D : \sum_{k=1}^{i} g_{k}^{(j)}(x) \leq 0, i = 1, \ldots, t, j = 1, \ldots, m\}$.

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

### 4.4 Constraints Satisfying Penalty Growth Condition

A natural question to ask is which classes of time-varying constraint satisfy Condition 2. In this section we present some useful examples. In particular, we consider the classes of constraints considered by [10] and [13] since these are the only previous works for time-varying constraints that report $O(\sqrt{t})$ regret and violation.

**Perturbed Constraints** In [10] the constraint functions considered are of the form:

$$g_{i}^{(j)}(x) = g^{(j)}(x) + b_{i}^{(j)}$$

with common Slater point and $b_{i}^{(j)}$ upper bounded $b_{i}^{(j)} \leq \bar{b}^{(j)}$. For this class of constraints we have that

$$h_{i}(x) = \sum_{j=1}^{m} \max \{0, \frac{1}{i} \sum_{k=1}^{i} (g^{(j)}(x) + b_{k}^{(j)})\} = \sum_{j=1}^{m} \max \{0, g^{(j)}(x) + \frac{1}{i} \sum_{k=1}^{i} b_{k}^{(j)}\}$$
Defining $b^{(j)}_i = \frac{1}{t} \sum_{k=1}^t b^{(j)}_k$ and $\Delta^{(j)}_i(x) = \frac{1}{t} \sum_{k=1}^t (b^{(j)}_k - b^{(j)}_i)$ then we can rewrite the penalty equivalently as

$$h_i(x) = \sum_{j=1}^m \max \{0, g^{(j)}(x) + b^{(j)}_i\} + \delta^{(j)}_i(x)$$

with $\delta^{(j)}_i(x) = \max \{0, g^{(j)}(x) + b^{(j)}_i + \Delta^{(j)}_i(x)\} - \max \{0, g^{(j)}(x) + b^{(j)}_i\}$. When $|\Delta^{(j)}_i(x)|$ is $O(1/\sqrt{t})$ then, by Lemma 5, so is $|\delta^{(j)}_i(x)|$. Hence, when $|\Delta^{(j)}_i(x)|$ is $O(1/\sqrt{t})$ then we can use the fact that Condition 2 holds for constraints $g^{(j)}(x) + \bar{b}^{(j)}_i \leq 0$ to show, by Corollary 2, that the Penalised FTRL update achieves $O(\sqrt{t})$ regret and violation with respect to benchmark set $\hat{X}^{max}_t = \{x : g^{(j)}(x) + \bar{b}^{(j)}_i \leq 0\}$. This corresponds to one extreme of [10]'s benchmark but Theorem 1 provides more general conditions under which it is applicable (in [10] the only two conditions given are when the constraints are either time-invariant or i.i.d.).

Alternatively, defining $\Delta^{(j)}_i(x) = \frac{1}{t} \sum_{k=1}^t (b^{(j)}_k - \bar{b}^{(j)}_i)$ and we can rewrite the penalty equivalently as

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

with $\delta^{(j)}_i(x) = \max \{0, g^{(j)}(x) + \bar{b}^{(j)}_i + \Delta^{(j)}_i(x)\} - \max \{0, g^{(j)}(x) + \bar{b}^{(j)}_i\}$. Observe that $\delta^{(j)}_i(x) \leq 0$ since $\Delta^{(j)}_i(x) \leq 0$. Hence, $\delta^{(j)}_i(x)$ does not add to the upper bound on the sum-constraint violation and so, by Corollary 2, that the Penalised FTRL update achieves $O(\sqrt{t})$ regret and violation with respect to benchmark set $\hat{X}^{max}_t = \{x : g^{(j)}(x) + \bar{b}^{(j)}_i \leq 0\}$. This corresponds to the other extreme of [10]'s benchmark, and in fact corresponds to the weak benchmark $\hat{X}^{min}_t$ and so is perhaps less interesting.

Families Of Constraints Suppose the time-varying constraint functions $g^{(j)}_i$ are selected from some family. That is, let $A^{(j)} = \{a^{(j)}_1, \ldots, a^{(j)}_{n_j}\}$ be a family of functions indexed by $k = 1, \ldots, n_j$ with $a^{(j)}_k : D \to \mathbb{R}$ $L_g$-Lipschitz and $|a^{(j)}_k(x)| \leq a^{max}_j$ for all $x \in D$. At time $i$ constraint function $g^{(j)}_i = a^{(j)}_k$ for some $k \in \{1, \ldots, n_j\}$, i.e. at each time step the constraint function $g^{(j)}_i$ is selected from family $A^{(j)}$. Let $n^{(j)}_{k,\tau}$ denote the number of times that function $a^{(j)}_k$ is visited up to time $\tau$ and $p^{(j)}_{k,\tau} = n^{(j)}_{k,\tau}/\tau$ the fraction of times that $a^{(j)}_k$ is visited. With this setup the penalty is

$$h_i(x) = \sum_{j=1}^m \max \{0, \frac{1}{i} \sum_{k=1}^i g^{(j)}_k\} = \sum_{j=1}^m \max \{0, \sum_{k=1}^{n_j} p^{(j)}_{k,i} a^{(j)}_k(x)\}$$
We proceed by rewriting the penalty equivalently as

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

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

**Condition 3 (1/√τ-Convergence)** For $\epsilon > 0$ there exists $t_0 > 0$ and $0 \leq \frac{p_k^{(j)}}{\ell} \leq 1$, $\sum_{j=1}^{m} \sum_{k=1}^{n_j} p_k^{(j)} = 1$ such that $|p_k^{(j)} - p_k^{(j)}| \leq \epsilon/\sqrt{\tau}$ 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 Penalised FTRL achieves $O(\sqrt{\tau})$ 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 $O(\sqrt{\tau})$ 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 domain $D = [-10, 10]$, loss function $f_\tau(x) = -2x$ and constraint function $g_\tau(x)$ alternates between $a_1(x) = -0.01$ and $a_2(x) = x$, equalling $a_2(x)$ at time $\tau$ with probability $1 - 0.1c/\ell^{1-\epsilon}$. Figure 2(a) shows the performance vs $c$ of the Penalised FTRL update with quadratic regulariser $R_\tau(x) = \sqrt{\ell}x^2$ 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 $0.5 < c < 1$ 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 that the violation $\sum_{i=1}^{t} g_i(x) = t(p_{1,\tau}a_1(x) + p_{2,\tau}a_2(x))$ stays small. When $c = 0.5$ then $p_{1,\tau} \to 1$, $p_{2,\tau} \to 0$ and the penalty term ensures $tp_{1,\tau}a_1(x)$ stays small while $tp_{2,\tau}a_2(x)$ is $O(\sqrt{\ell})$, thus $\sum_{i=1}^{t} g_i(x)$ is $O(\sqrt{\ell})$. When $0.5 < c < 1$ then again $p_{1,\tau} \to 1$, $p_{2,\tau} \to 0$ and the penalty term ensures $tp_{1,\tau}a_1(x)$ stays small but now $tp_{2,\tau}a_2(x)$ is larger than $O(\sqrt{\ell})$ and so $\sum_{i=1}^{t} g_i(x)$ is also larger than $O(\sqrt{\ell})$.

We claim that $1/\sqrt{\ell}$-convergence is sufficient for Penalised FTRL to achieve $O(\sqrt{\tau})$ 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 $1/\sqrt{\ell}$-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 (c.f. see [10]),

$$x_{t+1} = P_D(x_t - \alpha_t \frac{df_t}{dx_t}(x_t) + \lambda_t \frac{dg_t}{dx_t}(x_t)), \lambda_{t+1} = [\lambda_t + \alpha_t g_t(x_{t+1})]^+$$ (7)

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

1 Recall that $c \sum_{\tau=0}^{t} \frac{1}{\sqrt{\tau}} \approx c \int_{0}^{1} \frac{1}{\sqrt{\tau}} d\tau = t^c$ for $0 \leq c \leq 1$. Hence, with this choice $E[n_2, t] \approx 0.1t^c$ and $E[p_{2,\tau}] \approx 0.1t^{c-1}$. 

Fig. 2. Example illustrating role of $1/\sqrt{t}$-convergence in achieving $O(\sqrt{t})$ constraint violation.

**I.i.d Stochastic Constraints** In [13] i.i.d. constraint functions drawn from a family are considered and a primal-dual algorithm is presented that achieves $O(\sqrt{t})$ 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 Penalised FTRL achieves $O(\sqrt{t})$ 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 $|\frac{1}{T} \sum_{i=1}^{T} I_{k,i}^{(j)} - p_{k}^{(j)}| \leq 1/\sqrt{T}$ 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 $|\frac{1}{T} \sum_{i=1}^{T} I_{k,i}^{(j)} - \frac{1}{T} T_{k}^{(j)}| = \frac{1}{T} |\| - \frac{T_{k}^{(j)}}{T_{k}^{(j)}} - \frac{T_{k}^{(j)}}{T_{k}^{(j)}}| \leq \frac{1}{T}$. 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. We establish sufficient conditions for the new Penalised FTRL algorithm to achieve $O(\sqrt{t})$ regret and violation with respect to a strong benchmark $X_{t}^{\max}$. This matches the performance of the best existing primal-dual algorithms, 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:

\[ x_{t+1} = P_D(x_t - \alpha_t \left( \frac{df_t}{dx}(x_t) + \lambda_t \frac{dh_t}{dx}(x_t) \right)), \quad \lambda_{t+1} = [\lambda_t + \alpha_t h_t(x_{t+1})]^+ \]

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.

References


Appendix A: Proofs

5.1 Proof of Lemma 1

Proof. Firstly note that for feasible points $x \in X$ we have that $g^{(j)}(x) \leq 0$, $j = 1, \ldots, m$ 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 $w \not\in X$. Let $z$ be an interior point satisfying $g^{(j)}(z) < 0$, $j = 1, \ldots, m$; by assumption such a point exists. Let $\gamma_0 = \frac{f^* - f(z)}{G}$. It is sufficient to show that $F(w) > f^*$ for $\gamma \geq \gamma_0$ and $G = \max_{j \in \{1, \ldots, m\}} \{g^{(j)}(z)\}$.

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 $j = 1, \ldots, m$ and $g^{(j)}(v) = 0$ for at least one $j \in \{1, \ldots, m\}$). Such a point $v$ exists since $z$ lies in the interior of $X$ and $w \not\in X$. Let $A := \{j : j \in \{1, \ldots, m\}, g^{(j)}(v) = 0\}$ 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,

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

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.