Online Learning under Adversarial Nonlinear Constraints

Anonymous Author(s)
Affiliation
Address
email

Abstract

In many applications, learning systems are required to process continuous non-stationary data streams. We study this problem in an online learning framework and propose an algorithm that can deal with adversarial time-varying and nonlinear constraints. As we show in our work, the algorithm called Constraint Violation Velocity Projection (CVV-Pro) achieves $\sqrt{T}$ regret and converges to the feasible set at a rate of $1/\sqrt{T}$, despite the fact that the feasible set is slowly time-varying and a priori unknown to the learner. CVV-Pro only relies on local sparse linear approximations of the feasible set and therefore avoids optimizing over the entire set at each iteration, which is in sharp contrast to projected gradients or Frank-Wolfe methods. We also empirically evaluate our algorithm on two-player games, where the players are subjected to a shared constraint.

1 Introduction

Today's machine learning systems are able to combine computation, data, and algorithms at unprecedented scales, which opens up new and exciting avenues in many domains, such as computer vision, computer graphics, speech and text recognition, and robotics [Jordan and Mitchell, 2015]. One of the leading principles that has enabled this progress is the focus on relatively simple pattern recognition and empirical risk minimization approaches, which mostly rely on offline gradient-based optimization and stipulate that the training, validation, and test data are independent and identically distributed.

Somewhat overlooked in these developments is the role of non-stationarity and constraints [Jordan, 2019]. Indeed, emerging machine learning problems involve decision-making in the real world, which typically includes interactions with physical, social, or biological systems. These systems are not only time varying and affected by past interactions, but their behavior is often characterized via fundamental constraints. Examples include cyber-physical systems where constraints are imposed by the laws of physics, multi-agent systems that are subjected to a shared resource constraint, or a reinforcement learning agent that is subjected to safety and reliability constraints. In particular, in their seminal work Auer et al. [2002] gave a reduction for the multi-arm bandit setting to the full information online optimization setting, by employing the multiplicative weights framework [Littlestone and Warmuth, 1994]. This classical reduction was recently extended by Sun et al. [2017] to the contextual bandit setting with sequential (time-varying) risk constraints.

This motivates our work, which is in line with a recent trend in the machine learning community towards online learning, adaptive decision-making, and online optimization. More precisely, we study an online problem with slowly time-varying constraints, where in each time step t, the learner receives partial information on the current cost f_t and feasible set $\mathcal{C}_t := \{x \in \mathbb{R}^n \mid g_t(x) \geq 0\}$. The learner makes a decision x_t and incurs the loss $f_t(x_t)$. The quality of the learner's decision making

Submitted to 16th European Workshop on Reinforcement Learning (EWRL 2023). Do not distribute.

is measured by comparing to the best decision in hindsight, that is,

$$\sum_{t=1}^{T} f_t(x_t) - \min_{x^* \in \mathcal{C}_T} \sum_{t=1}^{T} f_t(x^*) \quad \text{subject to} \quad g_T(x_T) \geq -\frac{c}{\sqrt{T}}, \tag{1}$$

which will be shown to be bounded by $\mathcal{O}(\sqrt{T})$ for our algorithm. The functions f_t and g_t are restricted to $f_t \in \mathcal{F}$ and $g_t \in \mathcal{G}$ (as defined in Assumption 1.1) and $c > 0$ is an explicit constant. With minor modifications (slightly tightening the constraints), our algorithm also achieves $\mathcal{O}(\sqrt{T})$ regret in (1) for $c = 0$. It is important to note that our performance objective (1) is symmetric in the sense that the constraint $x \in \mathcal{C}_T$ applies to both the learner and the benchmark x^*. This contrasts prior work (see, e.g., Chen et al. [2017], Neely and Yu [2017], Yu et al. [2017], Sun et al. [2017], Liu et al. [2022]) where a different notion of constraint violations $\sum_{t=1}^{T} g_t(x_t) \geq -c\sqrt{T}$ is used for the learner, while the benchmark x^* is required to satisfy $g_t(x^*) \geq 0$ for all $t \in \{1, \ldots, T\}$. Unlike (1), this leads to an asymmetric regret formulation, since different requirements are imposed on the learner and the benchmark x^*.

Even more intriguing is the fact that our algorithm is unaware of the feasible sets a-priori, and obtains, at each iteration, only a local sparse approximation of $\mathcal{C}_t$ based on the first-order information of the violated constraints. The indices of all violated constraints at x_t will be captured by the index set $I(x_t) := \{i \in \{1, \ldots, m\} \mid g_{t,i}(x_t) \leq 0\}$, while $G(x_t) := [\nabla g_{t,i}(x_t)]_{i \in I(x_t)}$ denotes the matrix whose columns store the corresponding gradients. In order to guarantee a regret of $\mathcal{O}(\sqrt{T})$ in (1) we require the following assumptions.

Assumption 1.1. There exists $R > 0$ such that **1)** $\mathcal{F}$ is a class of convex functions, where every $f \in \mathcal{F}$ satisfies $||\nabla f(x)|| \leq L_{\mathcal{F}}, \forall x \in \mathcal{B}_{4R}$, with $||\cdot||$ the ℓ_2 norm and $\mathcal{B}_R$ the hypersphere of radius R centered at the origin; **2)** $\mathcal{G}$ is a class of concave $\beta_{\mathcal{G}}$-smooth functions, where every g satisfies $||\nabla g(x)|| \leq L_{\mathcal{G}}, \forall x \in \mathcal{B}_{4R}$; **3)** The feasible set $\mathcal{C}_t$ is non-empty and contained in $\mathcal{B}_R$ for all t.

We note that these assumptions are standard in online optimization [Hazan, 2016, Ch. 3]. The learner's task is nontrivial even in the case where the feasible set is time invariant. If the feasible set is time varying, additional assumptions are required that restrict the amount that the feasible set is allowed to change. These two assumptions, see i) and ii) below, are described by the following interaction protocol between the learner and the environment:

Assumption 1.2. (Interaction protocol) At each time step $t \in \{1, \ldots, T\}$:
1) the learner chooses x_t;
2) the environment chooses $f_t \in \mathcal{F}$ and $g_t \in \mathcal{G}$ such that i) $||g_t - g_{t-1}||_\infty = \mathcal{O}(1/t)$, with $||\cdot||_\infty$ the ℓ_∞ norm, and ii) $\mathcal{C}_t$ is contained in $\mathcal{Q}_t := \cap_{\ell=0}^{t-1} \mathcal{S}_\ell$, where $\mathcal{S}_t := \{x \in \mathbb{R}^n \mid G(x_t)^\top (x - x_t) \geq 0\}$ is a cone centered at x_t for $t \geq 1$ and $\mathcal{S}_0 = \mathbb{R}^n$ (the situation is illustrated in Figure 1);
3) the environment reveals to the learner partial information on cost $f_t(x_t)$, $\nabla f_t(x_t)$ and all violated constraints $g_{t,i}(x_t)$, $\nabla g_{t,i}(x_t)$ for $i \in I(x_t)$.

The requirements i) $||g_t - g_{t-1}||_\infty = \mathcal{O}(1/t)$; and ii) $\mathcal{C}_t \subset \mathcal{Q}_t$ restrict the feasible sets that the environment can choose. We note that despite the fact that $||g_t - g_{t-1}||_\infty = \mathcal{O}(1/t)$, $||g_t - g_1||_\infty = \mathcal{O}(\ln(t))$, which means that the sequence of functions g_t that defines $\mathcal{C}_t$ does *not* converge in general. As a result, C_t may evolve in such a way that the initial iterates $x_1, x_2, \ldots, x_{t_0}$ achieve a large cost compared to $\min_{x^* \in \mathcal{C}_T} \sum_{t=1}^{T} f_t(x^*)$, as these are constrained by the sets $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{t_0}$, which may be far away from $\mathcal{C}_T$. The second requirement ii) $\mathcal{C}_t \subset \mathcal{Q}_t$ avoids this situation and is therefore key for obtaining an $\mathcal{O}(\sqrt{T})$ regret.

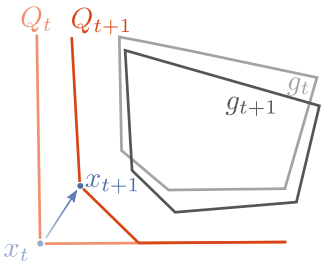

Figure 1: At each time step, the feasible set $\mathcal{C}_t$ contained in a polyhedral intersection $\mathcal{Q}_t$ changes slightly and is only partially revealed.

Our setup differs from traditional online convex optimization in the following two important ways: i) The environment chooses not only the functions f_t but also the nonlinear constraint functions g_t, ii) even if g_t is time-invariant, i.e., $g_t = g$ for all t the learner has only access to local information about the feasible set. That is, the information about the feasible set is only revealed piece-by-piece and needs to be acquired by the agent through repeated queries of a constraint violation oracle.

We propose an online algorithm that despite the lack of information about the feasible set, achieves $\mathcal{O}(\sqrt{T})$ regret, and will derive explicit non-asymptotic bounds for the regret and the convergence to $\mathcal{C}_T$. We thus conclude that our algorithm matches the performance of traditional online projected gradients or Frank-Wolfe schemes, while requiring substantially less information about the feasible set and allowing it to be time-varying. Perhaps equally important is the fact that instead of performing projections onto the full feasible set at each iteration, our algorithm only optimizes over a local sparse linear approximation. If constraints are nonlinear, which includes norm-constraints or constraints on the eigenvalues of a matrix, optimizing over the full feasible set at each iteration can be computationally challenging.

1.1 Related Work

Online learning has its roots in online or recursive implementations of algorithms, where due to the piece-by-piece availability of data, algorithms are often analyzed in a non i.i.d. setting. A central algorithm is the multiplicative weights scheme [Freud and Schapire, 1997], where a decider repeatedly chooses between a finite or countable number of options with the aim of minimizing regret. This online learning model not only offers a unifying framework for many classical algorithms [Blum, 1998], but represents a starting point for online convex optimization [Hazan, 2016], and adversarial bandits [Lattimore and Szepesvári, 2020]. Our approach extends this line of work by allowing the environment to not only choose the objective functions f_t, but also the constraints g_t. Due to the fact that our learner only obtains local information about the feasible set, our work is somewhat related to Levy and Krause [2019], Lu et al. [2022], Garber and Kretzu [2022], Mhammedi [2022], where the aim is to reduce the computational effort of performing online projected gradient steps or Frank-Wolfe updates. More precisely, Levy and Krause [2019] propose an algorithm that directly approximates projections, while requiring multiple queries of the constraint functions and their gradients. A slightly different constraint violation oracle is assumed in Garber and Kretzu [2022], where the learner can query separating hyperplanes between a given infeasible point and the feasible set. Algorithmically, both Garber and Kretzu [2022] and Levy and Krause [2019] depart from online gradient descent, where the latter computes projections via an approximate Frank-Wolf-type scheme. An alternative is provided by Mhammedi [2022] and Lu et al. [2022], where optimizations over the entire feasible set are simplified by querying only a set membership oracle based on the Minkowski functional. While our approach also avoids projections or optimizations over the entire feasible set, we introduce a different constraint violation oracle that returns a local sparse linear approximation of the feasible set. We call the constraint violation oracle only once every iteration and do not require a two-step procedure that involves multiple oracle calls. In addition, we also allow for adversarial time-varying constraints.

In addition, there has been important recent work that developed online optimization algorithms with constraints. In contrast to the primal formulation of our algorithm, these works are based on primal-dual formulations, where the algorithm is required to satisfy constraints on average, so called long-term constraints. The research can be divided into two lines of work Mahdavi et al. [2012], Jenatton et al. [2016], Yu and Neely [2020] and Yuan and Lamperski [2018], Yi et al. [2021] that use a set of weaker and stricter definitions for constraint violations and investigate time-invariant constraints, which contrasts our formulation that includes time-varying constraints. A third line of work by Chen et al. [2017], Neely and Yu [2017], Yu et al. [2017], Sun et al. [2017], Liu et al. [2022] focuses on time-varying constraints, where, however, the following weaker notion of constraint violation is used: $\sum_{t=1}^{T} g_t(x_t) \geq -c\sqrt{T}$, where t refers to time and x_t to the learner's decision. This metric allows constraint violations for many iterations, as long as these are compensated by strictly feasible constraints (in the worst case even with a single feasible constraint with a large margin). In contrast, our algorithm satisfies $g_t(x_t) \geq -c/\sqrt{t}$ for all iterations $t \in \{1, \ldots, T\}$, where c is an explicit constant independent of the dimension of the decision variable and the number of constraints. This means that we can explicitly bound the constraint violation at every iteration, whereas infeasible and strictly feasible iterates cannot compensate each other.

An important distinction to Neely and Yu [2017] is given by our performance metric (see also the discussion in Neely and Yu [2017] and Liu et al. [2022]). The work from Neely and Yu [2017] uses $\sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x^\star)$ as a performance measure, where the iterates x_t are required to satisfy $\sum_{t=1}^{T} g_t(x_t) \geq -c\sqrt{T}$ and $x^\star$ satisfies $g_t(x^\star) \geq 0$ *for all* $t \in \{1, \ldots, T\}$. This leads to a major asymmetry in the way regret is measured: while the iterates of the online algorithm only need to

satisfy a cumulative measure of constraint violation, the benchmark $x^\star$, which represents the best fixed decision in hindsight, is required to satisfy *all* constraints $g_t(x^\star) \geq 0$ for $t = \{1, \ldots, T\}$. The performance metric introduced in (1) is symmetric and imposes the same constraints on the learner as well as the benchmark x^*. These features make our algorithm a valuable addition to the algorithmic toolkit of online constrained optimization, which has also potential applications in bandit problems and related fields.

An important special case of our online learning model arises when the environment is represented by an adversarial player that competes with the learner. This corresponds to a repeated generalized Nash game due to the constraint that couples the decisions of the learner and its adversary. If the adversary plays best response, the resulting equilibria are characterized by quasi-variational inequalities [Facchinei and Kanzow, 2007] and there has been important recent work, for example by Jordan et al. [2023], Kim [2022], Facchinei and Kanzow [2010] that proposes different gradient and penalty methods for solving these inequalities. Our approach adopts a different perspective, rooted in online learning, which allows us to derive non-asymptotic convergence results for a first-order gradient-based algorithm that can be implemented in a straightforward manner. Our approach is also inspired by the recent work of Muehlebach and Jordan [2022], who propose a similar algorithm for the offline setting.

1.2 Main Contributions

We give an online optimization scheme under *unknown* non-linear constraints that achieves an optimal $\mathcal{O}(\sqrt{T})$ regret and converges to the latest feasible set at a rate of $\mathcal{O}(1/\sqrt{T})$. There are two variants of our problem formulation: The first deals with situations where constraints are unknown but fixed, the second allows constraints to be chosen in a time-varying and adversarial manner.

Our algorithm, named Constraint Violation Velocity Projection (CVV-Pro), has the following features:

1. It assumes access to a new type of oracle, which on input x_t, returns partial information on all currently violated constraints. Namely, the value $g_{t,i}(x_t)$ and the gradient $\nabla g_{t,i}(x_t)$ for *all* $i \in I(x_t)$.

2. It projects an adversarially generated negative cost gradient $-\nabla f_t(x_t)$ onto a velocity polytope $V_\alpha(x_t) := \left\{ v \in \mathbb{R}^n \mid [\nabla g_i(x_t)]^\top v \geq -\alpha g_i(x_t),\ \forall i \in I(x_t) \right\}$. Due to the linear and local structure of $V_\alpha(x_t)$, the projection can be computed efficiently.

3. In contrast to standard online methods that project in each round a candidate decision onto the feasible set, our method trades off feasibility for efficiency. In particular, it produces a sequence of decisions that converges at a rate of $\mathcal{O}(1/\sqrt{T})$ to the latest feasible set.

4. Our method handles time-varying adversarial constraints g_t, provided a decreasing rate of change $||g_{t+1} - g_t||_\infty \leq \mathcal{O}(1/t)$ and that each feasible set $\mathcal{C}_t$ belongs to $\mathcal{Q}_t$ (see Assumption 1.2). As we show in Section 3.1, an important special case where the assumption of decreasing rate of change is satisfied is given by $g_t = \frac{1}{t}\sum_{j=1}^{t} \tilde{g}_j$, i.e., when g_t represents an average of constraints $\tilde{g}_t$ over time.

1.3 Outline

Section 2 describes our algorithm and considers the situation where g_t is time invariant. This sets the stage for our main results in Section 3 that provide regret guarantees for our new online convex optimization setting with non-stationary, nonlinear, and unknown constraints. An important and interesting application of our algorithm are generalized Nash equilibrium problems, as will be illustrated with a numerical experiment in Section 4. The experiment will also highlight that the numerical results agree with the theoretical predictions.

2 Online learning under unknown, time-invariant, and nonlinear constraints

2.1 Online Gradient Descent

Online gradient descent [Hazan, 2016, Ch. 3.1] is a classical and perhaps the simplest algorithm that achieves optimal $\mathcal{O}(\sqrt{T})$ regret for the setting of a compact, convex, time-invariant, and a priori known feasible set. It consists of the following two operations: i) $y_{t+1} = x_t - \eta_t \nabla f_t(x_t)$ takes a step

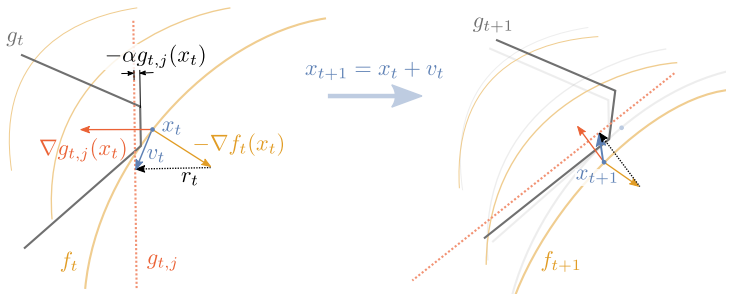

Figure 2: Illustration of the proposed (CVV-Pro) algorithm. **Left:** the constraint $g_{t,j}$ is violated by the current solution x_t. The cost gradient $-\nabla f_t(x_t)$ is projected onto the hyperplane (moved by $-\alpha g_{t,j}(x_t)$) with normal vector $\nabla g_{t,j}(x_t)$. This yields r_t (see Section 2.5), and results in the velocity projection v_t ($\eta = 1$ for clarity). **Right:** next iteration with updated x, where both f and g are changed. Then the procedure is applied recursively.

from the previous point in the direction of the previous cost gradient; and ii) $x_{t+1} = \mathrm{Proj}_{\mathcal{C}}(y_{t+1})$ projects y_{t+1} back to the feasible set $\mathcal{C}$, as y_{t+1} may be infeasible.

In this section, we generalize the online gradient descent algorithm to the setting where the feasible set is unknown a priori and has to be learned through repeated queries of a constraint violation oracle that only reveals local information.

2.2 Overview

In Section 2.3, we present the pseudo code of our algorithm. In Section 2.4, we give a structural result showing that Algorithm 1 under Assumption 1.1 and a bounded iterate assumption guarantees an optimal $\mathcal{O}(\sqrt{T})$ regret and converges to the feasible set at a rate of $\mathcal{O}(1/\sqrt{T})$. In Appendix C, we show that the bounded iterate assumption can be enforced algorithmically, by introducing an additional hypersphere constraint that attracts the sequence $\{x_t\}_{t\geq 1}$ to a fixed compact set.

2.3 Constraint Violation Velocity Projection (CVV-Pro)

We present below the pseudocode of Algorithm 1 for a fixed horizon length T, as it is standard in the literature [Hazan, 2016]. However, we note that our algorithm is oblivious to the horizon length T, i.e., it can run for any number of iterations without knowing T a priori.

Algorithm 1 Constraint Violation Velocity Projection (CVV-Pro)

Requirements: See Assumption 1.1
Input: $\alpha > 0$
Initialization: Step sizes $\left\{\eta_t = \frac{1}{\alpha\sqrt{t}}\right\}_{t\geq 1}$
for $t = 1$ **to** T **do**
  Play x_t **and observe**:
  cost information $f_t(x_t), \nabla f_t(x_t)$ and constraint information $\left\{\left(g_i(x_t), \nabla g_i(x_t)\right)\right\}_{i\in I(x_t)}$
  Construct the velocity polytope as follows

$$V_\alpha(x_t) := \left\{v \in \mathbb{R}^n \mid [\nabla g_i(x_t)]^\top v \geq -\alpha g_i(x_t),\ \forall i \in I(x_t)\right\},$$

  Solve the velocity projection problem $v_t = \arg\min_{v\in V_\alpha(x_t)} \frac{1}{2}\|v + \nabla f_t(x_t)\|^2$
  Update $x_{t+1} = x_t + \eta_t v_t$
end for

Let $x \in \mathcal{C}$ be an arbitrary decision. We show in Claim 2.2 that $\alpha(x - x_t) \in V_\alpha(x_t)$. Hence, the velocity set $V_\alpha(x_t)$ is always non-empty and well defined.

2.4 Structural Result

Here, we show that Algorithm 1 under Assumption 1.1 and a bounded iterate assumption, guarantees an optimal $\mathcal{O}(\sqrt{T})$ regret and converges to the feasible set at a rate $\mathcal{O}(1/\sqrt{T})$. The bounded iterate assumption will be removed subsequently, which however, will require a more complex analysis.

Theorem 2.1 (Structural)**.** *Suppose Assumption 1.1 holds and in addition $x_t \in \mathcal{B}_R$ for all $t \in \{1, \dots, T\}$. Then, on input $\alpha = L_{\mathcal{F}}/R$, Algorithm 1 with step sizes $\eta_t = \frac{1}{\alpha\sqrt{t}}$ guarantees the following for all $T \geq 1$:*

(regret) $\sum_{t=1}^{T} f_t(x_t) - \min_{x^\star \in \mathcal{C}} \sum_{t=1}^{T} f_t(x^\star) \leq 18 L_{\mathcal{F}} R \sqrt{T}$;

(feasibility) $g_i(x_t) \geq -8\left[\frac{L_{\mathcal{G}}}{R} + 2\beta_{\mathcal{G}}\right]\frac{R^2}{\sqrt{t}}$, *for all $t \in \{1, \dots, T\}$ and $i \in \{1, \dots, m\}$.*

2.5 Proof Sketch of Theorem 2.1

Our analysis establishes, in two steps, an important geometric property that connects the convex costs and the concave constraints via the velocity polytope $V_\alpha(x_t)$. This property will be crucial for deriving the regret and feasibility bounds.

In the first step, we leverage the constraints' concavity and show that the vector $\alpha(x^\star - x_t)$ belongs to the velocity polytope $V_\alpha(x_t)$.

Claim 2.2. *Suppose g_i is concave for every $i \in \{1, \dots, m\}$. Then $\alpha(x - x_t) \in V_\alpha(x_t)$ for all $x \in \mathcal{C}$. In addition, $x_t \notin \operatorname{int}(\mathcal{C})$ implies $[\nabla g_i(x_t)]^\top [x - x_t] \geq 0$ for all $x \in \mathcal{C}$.*

Proof. Let $x \in \mathcal{C}$ be an arbitrary feasible decision, satisfying $g_i(x) \geq 0$ for all $i \in \{1, \dots, m\}$. Since g_i is concave, we have $g_i(x_t) + [\nabla g_i(x_t)]^\top [x - x_t] \geq g_i(x) \geq 0$ and thus $[\nabla g_i(x_t)]^\top [x - x_t] \geq -g_i(x_t)$. The second conclusion follows by $x_t \notin \operatorname{int}(\mathcal{C})$, which implies $g_i(x_t) \leq 0$. □

In the second step, we show that $r_t^\top (x_t - x^\star) \leq 0$, where $r_t = v_t + \nabla f_t(x_t)$ is such that $-r_t$ belongs to the normal cone $N_{V_\alpha(x_t)}(v_t)$ of the velocity polytope $V_\alpha(x_t)$ evaluated at the projection v_t.

Lemma 2.3 (Main)**.** *Let v_t be the projection of $-\nabla f_t(x_t)$ onto the polytope $V_\alpha(x_t)$ such that $v_t = r_t - \nabla f_t(x_t) \in V_\alpha(x_t)$, where $-r_t \in N_{V_\alpha(x_t)}(v_t)$. Then, $-r_t^\top (x - x_t) \leq 0$ for all $x \in \mathcal{C}$.*

Proof. By definition, the normal cone $N_{V_\alpha(x_t)}(v_t)$ is given by $\{u \in \mathbb{R}^n \mid u^\top (v - v_t) \leq 0,\ \forall v \in V_\alpha(x_t)\}$. Then, by construction $-r_t \in N_{V_\alpha(x_t)}(v_t)$ and thus it holds for every $v \in V_\alpha(x_t)$ that $-r_t^\top [v - v_t] \leq 0$. The proof proceeds by case distinction:

Case 1. Suppose x_t is in the interior of $\mathcal{C}$. Then, $I(x_t) = \emptyset$, which implies $-\nabla f_t(x_t) \in V_\alpha(x_t) = \mathbb{R}^n$ and thus $r_t = 0$.

Case 2. Suppose x_t is on the boundary or outside of $\mathcal{C}$, i.e., $I(x_t) \neq \emptyset$. By Claim 2.2, we have $[\nabla g_i(x_t)]^\top [x - x_t] \geq 0$ for all $x \in \mathcal{C}$. By construction, $v_t \in V_\alpha(x_t)$ and thus $v(x) = v_t + x - x_t \in V_\alpha(x_t)$. The statement follows by applying $v = v(x)$ to $-r_t^\top [v - v_t] \leq 0$. □

Regret. To establish the first conclusion of Theorem 2.1 (regret), we combine the preceding geometric property with the analysis of online gradient descent. Since $f_t \in \mathcal{F}$ is convex, we upper bound the regret in terms of the gradient of f_t, namely $\sum_{t=1}^{T} f_t(x_t) - f_t(x^\star) \leq \sum_{t=1}^{T} [\nabla f_t(x_t)]^\top (x_t - x^\star)$ and then we show that the following inequality holds

$$
\begin{aligned}
[\nabla f_t(x_t)]^\top (x_t - x^\star) - \frac{\eta_t}{2}\|v_t\|^2 &= r_t^\top (x_t - x^\star) + \frac{\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2}{2\eta_t} \\
&\leq \frac{\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2}{2\eta_t}. \qquad (2)
\end{aligned}
$$

Moreover, in Appendix B (see Lemma B.2), we upper bound the velocity $\|v_t\| \leq \alpha\|x^\star - x_t\| + 2\|\nabla f_t(x_t)\|$. Combining Assumption 1.1 and $x_t \in \mathcal{B}_R$ yields a uniform bound $\|v_t\| \leq \mathcal{V}_\alpha$, where for $\alpha = L_{\mathcal{F}}/R$ we set $\mathcal{V}_\alpha := 4L_{\mathcal{F}}$. The desired regret follows by a telescoping argument and by convexity of the cost functions $f_t \in \mathcal{F}$.

Feasibility. For the second conclusion of Theorem 2.1 (convergence to the feasible set), we develop an inductive argument that proceeds in two steps. In Appendix B (see Claim B.6), we give a structural result that bounds the constraint functions from below. In particular, for every $i \in I(x_t)$ we have $g_i(x_{t+1}) \geq (1-\alpha\eta_t)g_i(x_t) - \eta_t^2\mathcal{V}_\alpha^2\beta_\mathcal{G}$ and for every $i \notin I(x_t)$ it holds that $g_i(x_{t+1}) \geq -\eta_{t+1}\mathcal{V}_\alpha[2L_\mathcal{G} + \mathcal{V}_\alpha\beta_\mathcal{G}/\alpha]$.

Using an inductive argument, we establish in Appendix B (see Lemma B.5) the following lower bound: $g_i(x_t) \geq -c\eta_t$ where $c = 2\mathcal{V}_\alpha(L_\mathcal{G} + \mathcal{V}_\alpha\beta_\mathcal{G}/\alpha)$. Choosing $\alpha = L_\mathcal{F}/R$ implies that $\mathcal{V}_\alpha = 4L_\mathcal{F}$. Then, the desired convergence rate to the feasible set follows for the step size $\eta_t = \frac{1}{\alpha\sqrt{t}}$, since

$$-c\eta_t = -\frac{2\mathcal{V}_\alpha}{\alpha\sqrt{t}}\Big[L_\mathcal{G} + \frac{\beta_\mathcal{G}\mathcal{V}_\alpha}{\alpha}\Big] = -8\Big[\frac{L_\mathcal{G}}{R} + 4\beta_\mathcal{G}\Big]\frac{R^2}{\sqrt{t}}.$$

3 Online Learning under Adversarial Nonlinear Constraints

3.1 Problem Formulation

In this section, we consider an online optimization problem with adversarially generated time-varying constraints. More precisely, at each time step t, the learner receives partial information on the current cost f_t and feasible set $\mathcal{C}_t$, and seeks to minimize (1). To make this problem well posed, we restrict the environment such that each feasible set $\mathcal{C}_t$ is contained in $\mathcal{Q}_t$ (see Section 1) and the rate of change between consecutive time-varying constraints *decreases* over time. We quantify a sufficient rate of decay with the following assumption.

Assumption 3.1 (TVC Decay Rate)**.** We assume that the adversarially generated sequence $\{g_t\}_{t\geq 1}$ of time-varying constraints is such that for every $x \in \mathcal{B}_{4R}$ and all $t \geq 1$, the following holds $\|g_{t+1}(x) - g_t(x)\|_\infty \leq \frac{98}{t+16}\big[\frac{L_\mathcal{G}}{R} + 3\beta_\mathcal{G}\big]R^2$.

We note that Assumption 3.1 essentially only requires $\|g_{t+1}(x) - g_t(x)\|_\infty \leq \mathcal{O}(1/t)$, as R can be chosen large enough such that the bound is satisfied. Of course, R will appear in our regret and feasibility bounds, but it will not affect the dependence on t or T (up to constant factors).

An important special case where Assumption 3.1 is satisfied, is summarized in the following Lemma. The proof is included in Appendix D (see Lemma D.7 and Lemma D.8).

Lemma 3.2. *Suppose the functions $\tilde{g}_{t,i}$ satisfy Assumption 1.1 and in addition there is a decision $x_{t,i} \in \mathcal{B}_R$ such that $\tilde{g}_{t,i}(x_{t,i}) = 0$ for every $t \geq 1$ and $i \in \{1,\dots,m\}$. Then the time-averaged constraints $g_{t,i}(x) := \frac{1}{t}\sum_{\ell=1}^{t}\tilde{g}_{\ell,i}(x)$ satisfy Assumption 1.1 and Assumption 3.1.*

3.2 Velocity Projection with Attractive Hypersphere Constraint

We show in Appendix C that the second assumption in Theorem 2.1, namely, "$x_t \in \mathcal{B}_R$ for all $t \geq 1$" can be enforced algorithmically. We achieve this in two steps.

1) Algorithmically, we introduce an additional hypersphere constraint $g_{m+1}(x_t) = \frac{1}{2}[R^2 - \|x_t\|^2]$ that attracts the decision sequence $\{x_t\}_{t\geq 1}$ to a hypersphere $\mathcal{B}_R$ and guarantees that it always stays inside a hypersphere $\mathcal{B}_{4R}$ with a slightly larger radius.

More precisely, we augment the velocity polytope in Step 3 of Algorithm 1 as follows: $V'_\alpha(x_t) = V_\alpha(x_t)$ if $\|x\| \leq R$, otherwise

$$V'_\alpha(x_t) = \{v \in V_\alpha(x_t) \mid [\nabla g_{m+1}(x_t)]^\top v \geq -\alpha g_{m+1}(x_t)\}.$$

2) Analytically, we give a refined inductive argument in Appendix D (see Lemma C.5), showing that $g_{m+1}(x_t) \geq -27R^2/\sqrt{t+15}$, $\|x_t\| \leq 4R$ and $\|v_t\| \leq 7L_\mathcal{F}$, for all $t \geq 1$.

3.3 Main Contribution

Our main contribution is to show that Algorithm 1 with the augmented velocity polytope $V'_\alpha(x_t)$, achieves optimal $\mathcal{O}(\sqrt{T})$ regret and satisfies $g_T(x_T) \geq -\Omega(1/\sqrt{T})$ convergence feasibility rate. Due to space limitations, we defer the proof to Appendix D.

Theorem 3.3 (Time-Varying Constraints)**.** *Suppose the functions $\{f_t, g_t\}_{t\geq 1}$ satisfy Assumptions 1.1, 1.2 and 3.1. Then, on input $R, L_{\mathcal{F}} > 0$ and $x_1 \in \mathcal{B}_R$, Algorithm 1 applied with $\alpha = L_{\mathcal{F}}/R$, augmented velocity polytope $V'_\alpha(\cdot)$ and step sizes $\eta_t = \frac{1}{\alpha\sqrt{t+15}}$ guarantees the following for all $T \geq 1$:*

(regret) $\sum_{t=1}^T f_t(x_t) - \min_{x\in\mathcal{C}_T} \sum_{t=1}^T f_t(x) \leq 246 L_{\mathcal{F}} R\sqrt{T}$;

(feasibility) $g_{t,i}(x_t) \geq -265\left[\frac{L_{\mathcal{G}}}{R} + 4\beta_{\mathcal{G}}\right]\frac{R^2}{\sqrt{t+15}}$, *for all $t \in \{1,\dots,T\}$ and $i \in \{1,\dots,m\}$;*

(attraction) $g_{m+1}(x_t) \geq -27\frac{R^2}{\sqrt{t+15}}$, *for all $t \in \{1,\dots,T\}$.*

Our regret analysis in Theorem 3.3 builds upon the following key structural result that generalizes Lemma 2.3 to time-varying constraints. In particular, in Appendix D (see Lemma D.3), we show that given the feasible set $\mathcal{C}_T \subset \mathcal{Q}_T$, it holds for every $x \in \mathcal{C}_T$ that $-r_t^\top(x - x_t) \leq 0$ for all $t \in \{1,\dots,T\}$. As a result, a similar argument as in (2) shows that the regret is bounded by $\mathcal{O}(\sqrt{T})$.

Moreover, we note that the linear and quadratic dependence on R in Theorem 3.3 is consistent in length units. Let the radius R be of length units ℓ, then the Lipschitz constant $L_{\mathcal{F}}$, which can be viewed as the supremum over the ℓ_2 norm of the gradient is of $1/\ell$ units, and the $\beta_{\mathcal{G}}$ smoothness constant (associated with Hessian) is of $1/\ell^2$ units. This means that the regret bound in Theorem 3.3 has the same units as f_t, while the feasibility bound has the same units as g_t.

4 Simulation examples

Two-player games with shared resources are an excellent example for demonstrating the effectiveness and importance of our online learning framework. We apply our algorithm and show numerical experiments that support our theoretical findings.

We choose random instances of a two player game with linear utility and constraints. In particular, we consider the following optimization problem

$$\min_{x\in\triangle_n} \max_{y\in\triangle_n} x^\top A y \quad \text{subject to} \quad C_x x + C_y y \leq 1, \tag{3}$$

where $\triangle_n = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1, x \geq 0\}$ is the probability simplex. Each component of the utility matrix $A \in \mathbb{R}^{n\times n}$ is sampled from the normal distribution and the constraint matrices $C_x, C_y \in [0,1]^{m\times n}$ have each of their components sampled uniformly at random from $[0,1]$.

4.1 Online Formulation

The problem in (3) can be modeled with our online learning framework (1) by choosing costs $f_t(x) := x^\top A y_t$ and time-averaged resource constraints $g_T(x) := \frac{1}{T}\sum_{t=1}^T \widetilde{g}_t(x)$, where the function $\widetilde{g}_t(x) := 1 - C_x x - C_y y_t$. Thus, the constraint in (3) is included as an average over the past iterations of y_t. The strategy for choosing y_t will be described below and, as we will see, the average of y_t over the past iterations converges. This ensures that the feasible set $\mathcal{C}_t$ (defined in (1)) is slowly time-varying, while the averages of x_t and y_t over past iterates converge to equilibria in (3). Further, by a refined version of Lemma 3.2 (see Lemma D.6 in Appendix D), the time-averaged constraints $g_T(x)$ satisfy Assumption 3.1.

In each iteration, Algorithm 1 seeks to minimize the online problem and commits to a decision x_t. The adversary computes the best response $\hat{y}_t$ with respect to the decision x_t by solving $\arg\max_{y\in\triangle_n} x_t^\top A y$. To make the dynamics more interesting, the adversary then commits with probability 0.8 to $\hat{y}_t$ and with probability 0.2 to a random decision r_t, i.e., $y_t = 0.8\hat{y}_t + 0.2 r_t$ where $r_t \overset{u.a.r.}{\sim} \triangle_n$.

As both players optimize over the probability simplex ($x, y \in \triangle_n$), the sequence of decisions $\{x_t\}_{t\geq 1}$ is automatically bounded. Thus, we can apply Theorem 3.3 with the original velocity polytope, as discussed in Appendix C. We implemented our algorithm with $\eta_t = 1/(\alpha\sqrt{t})$ and $\alpha = 100$.

4.2 Experimental Results

We report results from numerical simulations with decision dimension $n = 100$, $m = 10$ shared resource constraints, $T = 4000$ iterations, and five independently sampled instances of

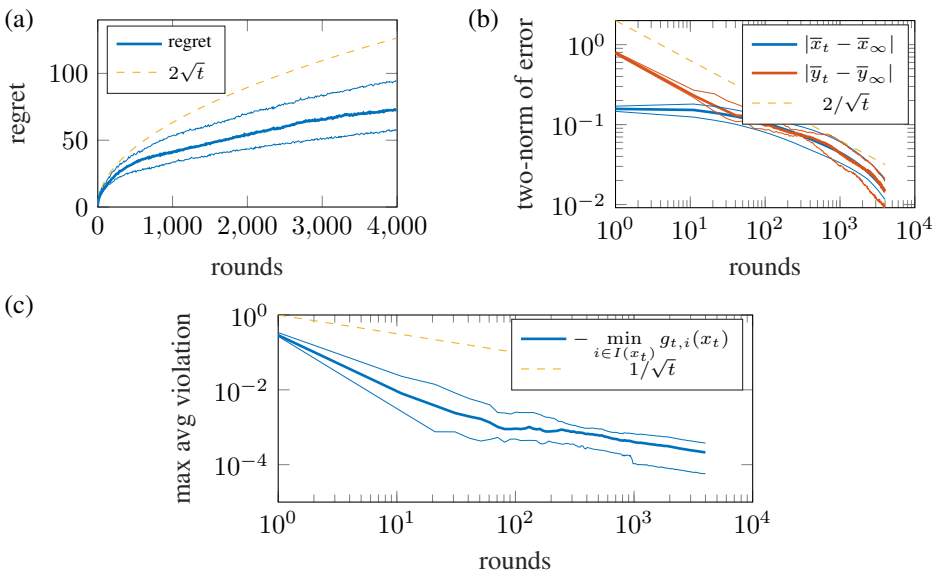

Figure 3: **(a)** The CVV-Pro algorithm is executed on five random instances of the two-player game with shared resources (Section 4.1). The regret follows the predicted $\mathcal{O}(\sqrt{T})$ slope. The thick line is the mean and the thin lines indicate the minimum and the maximum over the five runs. **(b)** The CVV-Pro algorithm achieves a convergence rate of $\mathcal{O}(1/\sqrt{t})$ for the averaged decisions $\overline{x}_t := \frac{1}{t}\sum_{\ell=1}^{t} x_\ell$ towards $\overline{x}_\infty$. In our experiment, we set $\overline{x}_\infty := \overline{x}_{10000}$. Similar behavior is reported for the averaged decisions $\overline{y}_t$ of the adversary. **(c)** The maximal constraint violation expressed by $-\min_{i \in I(x_t)} g_{t,i}(x_t)$ converges at a rate of $\mathcal{O}(1/\sqrt{t})$, as predicted by our theoretical results.

the two-player game. The learner's regret, depicted in Figure 3a, shows a clear correspondence with the theoretical prediction of $\mathcal{O}(\sqrt{T})$. Figure 3b presents the maximal constraint violation $-\min_{i \in I(x_T)} \frac{1}{T}\sum_{t=1}^{T} \widetilde{g}_{t,i}(x_T)$, which follows the predicted $\mathcal{O}(1/\sqrt{T})$ convergence rate. We also conclude from Figure 3c that the learner's averaged decisions $\overline{x}_T = \frac{1}{T}\sum_{t=1}^{T} x_t$ converge at a rate of $\mathcal{O}(1/\sqrt{T})$. Similarly, the averaged decisions $\overline{y}_T$ of the adversary also converge at a rate of $\mathcal{O}(1/\sqrt{T})$. We note that there is little variability in the results despite the different realizations of the matrices A, C_x, C_y.

Contrasting CVV-Pro and Online Gradient Descent In Appendix A, we show that our (CVV-Pro) algorithm outperforms the standard Online Gradient Descent algorithm in the two-player game from above. In particular, our algorithm achieves a lower regret and a runtime improvement of about 60%. Further, the percentage of violated constraints decreases rapidly and plateaus at 20%.

The amount of improvement in execution time is likely to be greater for higher-dimensional problems, where fewer constraints tend to be active at each iteration. Moreover, when the constraints are nonlinear, which includes ℓ_p norm or spectral constraints, optimizing over the full feasible set can be computationally challenging. In contrast, the velocity projection step in CCV-Pro is always a convex quadratic program with linear constraints, regardless of the underlying feasible set.

5 Conclusion

We propose an online algorithm that, despite the lack of information about the feasible set, achieves $\mathcal{O}(\sqrt{T})$ regret. We further ensure convergence of violated constraint $-\min\{g_T(x_T), 0\}$ at a rate of $\mathcal{O}(1/\sqrt{T})$ and derive explicit constants for all our bounds that hold for all $T \geq 1$. We thus conclude that our algorithm matches the performance of traditional online projected gradients or Frank-Wolfe schemes, while requiring substantially less information about the feasible set and allowing the feasible set to be time-varying. Perhaps equally important is the fact that instead of performing projections onto the full feasible set at each iteration, our algorithm only optimizes over a local sparse linear approximation. We show the applicability of our algorithm in numeric simulations of random two-player games with shared resources.

References

Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1):48–77, 2002. doi: 10.1137/S0097539701398375. URL https://doi.org/10.1137/S0097539701398375.

Avrim L. Blum. On-line algorithms in machine learning. *Online Algorithms: the state of the art*, pages 306–325, 1998.

Tianyi Chen, Qing Ling, and Georgios B. Giannakis. An online convex optimization approach to proactive network resource allocation. *IEEE Transactions on Signal Processing*, 65(24):6350–6364, 2017.

Francisco Facchinei and Christian Kanzow. Generalized Nash equilibrium problems. *4OR*, 5:173–210, 2007.

Francisco Facchinei and Christian Kanzow. Penalty methods for the solution of generalized Nash equilibrium problems. *SIAM Journal on Optimization*, 20(5):2228–2253, 2010.

Yoav Freud and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

Dan Garber and Ben Kretzu. New projection-free algorithms for online convex optimization with adaptive regret guarantees. *Proceedings of Machine Learning Research*, 178:2326–2359, 2022.

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

Rodolphe Jenatton, Jim Huang, and Cedric Archambeau. Adaptive algorithms for online convex optimization with long-term constraints. *Proceedings of Machine Learning Research*, 48:402–411, 2016.

M. I. Jordan. Artificial intelligence – the revolution hasn't happened yet. *Harvard Data Science Review*, 1(1):1–8, 2019.

M. I. Jordan and T. M. Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.

Michael I. Jordan, Tianyi Lin, and Manolis Zampetakis. First-order algorithms for nonlinear generalized Nash equilibrium problems. *arXiv:2204.03132v2*, pages 1–44, 2023.

Jong Gwang Kim. Equilibrium computation of generalized Nash games: A new Lagrangian-based approach. *arXiv:2106.00109v2*, pages 1–36, 2022.

Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

Kfir Y. Levy and Andreas Krause. Projection free online learning over smooth sets. *Proceedings of Machine Learning Research*, 89:1458–1466, 2019.

Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Inf. Comput.*, 108(2): 212–261, 1994. doi: 10.1006/inco.1994.1009. URL https://doi.org/10.1006/inco.1994.1009.

Qingsong Liu, Wenfei Wu, Longbo Huang, and Zhixuan Fang. Simultaneously achieving sublinear regret and constraint violations for online convex optimization with time-varying constraints. *Performance Evaluation*, 49(3):4–5, 2022.

Zhou Lu, Nataly Brukhim, Paula Gradu, and Elad Hazan. Projection-free adaptive regret with membership oracles. *arXiv.2211.12638*, pages 1–18, 2022.

Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. Trading regret for efficiency: Online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(81):2503–2528, 2012.

Zakaria Mhammedi. Efficient projection-free online convex optimization with membership oracle. *Proceedings of Machine Learning Research*, 178:5314–5390, 2022.

Michael Muehlebach and Michael I. Jordan. On constraints in first-order optimization: A view from non-smooth dynamical systems. *Journal of Machine Learning Research*, 23(256):1–47, 2022.

Michael J. Neely and Hao Yu. Online convex optimization with time-varying constraints. *arXiv:1702.04783*, pages 1–18, 2017.

Wen Sun, Debadeepta Dey, and Ashish Kapoor. Safety-aware algorithms for adversarial contextual bandit. *Proceedings of Machine Learning Research*, 70:3280–3288, 2017.

Xinlei Yi, Xiuxian Li, Tao Yang, Lihua Xie, Tianyou Chai, and Karl H. Johansson. Regret and cumulative constraint violation analysis for online convex optimization with long term constraints. *Proceedings of Machine Learning Research*, 139:1–11, 2021.

Hao Yu and Michael J. Neely. A low complexity algorithm with $o(\sqrt{T})$ regret and $o(1)$ constraint violations for online convex optimization with long term constraints. *Journal of Machine Learning Research*, 21(1):1–24, 2020.

Hao Yu, Michael Neely, and Xiaohan Wei. Online convex optimization with stochastic constraints. *Advances in Neural Information Processing Systems*, 30:1–11, 2017.

Jianjun Yuan and Andrew Lamperski. Online convex optimization for cumulative constraints. *Advances in Neural Information Processing Systems*, 31:1–10, 2018.

Supplementary Material for Online Learning under Adversarial Nonlinear Constraints

A Contrasting CVV-Pro and OGD: A Comparative Study

In this section, we compare the runtime performance and regret guarantee of the standard Online Gradient Descent (OGD) algorithm and our (CVV-Pro) algorithm in the two-player game setting (defined in Section 4). More concretely, we consider shared constraints of the form $C_x x + C_y y \leq b$. We report results from numerical simulations with decision dimension $n = 1000$, $m = 100$ shared resource constraints, capacity $b = 1.3$, $T = 2000$ iterations, and 5 independently sampled instances of the two-player game. We report below the results:

Regret: The 25th percentile of OGD has a higher regret around iteration 1400 than the function $5\sqrt{t}$ and stays above it. In contrast, CVV-Pro achieves better regret, with the 75th percentile being strictly bounded by the function $5\sqrt{t}$, see Figure S1a.

$\%$ **Constraints Violation** In each iteration, CVV-Pro requires an oracle access only to the currently violated constraints. The percentage of violated constraints first increases from 0.01% to 57% in the first four iterations, and then decreases rapidly to plateau at 20%, see Figure S1b.

Runtime: In Figure S1c, we report the average runtime per iteration for computing a projection. Since CVV-Pro solves the velocity projection problem with a decreasing number of constraints, it achieves a faster average runtime of 0.11 ± 0.01s compared to OGD, which requires solving the full projection problem each time and runs in 0.18 ± 0.01s. Thus, for the two-player game with shared constraints, our algorithm CVV-Pro achieves a runtime improvement of around 60% over OGD. Further, we report in Figure S1d the total cumulative runtime of CVV-Pro and OGD for computing the projection.

The amount of improvement in execution time is likely to be greater for higher-dimensional problems, where fewer constraints tend to be active at each iteration. Moreover, there are important situations, for example if constraints are non-convex, where projections are very difficult to compute (and/or might not even be well defined). In contrast, the velocity projection step in CCV-Pro is always a convex problem, regardless of whether the underlying feasible set is convex or not.

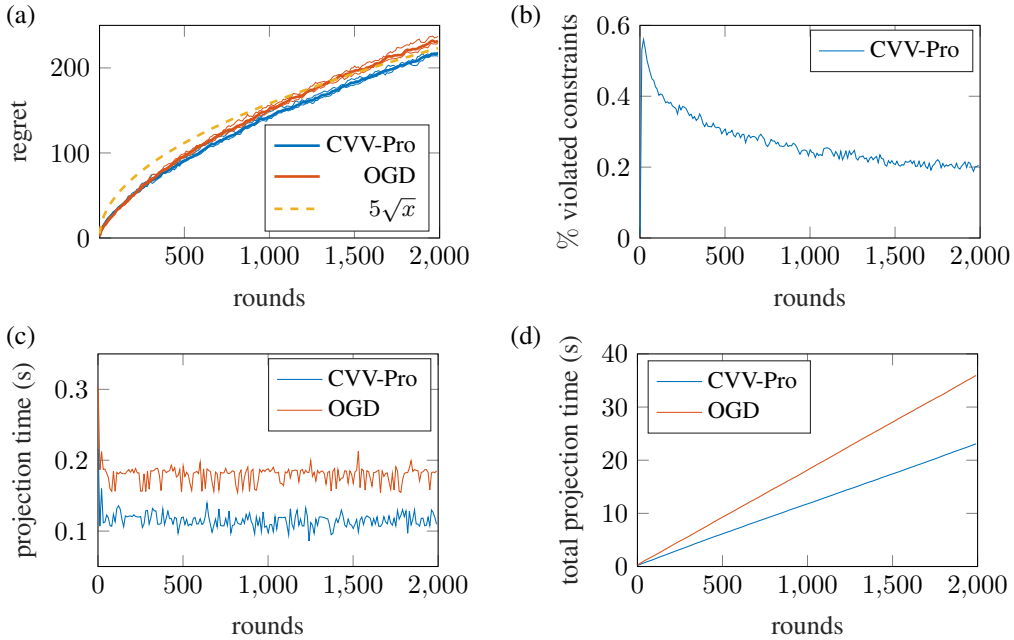

Figure S1: The figure contrasts CVV-Pro and OGD by comparing the resulting regret (a) and execution time (c,d). Panel (b) shows how the number of violated constraints evolves over time.

B Proof of Theorem 2.1

In this section, we consider an online optimization problem with time-invariant constraints and a bounded iterate assumption. The bounded iterate assumption will be removed subsequently in Section C, which however, will require a more complex analysis.

We restate Theorem 2.1 below for the convenience of the reader.

Theorem B.1 (Structural)**.** *Suppose Assumption 1.1 holds and in addition $x_t \in \mathcal{B}_R$ for all $t \in \{1, \dots, T\}$. Then, on input $\alpha = L_{\mathcal{F}}/R$, Algorithm 1 with step sizes $\eta_t = \frac{1}{\alpha\sqrt{t}}$ guarantees the following for all $T \geq 1$:*

(regret) $\sum_{t=1}^{T} f_t(x_t) - \min_{x^\star \in \mathcal{C}} \sum_{t=1}^{T} f_t(x^\star) \leq 18 L_{\mathcal{F}} R \sqrt{T}$;

(feasibility) $g_i(x_t) \geq -8 \left[\frac{L_{\mathcal{G}}}{R} + 2\beta_{\mathcal{G}}\right] \frac{R^2}{\sqrt{t}}$, *for all $t \in \{1, \dots, T\}$ and $i \in \{1, \dots, m\}$.*

The rest of this section is devoted to proving the preceding statement.

B.1 Structural Properties

Lemma B.2. *Suppose g_i is concave for every $i \in \{1, \dots, m\}$. Then, for any $\alpha > 0$ and all $x \in \mathcal{C}$ the following holds*

$$\max_{t \geq 0} \|v_t\| \leq \alpha \|x - x_t\| + 2\|\nabla f_t(x_t)\|.$$

In particular, when f_t satisfies $\|\nabla f_t(z)\| \leq L_{\mathcal{F}}$ for all $z \in \mathcal{B}_{cR}$, it follows that $\|v_t\| \leq (c+1)\alpha R + 2L_{\mathcal{F}}$ for any $x \in \mathcal{B}_R$, $x_t \in \mathcal{B}_{cR}$, and $c > 0$.

Proof. By Claim 2.2, we have $\alpha(x - x_t) \in V_\alpha(x_t)$ for every $x \in \mathcal{C}$. Combining the triangle inequality with the fact that v_t is an optimal solution of the velocity projection problem in Step 8, yields

$$\begin{aligned} \|v_t\| - \|\nabla f_t(x_t)\| &\leq \|v_t + \nabla f_t(x_t)\| \\ &\leq \|\alpha(x - x_t) + \nabla f_t(x_t)\| \\ &\leq \alpha\|x - x_t\| + \|\nabla f_t(x_t)\|. \end{aligned}$$

Using $x \in \mathcal{B}_R$, $x_t \in \mathcal{B}_{cR}$ and $\|\nabla f_t(x_t)\| \leq L_{\mathcal{F}}$, we conclude

$$\|v_t\| \leq \alpha\|x - x_t\| + 2\|\nabla f_t(x_t)\| \leq (c+1)\alpha R + 2L_{\mathcal{F}}.$$

□

B.2 Cost Regret

Lemma B.3. *Suppose Assumption 1.1 holds and $x_t \in \mathcal{B}_{cR}$ for all $t \in \{1, \dots, T\}$ with $c \in (0, 4]$. Let $d \geq 0$ be a constant. Then, Algorithm 1 applied with $\alpha = L_{\mathcal{F}}/R$ and step sizes $\eta_t = \frac{1}{\alpha\sqrt{t+d}}$, guarantees the following for all $T \geq 1$:*

$$R_T = \sum_{t=1}^{T} f_t(x_t) - \min_{x^\star \in \mathcal{C}} \sum_{t=1}^{T} f_t(x^\star) \leq \sqrt{d+1} \left[(c+3)^2 + \frac{1}{2}(c+1)^2\right] L_{\mathcal{F}} R \sqrt{T}.$$

In particular, for $c = 1$ and $d = 0$ we have $R_T \leq 18 L_{\mathcal{F}} R \sqrt{T}$.

Proof. We denote an optimal decision in hindsight by $x^\star \in \arg\min_{x \in \mathcal{C}} \sum_{t=1}^{T} f_t(x)$. For any points $x^\star$, x_t we have $f_t(x_t) - f_t(x^\star) \leq [\nabla f_t(x_t)]^\top (x_t - x^\star)$, since f_t is convex. Summing over the number of rounds t results in

$$\sum_{t=1}^{T} f_t(x_t) - f_t(x^\star) \leq \sum_{t=1}^{T} [\nabla f_t(x_t)]^\top (x_t - x^\star).$$

We proceed by upper bounding the expression $[\nabla f_t(x_t)]^\top (x_t - x^\star)$. Using $x_{t+1} = x_t + \eta_t v_t$ and $v_t = r_t - \nabla f_t(x_t)$, we have

$$\begin{aligned} \|x_{t+1} - x^\star\|^2 &= \|x_t + \eta_t (r_t - \nabla f_t(x_t)) - x^\star\|^2 \\ &= \|x_t - x^\star\|^2 + \eta_t^2 \|r_t - \nabla f_t(x_t)\|^2 + 2\eta_t \left[r_t - \nabla f_t(x_t)\right]^\top (x_t - x^\star). \end{aligned}$$

Then, Lemma 2.3 gives $r_t^\top(x_t - x^\star) \le 0$ and thus

$$\begin{aligned}
[\nabla f_t(x_t)]^\top(x_t - x^\star) &= r_t^\top(x_t - x^\star) + \frac{\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2}{2\eta_t} + \frac{\eta_t}{2}\|v_t\|^2 \\
&\le \frac{\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2}{2\eta_t} + \frac{\eta_t}{2}\|v_t\|^2.
\end{aligned}$$

Since $x^\star \in \mathcal{B}_R$ and $x_t \in \mathcal{B}_{cR}$ for all $t \in \{1, \dots, T\}$, by Lemma B.2 it follows for all $t \in \{1, \dots, T\}$ that

$$\|v_t\| \le (c+1)\alpha R + 2L_{\mathcal{F}} = (c+3)L_{\mathcal{F}} =: \mathcal{V}_\alpha. \tag{S1}$$

Summing over the whole sequence, using the fact that $\eta_t = \frac{1}{\alpha\sqrt{t+d}}$ is a decreasing positive sequence and applying Claim B.4, $x^\star \in \mathcal{B}_R$, $x_t \in \mathcal{B}_{cR}$, and (S1), yields

$$\begin{aligned}
2\sum_{t=1}^{T}[\nabla f_t(x_t)]^\top(x_t - x^\star) &\le \sum_{t=1}^{T} \frac{\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2}{\eta_t} + \eta_t\|v_t\|^2 \\
&\le \mathcal{V}_\alpha^2\left(\sum_{t=1}^{T}\eta_t\right) + \frac{(c+1)^2R^2}{\eta_T} \\
&\le (c+3)^2L_{\mathcal{F}}^2\frac{2}{\alpha}\sqrt{T+d} + (c+1)^2L_{\mathcal{F}}R\sqrt{T+d} \\
&= \left[2(c+3)^2 + (c+1)^2\right]L_{\mathcal{F}}R\sqrt{T+d},
\end{aligned}$$

where last inequality uses

$$\sum_{t=1}^{T}\eta_t = \frac{1}{\alpha}\sum_{t=1}^{T}\frac{1}{\sqrt{t+d}} < \frac{1}{\alpha}\sum_{t=1}^{T+d}\frac{1}{\sqrt{t}} \le \frac{2}{\alpha}\sqrt{T+d}.$$

The statement follows by combining the fact that $\sqrt{T+d} \le \sqrt{d+1}\sqrt{T}$ for any $d \ge 0$ and all $T \ge 1$, and

$$\sum_{t=1}^{T} f_t(x_t) - f_t(x^\star) \le \sum_{t=1}^{T}[\nabla f_t(x_t)]^\top(x_t - x^\star) \le \sqrt{d+1}\left[(c+3)^2 + \frac{1}{2}(c+1)^2\right]L_{\mathcal{F}}R\sqrt{T}.$$

□

Claim B.4 (Series)**.** *For any positive sequence $\{a_t\}_{t=1}^{T+1}$ and any decreasing positive sequence $\{\eta_t\}_{t=1}^{T}$, it holds that*

$$\sum_{t=1}^{T}\frac{a_t - a_{t+1}}{\eta_t} \le \frac{A}{\eta_T}, \quad \textit{where} \quad A := \max_{t=\{1,\dots,T\}} a_t.$$

Proof. Observe that

$$\begin{aligned}
\sum_{t=1}^{T}\frac{a_t - a_{t+1}}{\eta_t} &= \frac{a_1 - a_2}{\eta_1} + \frac{a_2 - a_3}{\eta_2} + \frac{a_3 - a_4}{\eta_3} + \dots + \frac{a_T - a_{T+1}}{\eta_T} \\
&= \frac{a_1}{\eta_1} - \frac{a_{T+1}}{\eta_T} + \sum_{i=2}^{\top} a_i\left(\frac{1}{\eta_i} - \frac{1}{\eta_{i-1}}\right) \\
&\le \frac{A}{\eta_T},
\end{aligned}$$

where the last inequality follows by

$$\sum_{i=2}^{\top} a_i\left(\frac{1}{\eta_i} - \frac{1}{\eta_{i-1}}\right) \le A\sum_{i=2}^{\top}\left(\frac{1}{\eta_i} - \frac{1}{\eta_{i-1}}\right) = A\left(\frac{1}{\eta_T} - \frac{1}{\eta_1}\right) \le \frac{A}{\eta_T} - \frac{a_1}{\eta_1}.$$

□

B.3 Convergence Rate of Constraint Violations

Lemma B.5 (Convergence Rate of Constraint Violations)**.** *Suppose Assumption 1.1 holds and* $\{x_t\}_{t\geq 1} \in \mathcal{B}_{cR}$ *with* $x_1 \in \mathcal{B}_R$ *and* $c \in (0, 4]$*. Then, for any* $\alpha > 0$ *and* $d \geq 0$*, step sizes* $\eta_t = 1/(\alpha\sqrt{t+d})$ *and* $\mathcal{V}_\alpha > 0$ *such that* $\|v_t\| \leq \mathcal{V}_\alpha$ *for all* $t \geq 1$*, it follows for every* $i \in \{1, \ldots, m\}$ *and* $t \geq 1$ *that*

$$g_i(x_t) \geq -c_1\eta_t,$$

where

$$c_1 = \mathcal{V}_\alpha\left[2L_\mathcal{G} + \frac{\beta_G\mathcal{V}_\alpha}{\alpha}\right] + \mathcal{Z}_d \quad \textit{and} \quad \mathcal{Z}_d = \left(1 - \frac{1}{\sqrt{d+1}}\right)\sqrt{d+2}\left[\frac{L_\mathcal{G}}{R} + \beta_\mathcal{G}\right]2\alpha R^2.$$

In particular, when Assumption 1.1 holds, $\{x_t\}_{t\geq 1} \in \mathcal{B}_R$*,* $\alpha = L_\mathcal{F}/R$ *and* $d = 0$*, it follows that*

$$g_i(x_t) \geq -8\left[\frac{L_\mathcal{G}}{R} + 2\beta_G\right]\frac{R^2}{\sqrt{t}} \quad \textit{for all } t \geq 1.$$

Proof. The proof is by induction on t. We start with the base case $t = 1$. The proof proceeds by case distinction.

Case 1. Suppose $i \in \{1, \ldots, m\}\backslash I(x_1)$, i.e., $g_i(x_1) > 0$. Then, by Claim B.6 Part ii) we have

$$g_i(x_2) \geq -\eta_2\mathcal{V}_\alpha\left[2L_\mathcal{G} + \frac{\mathcal{V}_\alpha\beta_\mathcal{G}}{\alpha\sqrt{1+d}}\right] \geq -c_1\eta_2.$$

Case 2. Suppose $i \in I(x_1)$, i.e., $g_i(x_1) \leq 0$. By combining $x_1 \in \mathcal{B}_R$ and g_i is concave $\beta_\mathcal{G}$-smooth, it follows for every $x \in \mathcal{C} \subseteq \mathcal{B}_R$ that

$$\begin{aligned} g_i(x_1) &\geq g_i(x) + \nabla g_i(x)^T(x_1 - x) - \frac{\beta_\mathcal{G}}{2}\|x_1 - x\|^2 \\ &\geq -2L_\mathcal{G}R - 2\beta_\mathcal{G}R^2 \\ &= -\eta_1\sqrt{d+1}\left[\frac{L_\mathcal{G}}{R} + \beta_\mathcal{G}\right]2\alpha R^2 \geq -c_1\eta_2. \end{aligned}$$

Using $\eta_t = 1/(\alpha\sqrt{t+d})$, $\eta_1/\eta_2 \leq \sqrt{2}$ and $\eta_1^2\frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{2} \leq \eta_2^2\mathcal{V}_\alpha^2\beta_\mathcal{G} = \eta_2\frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{\alpha\sqrt{d+2}}$, it follows by Claim B.6 Part i) that

$$\begin{aligned} g_i(x_2) &\geq (1 - \alpha\eta_1)g_i(x_1) - \eta_1^2\frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{2} \\ &\geq -\eta_2\left[\left(1 - \frac{1}{\sqrt{d+1}}\right)\sqrt{d+2}\left[\frac{L_\mathcal{G}}{R} + \beta_\mathcal{G}\right]2\alpha R^2 + \frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{\alpha\sqrt{d+2}}\right] \geq -c_1\eta_2. \end{aligned}$$

Our inductive hypothesis is $g_i(x_t) \geq -c_1\eta_t$ for all i. We now show that it holds for $t+1$.

Case 1. Suppose $i \in \{1, \ldots, m\}\backslash I(x_1)$, i.e., $g_i(x_t) > 0$. Then by Claim B.6 Part ii)

$$g_i(x_{t+1}) \geq -\eta_{t+1}\mathcal{V}_\alpha\left[2L_\mathcal{G} + \frac{\beta_\mathcal{G}\mathcal{V}_\alpha}{\alpha\sqrt{d+1}}\right] \geq -c_1\eta_{t+1}.$$

Case 2. Suppose $i \in I(x_t)$, i.e., $g_i(x_t) \leq 0$. Combining Claim B.6 Part ii) and the inductive hypothesis we have

$$\begin{aligned} g_i(x_{t+1}) &\geq (1 - \alpha\eta_t)g_i(x_t) - \eta_t^2\frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{2} \\ &\geq -c_1\eta_t + c_1\alpha\eta_t^2 - \eta_t^2\frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{2} \\ &= -c_1\eta_{t+1} + c_1\eta_{t+1} - c_1\eta_t + c_1\alpha\eta_t^2 - \eta_t^2\frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{2} \\ &= -c_1\eta_{t+1} + c_1\eta_t\left[\frac{\eta_{t+1}}{\eta_t} - 1 + \alpha\eta_t - \eta_t\frac{\mathcal{V}_\alpha^2\beta_\mathcal{G}}{2c_1}\right]. \end{aligned}$$

Since $c_1\eta_t > 0$, it suffices to show that

$$\alpha - \frac{\eta_t - \eta_{t+1}}{\eta_t^2} \geq \frac{\mathcal{V}_\alpha^2 \beta_{\mathcal{G}}}{c_1} \tag{S2}$$

or equivalently (using $\eta_t = \frac{1}{\alpha\sqrt{t+d}}$ for $t \geq 1$)

$$\alpha - \alpha\sqrt{\frac{t+d}{t+d+1}}\left(\sqrt{t+d+1} - \sqrt{t+d}\right) \geq \frac{\mathcal{V}_\alpha^2 \beta_{\mathcal{G}}}{2c_1}.$$

Straightforward checking shows that $\max_{t\geq 1}\sqrt{\frac{t}{t+1}}\left(\sqrt{t+1} - \sqrt{t}\right) < \frac{1}{3}$. Hence, inequality (S2) is implied for $c_1 \geq \beta_{\mathcal{G}}\mathcal{V}_\alpha^2/\alpha$ and thus $g_i(x_{t+1}) \geq -c_1\eta_{t+1}$.

Furthermore, for $c = 1$ and $\alpha = L_{\mathcal{F}}/R$, by Lemma B.2, we can set $\mathcal{V}_\alpha = 4L_{\mathcal{F}}$. Then, for $d = 0$ we have $g_i(x_t) \geq -8\left[\frac{L_{\mathcal{G}}}{R} + 2\beta_{\mathcal{G}}\right]\frac{R^2}{\sqrt{t}}$ for all $t \geq 1$. □

Claim B.6 (Constraint Violation)**.** *Suppose g_i is concave, $\beta_{\mathcal{G}}$-smooth and satisfies $\|\nabla g_i(x)\| \leq L_{\mathcal{G}}$ for all $x \in \mathcal{B}_{cR}$ and $i \in \{1, \ldots, m\}$, where $c > 0$ is a constant. Suppose further that there exists a constant $\mathcal{V}_\alpha > 0$ such that $x_t \in \mathcal{B}_{cR}$ and $\|v_t\| \leq \mathcal{V}_\alpha$, for all $t \geq 1$. Then, for all $t \geq 1$ we have*

i) $g_i(x_{t+1}) \geq (1 - \alpha\eta_t)g_i(x_t) - \eta_t^2\mathcal{V}_\alpha^2\beta_{\mathcal{G}}/2$ for every $i \in I(x_t)$;

ii) $g_i(x_{t+1}) \geq -\eta_{t+1}\mathcal{V}_\alpha\left[2L_{\mathcal{G}} + \mathcal{V}_\alpha\beta_{\mathcal{G}}/(\alpha\sqrt{1+d})\right]$ for every $i \in \{1, \ldots, m\}\backslash I(x_t)$.

Proof. The proof proceeds by case distinction.

Case 1. Suppose $i \in I(x_t)$, i.e., $g_i(x_t) \leq 0$. By combining the facts that g_i is concave and $\beta_{\mathcal{G}}$-smooth, $x_{t+1} = x_t + \eta_t v_t$ and $[\nabla g_i(x_t)]^\top v_t \geq -\alpha g_i(x_t)$, it follows that

$$\begin{aligned} g_i(x_{t+1}) &\geq g_i(x_t) + [\nabla g_i(x_t)]^\top[x_{t+1} - x_t] - \frac{\beta_{\mathcal{G}}}{2}\|x_{t+1} - x_t\|_2^2 \\ &\geq (1 - \alpha\eta_t)g_i(x_t) - \eta_t^2\mathcal{V}_\alpha^2\frac{\beta_{\mathcal{G}}}{2}. \end{aligned} \tag{S3}$$

Case 2. Suppose $i \notin I(x_t)$, i.e., $g_i(x_t) > 0$. Using $\|\nabla g_i(x)\| \leq L_{\mathcal{G}}$ for $x_t \in \mathcal{B}_{cR}$, we have

$$[\nabla g_i(x_t)]^\top[x_{t+1} - x_t] \leq \|\nabla g_i(x_t)\|\|x_{t+1} - x_t\| \leq \eta_t L_{\mathcal{G}}\mathcal{V}_\alpha.$$

Hence,

$$\begin{aligned} g_i(x_{t+1}) &\geq g_i(x_t) + [\nabla g_i(x_t)]^\top[x_{t+1} - x_t] - \frac{\beta_{\mathcal{G}}}{2}\|x_{t+1} - x_t\|_2^2 \\ &\geq -\eta_t L_{\mathcal{G}}\mathcal{V}_\alpha - \eta_t^2\mathcal{V}_\alpha^2\frac{\beta_{\mathcal{G}}}{2} \\ &= -\eta_{t+1}\frac{\eta_t}{\eta_{t+1}}\mathcal{V}_\alpha\left[L_{\mathcal{G}} + \frac{\eta_t}{2}\mathcal{V}_\alpha\beta_{\mathcal{G}}\right] \\ &> -\eta_{t+1}\mathcal{V}_\alpha\left[2L_{\mathcal{G}} + \frac{\mathcal{V}_\alpha\beta_{\mathcal{G}}}{\alpha\sqrt{1+d}}\right], \end{aligned}$$

where the last inequality follows by $\eta_t \leq \eta_1 = 1/(\alpha\sqrt{1+d})$ and

$$\max_{\ell\geq 1}\frac{\eta_\ell}{\eta_{\ell+1}} \leq \max_{\ell\geq 1}\sqrt{\frac{\ell+1}{\ell}} = \sqrt{2}.$$

□

C Guaranteeing a Bounded Decision Sequence

We now show that the second assumption in Theorem 2.1, namely, "$x_t \in \mathcal{B}_R$ for all $t \in \{1, \ldots, T\}$" can be enforced algorithmically. We achieve this by introducing an additional hypersphere constraint $g_{m+1}(x_t) = \frac{1}{2}[R^2 - \|x_t\|^2]$ that attracts the decision sequence $\{x_t\}_{t\geq 1}$ to a hypersphere $\mathcal{B}_R$ and

guarantees that it always stays inside a hypersphere $\mathcal{B}_{4R}$ with a slightly larger radius. Technically, we modify the velocity polytope in Step 3 of Algorithm 1 as follows: $V'_\alpha(x_t) = V_\alpha(x_t)$ if $\|x\| \le R$, and otherwise

$$V'_\alpha(x_t) = \{v \in V_\alpha(x_t) \mid [\nabla g_{m+1}(x_t)]^\top v \ge -\alpha g_{m+1}(x_t)\}.$$

We are now ready to state our main algorithmic result for the setting of time-invariant constraints.

Theorem C.1 (Algorithm)**.** *Suppose Assumption 1.1 holds. Then, on input $R, L_{\mathcal{F}} > 0$, $\alpha = L_{\mathcal{F}}/R$ and $x_1 \in \mathcal{B}_R$, Algorithm 1 with augmented velocity polytope $V'_\alpha(\cdot)$ and step sizes $\eta_t = \frac{1}{\alpha\sqrt{t+15}}$ guarantees the following for all $T \ge 1$:*

- ***(regret)*** $\sum_{t=1}^T f_t(x_t) - \min_{x^\star \in \mathcal{C}} \sum_{t=1}^T f_t(x^\star) \le 246 L_{\mathcal{F}} R\sqrt{T}$;
- ***(feasibility)*** $g_i(x_t) \ge -21\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_G\right]\frac{R^2}{\sqrt{t+15}}$, *for all $t \in \{1,\dots,T\}$ and $i \in \{1,\dots,m\}$;*
- ***(attraction)*** $g_{m+1}(x_t) \ge -27\frac{R^2}{\sqrt{t+15}}$ *for all $t \in \{1,\dots,T\}$.*

In addition, $\|x_t\| \le 4R$ and $\|v_t\| \le 7L_{\mathcal{F}}$, for all $t \ge 1$.

To ensure convergence of the hypersphere constraint $-\min\{g_{m+1}(x_t), 0\}$ at a rate of $\mathcal{O}(1/\sqrt{t})$, we use an inductive argument similar to Lemma B.5. We note that compared to the simplified setting of Appendix B, our analysis requires an additional refined inductive argument, which is summarized in Lemma C.5.

C.1 Hypersphere constraint

Definition C.2. We consider the following hypersphere constraint, parameterized by $R > 0$,

$$g_{m+1}(x) = \frac{1}{2}[R^2 - \|x_t\|^2].$$

By construction, g_{m+1} is concave and 1-smooth.

Claim C.3. *Suppose g_i is concave for every $i \in \{1,\dots,m\}$ such that $\mathcal{C} \subseteq \mathcal{B}_R$ and f_t is convex such that $\|\nabla f_t(x)\| \le L_{\mathcal{F}}$ for all $x \in \mathcal{B}_{cR}$, where $c > 0$ is a constant. Then for any decision $x_t \in \mathcal{B}_{cR}$, it holds that*

$$\|v_t\| \le \alpha\|x_t\| + (\alpha R + 2L_{\mathcal{F}}) \qquad \textit{and} \qquad \frac{1}{2}\|v_t\|^2 < -2\alpha^2 g_{m+1}(x) + \left[\alpha^2 R^2 + (\alpha R + 2L_{\mathcal{F}})^2\right].$$

Proof. Due to the fact that g_{m+1} and g_i are concave for every $i \in \{1,\dots,m\}$, it follows by Lemma B.2 that

$$\begin{aligned} \|v_t\| &\le 2\|\nabla f_t(x_t)\| + \alpha\|x^\star - x_t\| \\ &\le \alpha\|x_t\| + \alpha R + 2L_{\mathcal{F}}. \end{aligned}$$

Further, by definition of $g_{m+1}(x)$ we have

$$\begin{aligned} \frac{1}{2}\|v_t\|^2 &\le \frac{1}{2}\left[\alpha\|x_t\| + (\alpha R + 2L_{\mathcal{F}})\right]^2 \\ &\le \alpha^2\|x_t\|^2 + (\alpha R + 2L_{\mathcal{F}})^2 \\ &= -2\alpha^2 g_{m+1}(x) + \left[\alpha^2 R^2 + (\alpha R + 2L_{\mathcal{F}})^2\right]. \end{aligned}$$

□

Claim C.4. *Suppose the assertions in Claim C.3 hold. Let the step sizes be $\{\eta_t = \frac{1}{\alpha\sqrt{t+15}}\}_{t\ge1}$ and $\alpha = L_{\mathcal{F}}/R$. Then, we have*

i) If $g_{m+1}(x_t) > 0$ then $g_{m+1}(x_{t+1}) \ge -\eta_t \cdot 6L_{\mathcal{F}}R$; and

ii) If $g_{m+1}(x_t) \le 0$ then $g_{m+1}(x_{t+1}) \ge \left(1 - \frac{\alpha}{2}\eta_t\right)g_{m+1}(x_t) - \eta_t^2 10 L_{\mathcal{F}}^2$.

Proof. The proof is by case distinction.

Case 1. Suppose $g_{m+1}(x_t) > 0$. Using $\|x_t\| < R$ it follows by Claim C.3 that

$$\|v_t\| \leq 2\left(\alpha R + L_{\mathcal{F}}\right) = 4L_{\mathcal{F}}.$$

Using g_{m+1} is concave and 1-smooth, $g_{m+1}(x_t) > 0$, $\nabla g_{m+1}(x_t) = -x_t$ and $\|x_t\| < R$, we have

$$\begin{aligned} g_{m+1}(x_{t+1}) &\geq g_{m+1}(x_t) + \nabla g_{m+1}(x_t)^\top (x_{t+1} - x_t) - \frac{1}{2}\|x_{t+1} - x_t\|^2 \\ &\geq -\eta_t R\|v_t\| - \frac{1}{2}\eta_t^2\|v_t\|^2 \\ &\geq -\eta_t \cdot 6L_{\mathcal{F}}R \\ &\geq -\eta_{t+1} \cdot 7L_{\mathcal{F}}R, \end{aligned}$$

where we used

$$\frac{1}{2}\eta_t 16L_{\mathcal{F}}^2 = \frac{8}{\sqrt{t+15}}L_{\mathcal{F}}R \leq 2L_{\mathcal{F}}R.$$

Case 2. Suppose $g_{m+1}(x_t) \leq 0$, i.e., $\|x_t\| \geq R$. Using $\alpha^2R^2 + (\alpha R + 2L_{\mathcal{F}})^2 = 10L_{\mathcal{F}}^2$, it follows by Claim C.3 that

$$\frac{1}{2}\|v_t\|^2 < -2\alpha^2 g_{m+1}(x) + 10L_{\mathcal{F}}^2.$$

Combining g_{m+1} is concave and 1-smooth, and $\nabla g_{m+1}(x_t)^\top v_t \geq -\alpha g_{m+1}(x_t)$ yields

$$\begin{aligned} g_{m+1}(x_{t+1}) &\geq g_{m+1}(x_t) + \nabla g_{m+1}(x_t)^\top (x_{t+1} - x_t) - \frac{1}{2}\|x_{t+1} - x_t\|^2 \\ &\geq (1 - \alpha\eta_t) g_{m+1}(x_t) - \frac{1}{2}\eta_t^2\|v_t\|^2 \\ &> \left(1 - \alpha\eta_t + 2\alpha^2\eta_t^2\right) g_{m+1}(x_t) - \eta_t^2 10L_{\mathcal{F}}^2 \\ &\geq \left(1 - \frac{\alpha}{2}\eta_t\right) g_{m+1}(x_t) - \eta_t^2 10L_{\mathcal{F}}^2, \end{aligned}$$

where the last inequality follows by: $-\eta_t\alpha + 2\eta_t^2\alpha^2 \leq -\eta_t\frac{\alpha}{2}$, which is implied by $\eta_t = \frac{1}{\alpha\sqrt{t+15}}$. □

Lemma C.5 (Main). *Suppose the assertions in Claim C.3 hold for $c = 4$. Given $\alpha = L_{\mathcal{F}}/R$, step sizes $\{\eta_t = \frac{1}{\alpha\sqrt{t+15}}\}_{t\geq 1}$ and an arbitrary initial decision x_1 with $\|x_1\| < R$, then it holds that*

$$g_{m+1}(x_t) \geq -27\frac{R^2}{\sqrt{t}}, \qquad \|x_t\| \leq 4R, \qquad \|v_t\| \leq 7L_{\mathcal{F}}, \qquad \textit{for all } t \geq 1. \tag{S4}$$

Proof. The proof is by induction on $t \geq 1$.

Part I) We show first that $g_{m+1}(x_{t+1}) \geq -c_0\eta_{t+1}$, for some $c_0 > 0$. The proof proceeds by case distinction.

Case 1. Suppose $g_{m+1}(x_t) > 0$, then by Claim C.4 we have

$$g_{m+1}(x_{t+1}) \geq -\eta_t \cdot 6L_{\mathcal{F}}R, \qquad (\text{implying } c_0 \geq 6L_{\mathcal{F}}R).$$

Case 2. Suppose $g_{m+1}(x_t) \leq 0$. Let $A := 10L_{\mathcal{F}}^2$, then by combining Claim C.4 and the inductive hypothesis, we have

$$\begin{aligned} g_{m+1}(x_{t+1}) &\geq \left(1 - \eta_t\frac{\alpha}{2}\right) g_{m+1}(x_t) - \eta_t^2 A \\ &\geq -\left(1 - \eta_t\frac{\alpha}{2}\right) c_0\eta_t - \eta_t^2 A \\ &= -c_0\eta_t - \left(A - \frac{\alpha}{2}c_0\right)\eta_t^2 \\ &= -c_0\eta_{t+1} - c_0\eta_t + c_0\eta_{t+1} - \left(A - \frac{\alpha}{2}c_0\right)\eta_t^2 \\ &= -c_0\eta_{t+1} + c_0\eta_t\left[-1 + \frac{\eta_{t+1}}{\eta_t} - \eta_t\left(\frac{A}{c_0} - \frac{\alpha}{2}\right)\right]. \end{aligned}$$

Since $c_0\eta_t > 0$, it suffices to show that

$$-1 + \frac{\eta_{t+1}}{\eta_t} - \eta_t\left(\frac{A}{c_0} - \frac{\alpha}{2}\right) \geq 0 \iff \frac{\alpha}{2} - \frac{\eta_t - \eta_{t+1}}{\eta_t^2} \geq \frac{A}{c_0}.$$

The previous condition is equivalent to (using $\eta_t = \frac{1}{\alpha\sqrt{t+15}}$ for $t \geq 1$)

$$\alpha\left[\frac{1}{2} - \frac{\sqrt{t+15}}{\sqrt{t+16}}\left[\sqrt{t+16} - \sqrt{t+15}\right]\right] \geq \frac{A}{c_0}.$$

Straightforward checking shows that $\max_{t\geq 16}\sqrt{\frac{t}{t+1}}\left(\sqrt{t+1} - \sqrt{t}\right) < 0.12$ and thus

$$c_0 \geq 2.7\frac{A}{\alpha} = 27L_{\mathcal{F}}R.$$

Hence, for $c_0 = 27L_{\mathcal{F}}R$ it holds that $g_{m+1}(x_{t+1}) \geq -c_0\eta_{t+1}$. We set c_0 to the maximum over the preceding two case, i.e.,

$$c_0 := \max\left\{7L_{\mathcal{F}}R,\ 27L_{\mathcal{F}}R\right\},$$

and obtain

$$g_{m+1}(x_t) \geq -c_0\eta_t = -\frac{27R^2}{\sqrt{t+15}}.$$

Part II) We now show that $\|x_{t+1}\| \leq 4R$. Combining **Part I)** and the definition of step size $\eta_t = \frac{1}{\alpha\sqrt{t+15}}$, we have

$$\frac{1}{2}\left[R^2 - \|x_{t+1}\|^2\right] = g_{m+1}(x_{t+1}) \geq -c_0\eta_{t+1} \geq -c_0\eta_1 = -\frac{c_0}{4\alpha}$$

and thus

$$\|x_{t+1}\|^2 \leq R^2 + \frac{c_0}{2\alpha} < 15R^2 < (4R)^2.$$

Part III) By Claim C.3, it follows that

$$\|v_{t+1}\| \leq \frac{L_{\mathcal{F}}}{R}\|x_{t+1}\| + 3L_{\mathcal{F}} < 7L_{\mathcal{F}}.$$

□

C.2 Concluding Remarks

By Lemma C.5, the decision sequence $\{x_t\}_{t\geq 1}$ is attracted to the hypersphere $\mathcal{B}_R$ and always stays inside a slightly larger hypersphere $\mathcal{B}_{4R}$.

Then, by Lemma B.3 applied with $c = 4$, $d = 15$, $\alpha = L_{\mathcal{F}}/R$ and step size $\eta_t = 1/(\alpha\sqrt{t+d})$ we obtain

$$\begin{aligned}\text{Regret}_T &\leq \sqrt{15+1}\left[(4+3)^2 + \frac{1}{2}(4+1)^2\right]L_{\mathcal{F}}R\sqrt{T} \\ &= 246L_{\mathcal{F}}R\sqrt{T}.\end{aligned}$$

Moreover, by Lemma B.5, we have $\mathcal{Z}_d = \frac{3}{2}\sqrt{17}\left[\frac{L_{\mathcal{G}}}{R} + \beta_{\mathcal{G}}\right]L_{\mathcal{F}}R$ and

$$c_1 = \mathcal{V}_\alpha\left[2L_{\mathcal{G}} + \frac{\beta_G\mathcal{V}_\alpha}{\alpha}\right] + \mathcal{Z}_d \leq 21\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_G\right]L_{\mathcal{F}}R.$$

Hence, the convergence rate to the feasible $\mathcal{C}$ satisfies for every $t \geq 1$ and $i \in \{1, \ldots, m\}$

$$g_i(x_t) \geq -c_1\eta_t \geq -21\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_G\right]\frac{R^2}{\sqrt{t+15}}.$$

D Proof of Theorem 3.3

In this section, we consider an online optimization problem with adversarially generated time-varying constraints. More precisely, at each time step t, the learner receives partial information on the current cost f_t and feasible set $\mathcal{C}_t$, and seeks to minimize (1). To make this problem well posed, we restrict the environment such that each feasible set $\mathcal{C}_t$ is contained in $\mathcal{Q}_t$ (see Section 1) and the rate of change between consecutive time-varying constraints *decreases* over time. We quantify a sufficient rate of decay in Assumption 3.1, which we restate below for the convenience of the reader.

Assumption D.1 (TVC Decay Rate)**.** We assume that the adversarially generated sequence $\{g_t\}_{t\geq 1}$ of time-varying constraints are such that for every $x \in \mathcal{B}_{4R}$ and all $t \geq 1$, the following holds $\|g_{t+1}(x) - g_t(x)\|_\infty \leq \frac{98}{t+16}\left[\frac{L_\mathcal{G}}{R} + 3\beta_\mathcal{G}\right]R^2$.

We note that Assumption D.1 essentially only requires $\|g_{t+1}(x) - g_t(x)\|_\infty \leq \mathcal{O}(1/t)$, as R can be chosen large enough such that the bound is satisfied. Of course, R will appear in our regret and feasibility bounds, but it will not affect the dependence on t or T (up to constant factors).

We restate Theorem 3.3 below for the convenience of the reader.

Theorem D.2 (Time-Varying Constraints)**.** *Suppose the functions $\{f_t, g_t\}_{t\geq 1}$ satisfy Assumptions 1.1, 1.2 and D.1. Then, on input $R, L_\mathcal{F} > 0$ and $x_1 \in \mathcal{B}_R$, Algorithm 1 applied with $\alpha = L_\mathcal{F}/R$, augmented velocity polytope $V'_\alpha(\cdot)$ and step sizes $\eta_t = \frac{1}{\alpha\sqrt{t+15}}$ guarantees the following for all $T \geq 1$:*

- ***(regret)*** $\sum_{t=1}^T f_t(x_t) - \min_{x^\star\in\mathcal{C}} \sum_{t=1}^T f_t(x^\star) \leq 246 L_\mathcal{F} R\sqrt{T}$;
- ***(feasibility)*** $g_{t,i}(x_t) \geq -265\left[\frac{L_\mathcal{G}}{R} + 4\beta_\mathcal{G}\right]\frac{R^2}{\sqrt{t+15}}$, *for all $t \in \{1,\dots,T\}$ and $i \in \{1,\dots,m\}$;*
- ***(attraction)*** $g_{m+1}(x_t) \geq -27\frac{R^2}{\sqrt{t+15}}$ *for all $t \in \{1,\dots,T\}$.*

Outline This section is organized as follows. In Subsection D.1, we introduce a key geometric property that allows us to generalize the standard online gradient descent analysis to the setting of time-varying constraints. In Subsection D.2, we give an overview of our proof approach for Theorem 3.3. In Subsection D.3, we present the analysis that quantifies the convergence rate to the feasible set for the setting of slowly time-varying constraints. Finally, in Subsection D.4, we give an important special case, slightly generalizing Lemma 3.2, for which Assumption D.1 is satisfied.

D.1 Key Geometric Property

Our regret analysis builds upon the following key geometric property that generalizes Lemma 2.3 to time-varying constraints. We show that for any subset $\mathcal{C}_T$ of the polyhedral intersection $\mathcal{Q}_T$, every decision $x \in \mathcal{C}_T$ satisfies the normal cone constraint $-r_t^\top(x - x_t) \leq 0$, for every pair (x_t, r_t) in the decision sequence $\{(x_t, r_t)\}_{t=1}^T$ up to step T. As a result, a similar argument as in (2) yields $\mathcal{O}(\sqrt{T})$ regret in the time-varying constraint setting.

Lemma D.3 (Polyhedral Intersection)**.** *Let $\mathcal{C}_T$ be any subset of the polyhedral intersection $\mathcal{Q}_T$. Then, every decision $x \in \mathcal{C}_T$ satisfies the normal cone constraint $-r_t^\top(x - x_t) \leq 0$, $\forall t \in \{1,\dots,T\}$.*

Proof. Using $\mathcal{S}_0 = \mathbb{R}^n$, $\mathcal{C}_T$ is contained in $\cap_{t=1}^{T-1}\{x \in \mathbb{R}^n \mid G(x_t)^\top(x - x_t) \geq 0\}$. Since $x \in \mathcal{C}_T$, it follows by Lemma 2.3 that $r_T^\top(x - x_T) \leq 0$. The proof proceeds by case distinction. Let $t \in \{1,\dots,T-1\}$ be arbitrary. Suppose $x_t \in \mathcal{C}_t$, then by Part 1 in the proof of Lemma 2.3 we have $r_t = 0$. Suppose $x_t \notin \mathcal{C}_t$, then $x \in \mathcal{C}_T$ implies $G(x_t)^\top(x - x_t) \geq 0$ or equivalently $\nabla g_{t,i}(x_t)^\top(x - x_t) \geq 0$ for all $i \in I(x_t)$. Since $v_t \in V_\alpha(x_t)$, it follows that $v(x) = v_t + x - x_t \in V_\alpha(x_t)$. Moreover, the vector $-r_t$ belongs to the normal cone $N_{V_\alpha(x_t)}(v_t)$, which implies $-r_t^\top(v - v_t) \leq 0$ for all $v \in V_\alpha(x_t)$. In particular, for $v(x)$ we have $-r_t^\top(x - x_t) \leq 0$. $\square$

D.2 Proof Overview of Theorem 3.3

By Assumption 1.1, the slowly time-varying constraints $g_{t,i}(x)$ are concave and $\beta_{\mathcal{G}}$-smooth such that $\|\nabla g_{t,i}(x)\| \leq L_{\mathcal{G}}$ for all $x \in \mathcal{B}_{4R}$, $t \geq 1$ and $i \in \{1, \ldots, m\}$. By construction, see Lemma C.5, $\eta_t = 1/(\alpha\sqrt{t+15})$, $\alpha = L_{\mathcal{F}}/R$ and $\mathcal{V}_\alpha = 7L_{\mathcal{F}}$ implies that $\eta_{t+1}\mathcal{V}_\alpha = 7R/\sqrt{t+16}$. We note that Lemma C.5 still holds for time-varying constraints, which implies $\|x_t\| \leq 4R$ and $\|v_t\| \leq 7L_{\mathcal{F}}$.

Further, by Assumption D.1 we have for every $x \in \mathcal{B}_{4R}$, $t \geq 1$ and $i \in \{1, \ldots, m\}$ that

$$|g_{t+1,i}(x) - g_{t,i}(x)| \leq \frac{98}{t+16}\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right] R^2 = 2\eta_{t+1}^2\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2. \tag{S5}$$

Then, applying the preceding inequality and using similar arguments as in Part 2) of Section 2.5, we give in Corollary D.5 bounds on the slowly time-varying constraints $g_{t,i}(x)$ from below. In particular, we show that

$$g_{t+1,i}(x_{t+1}) \geq (1-\alpha\eta_t)g_{t,i}(x_t) - \eta_t^2\left[2\frac{L_{\mathcal{G}}}{R} + 7\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2 \qquad \text{for all } i \in I(x_t),$$

and

$$g_{t+1,i}(x_{t+1}) \geq -\eta_{t+1}7\mathcal{V}_\alpha\left[L_{\mathcal{G}} + \frac{\beta_{\mathcal{G}}\mathcal{V}_\alpha}{4\alpha}\right] \qquad \text{for all } i \in \{1, \ldots, m\}\backslash I(x_t).$$

Using a similar inductive argument as in Lemma B.5, we show in Lemma D.4 that in the setting of slowly time-varying constraints, the following feasibility convergence rate holds

$$g_{t,i}(x_t) \geq -\left[265\frac{L_{\mathcal{G}}}{R} + 927\beta_{\mathcal{G}}\right]\frac{R^2}{\sqrt{t+15}}, \qquad \text{for all } t \in \{1, \ldots, T\} \text{ and } i \in \{1, \ldots, m\}.$$

Then, the regret and the attraction to the feasible sets follow as in Theorem C.1.

D.3 Slowly Time-Varying Constraints

Lemma D.4 (Slowly TVC)**.** *Suppose Assumption 1.1 holds, $x_1 \in \mathcal{B}_R$, $\alpha = L_{\mathcal{F}}/R$ and step sizes $\eta_t = 1/(\alpha\sqrt{t+15})$. Then, for every $i \in \{1, \ldots, m\}$ and $T \geq 1$ we have*

$$g_{t,i}(x_t) \geq -\left[265\frac{L_{\mathcal{G}}}{R} + 927\beta_{\mathcal{G}}\right]\frac{R^2}{\sqrt{t+15}}, \qquad \textit{for all } t \in \{1, \ldots, T\} \textit{ and } i \in \{1, \ldots, m\}.$$

Proof. The proof is by induction on t. We start with the base case $t = 1$. The proof proceeds by case distinction.

Case 1. Suppose $i \in \{1, \ldots, m\}\backslash I(x_1)$, i.e., $g_{1,i}(x_1) > 0$. Then, by Corollary D.5 Part ii) we have

$$g_{2,i}(x_2) \geq -\eta_2 7\mathcal{V}_\alpha\left[L_{\mathcal{G}} + \frac{\beta_{\mathcal{G}}\mathcal{V}_\alpha}{4\alpha}\right] \geq -\eta_2\left[49\frac{L_{\mathcal{G}}}{R} + 86\beta_{\mathcal{G}}\right]L_{\mathcal{F}}R.$$

Case 2. Suppose $i \in I(x_1)$, i.e., $g_{1,i}(x_1) \leq 0$. By combining $x_1 \in \mathcal{B}_R$ and $g_{1,i}$ is concave $\beta_{\mathcal{G}}$-smooth, it follows for every $x \in \mathcal{C}_1 \subseteq \mathcal{B}_R$ that

$$\begin{aligned} g_{1,i}(x_1) &\geq g_{1,i}(x) + \nabla g_{1,i}(x)^T(x_1 - x) - \frac{\beta_{\mathcal{G}}}{2}\|x_1 - x\|^2 \\ &\geq -2L_{\mathcal{G}}R - 2\beta_{\mathcal{G}}R^2 \\ &= -\eta_1\left[8\frac{L_{\mathcal{G}}}{R} + 8\beta_{\mathcal{G}}\right]L_{\mathcal{F}}R. \end{aligned}$$

Using $\eta_t = 1/(\alpha\sqrt{t+15})$ and $\eta_1/\eta_2 \leq \sqrt{2}$, it follows that

$$(1-\alpha\eta_1)g_{1,i}(x_1) \geq -\eta_1\left[\frac{L_{\mathcal{G}}}{R} + \beta_{\mathcal{G}}\right]6L_{\mathcal{F}}R \geq -\eta_2\left[9\frac{L_{\mathcal{G}}}{R} + 9\beta_{\mathcal{G}}\right]L_{\mathcal{F}}R$$

and

$$\eta_1^2\left[2\frac{L_{\mathcal{G}}}{R} + 7\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2 \leq \eta_2^2\left[4\frac{L_{\mathcal{G}}}{R} + 14\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2 \leq \eta_2\left[49\frac{L_{\mathcal{G}}}{R} + 172\beta_{\mathcal{G}}\right]L_{\mathcal{F}}R.$$

Then, by Corollary D.5 Part i) we have

$$\begin{aligned} g_{2,i}(x_2) &\geq (1-\alpha\eta_1)g_{1,i}(x_1) - \eta_1^2\left[2\frac{L_{\mathcal{G}}}{R} + 7\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2 \\ &\geq -\eta_2\left[58\frac{L_{\mathcal{G}}}{R} + 181\beta_{\mathcal{G}}\right]L_F R. \end{aligned}$$

Our inductive hypothesis is $g_{t,i}(x_t) \geq -c_2\eta_t$ for all i. We now show that it holds for $t+1$.

Case 1. Suppose $i \in \{1,\ldots,m\}\backslash I(x_1)$, i.e., $g_i(x_t) > 0$. Then by Corollary D.5 ii)

$$g_{t+1,i}(x_{t+1}) \geq -\eta_{t+1}7\mathcal{V}_\alpha\left[L_{\mathcal{G}} + \frac{\beta_{\mathcal{G}}\mathcal{V}_\alpha}{4\alpha}\right] \geq -\eta_{t+1}\left[49\frac{L_{\mathcal{G}}}{R} + 86\beta_{\mathcal{G}}\right]L_{\mathcal{F}}R.$$

Case 2. Suppose $i \in I(x_t)$, i.e., $g_i(x_t) \leq 0$. Let $A = \left[2\frac{L_{\mathcal{G}}}{R} + 7\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2$. By combining Corollary D.5 Part i), the inductive hypothesis and using similar arguments as in the proof of Lemma C.5 Case 2, yields

$$g_{t+1,i}(x_{t+1}) \geq -c_2\eta_{t+1}, \quad \text{where} \quad c_2 = 2.7\frac{A}{\alpha} = \left[265\frac{L_{\mathcal{G}}}{R} + 927\beta_{\mathcal{G}}\right]L_{\mathcal{F}}R.$$

The feasibility convergence rate is then given by

$$g_{t,i}(x_t) \geq -\left[265\frac{L_{\mathcal{G}}}{R} + 927\beta_{\mathcal{G}}\right]\frac{R^2}{\sqrt{t+15}}.$$

□

Corollary D.5. *Suppose Assumptions 1.1 and Assumption D.1 hold. Let $\alpha = L_{\mathcal{F}}/R$, $\mathcal{V}_\alpha = 7L_{\mathcal{F}}$ and step sizes $\eta_t = 1/(\alpha\sqrt{t+15})$. Then, for every $t \geq 1$ we have*

i) $g_{t+1,i}(x_{t+1}) \geq (1-\alpha\eta_t)g_{t,i}(x_t) - \eta_t^2\left[2\frac{L_{\mathcal{G}}}{R} + 7\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2$ for all $i \in I(x_t)$; and

ii) $g_{t+1,i}(x_{t+1}) \geq -\eta_{t+1}7\mathcal{V}_\alpha\left[L_{\mathcal{G}} + \frac{\beta_{\mathcal{G}}\mathcal{V}_\alpha}{4\alpha}\right]$ for all $i \in \{1,\ldots,m\}\backslash I(x_t)$.

Proof. Combining Assumption D.1 and (S5) gives

$$g_{t+1,i}(x_{t+1}) \geq g_{t,i}(x_{t+1}) - 2\eta_{t+1}^2\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2.$$

Then, by Claim B.6, it follows for every $i \in I(x_t)$ that

$$\begin{aligned} g_{t+1,i}(x_{t+1}) &\geq g_{t,i}(x_{t+1}) - 2\eta_{t+1}^2\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2. \\ &\geq (1-\alpha\eta_t)g_{t,i}(x_t) - \eta_t^2\frac{\mathcal{V}_\alpha^2\beta_{\mathcal{G}}}{2} - \eta_t^2\left[2\frac{L_{\mathcal{G}}}{R} + 6\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2 \\ &> (1-\alpha\eta_t)g_{t,i}(x_t) - \eta_t^2\left[2\frac{L_{\mathcal{G}}}{R} + 7\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2, \end{aligned}$$

and for every $i \in \{1,\ldots,m\}\backslash I(x_t)$ that

$$\begin{aligned} g_{t+1,i}(x_{t+1}) &\geq g_{t,i}(x_{t+1}) - 2\eta_{t+1}^2\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2 \\ &\geq -\eta_{t+1}\mathcal{V}_\alpha\left[2L_{\mathcal{G}} + \frac{\beta_{\mathcal{G}}\mathcal{V}_\alpha}{4\alpha}\right] - \eta_{t+1}\mathcal{V}_\alpha\left[\frac{L_{\mathcal{G}}\mathcal{V}_\alpha}{2\alpha R} + \frac{3\beta_{\mathcal{G}}\mathcal{V}_\alpha}{2\alpha}\right] \\ &\geq -\eta_{t+1}7\mathcal{V}_\alpha\left[L_{\mathcal{G}} + \frac{\beta_{\mathcal{G}}\mathcal{V}_\alpha}{4\alpha}\right]. \end{aligned}$$

where we used that $\alpha = L_{\mathcal{F}}/R$ and $\mathcal{V}_\alpha = 7L_{\mathcal{F}}$ implies $\frac{L_{\mathcal{G}}\mathcal{V}_\alpha}{R\alpha} = 7L_{\mathcal{G}}$. □

D.4 Average Time-Varying Constraints

An important special case where Assumption D.1 is satisfied, is summarized in the following slightly more general version of Lemma 3.2.

Lemma D.6. *Suppose the functions $\tilde{g}_{t,i}$ satisfy Assumption 1.1 and in addition there is a decision $x_{t,i} \in \mathcal{B}_R$ such that $|\widetilde{g}_{t,i}(x_{t,i})| \leq \frac{1}{2}\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]R^2$, for every $t \geq 1$ and $i \in \{1, \dots, m\}$. Then the following average time-varying constraints, satisfy Assumption 1.1 and Assumption D.1:*

$$g_{t,i}(x) := \frac{1}{t}\sum_{\ell=1}^{t} \widetilde{g}_{\ell,i}(x) \in \mathbb{R}^m. \tag{S6}$$

The rest of this subsection is devoted to proving Lemma D.6. We achieve this in two steps. We start by showing in Lemma D.7 that the average time-varying constraints satisfy Assumption 1.1, and then in Lemma D.8 we demonstrate that they also satisfy Assumption D.1.

Lemma D.7. *Suppose $\widetilde{g}_{t,i}$ is concave $\beta_{\mathcal{G}}$-smooth such that $\|\nabla \widetilde{g}_{t,i}(x)\| \leq L_{\mathcal{G}}$ for all $x \in \mathcal{B}_{4R}$, $t \geq 1$ and $i \in \{1, \dots, m\}$. Then, the average function*

$$g_{t,i}(x) := \frac{1}{t}\sum_{\ell=1}^{t} \widetilde{g}_{\ell,i}(x)$$

is concave and $\beta_{\mathcal{G}}$-smooth and $\|\nabla g_{t,i}(x)\| \leq L_{\mathcal{G}}$ holds for all $x \in \mathcal{B}_{4R}$, $t \geq 1$ and $i \in \{1, \dots, m\}$.

Proof. By assumption, each $\widetilde{g}_{\ell,i}$ is concave and $\beta_{\mathcal{G}}$-smooth, which implies

$$\widetilde{g}_{\ell,i}(x_{t+1}) \geq \widetilde{g}_{\ell,i}(x_t) + [\nabla \widetilde{g}_{\ell,i}(x_t)]^\top [x_{t+1} - x_t] - \frac{\beta_{\mathcal{G}}}{2}\|x_{t+1} - x_t\|^2.$$

Summing over all $\ell \in \{1, ..., t\}$ yields

$$\frac{1}{t}\sum_{\ell=1}^{t} \widetilde{g}_{\ell,i}(x_{t+1}) \geq \frac{1}{t}\sum_{\ell=1}^{t} \widetilde{g}_{\ell,i}(x_t) + \left[\frac{1}{t}\sum_{\ell=1}^{t} \nabla \widetilde{g}_{\ell,i}(x_t)\right]^\top [x_{t+1} - x_t] - \frac{1}{t}\sum_{\ell=1}^{t} \frac{\beta_{\mathcal{G}}}{2}\|x_{t+1} - x_t\|^2,$$

since $\frac{1}{t}\sum_{\ell=1}^{t} \nabla \widetilde{g}_{\ell,i}(x) = \nabla g_{t,i}(x)$, which is equivalent to

$$g_{t,i}(x_{t+1}) \geq g_{t,i}(x_t) + [\nabla g_{t,i}(x)]^\top [x_{t+1} - x_t] - \frac{\beta_{\mathcal{G}}}{2}\|x_{t+1} - x_t\|_2^2.$$

Hence, $g_{t,i}$ is concave and $\beta_{\mathcal{G}}$-smooth.

Moreover, since $\|\nabla \widetilde{g}_{t,i}(x)\| \leq L_{\mathcal{G}}$ for all $x \in \mathcal{B}_{4R}$, we have

$$\|\nabla g_{t,i}(x)\| = \left\|\frac{1}{t}\sum_{\ell=1}^{t} \nabla \widetilde{g}_{\ell,i}(x)\right\| \leq \frac{1}{t}\sum_{\ell=1}^{t} \|\nabla \widetilde{g}_{\ell,i}(x)\| \leq L_{\mathcal{G}}.$$

□

We show next that the average time-varying constraints satisfy Assumption D.1.

Lemma D.8 (Average TVC)**.** *Suppose $\widetilde{g}_{t,i}$ is concave $\beta_{\mathcal{G}}$-smooth such that $\|\nabla \widetilde{g}_{t,i}(x)\| \leq L_{\mathcal{G}}$ for all $x \in \mathcal{B}_{4R}$, $t \geq 1$ and $i \in \{1, \dots, m\}$. Further, suppose for every $t \geq 1$ and $i \in \{1, \dots, m\}$, there exists a decision $x_{t,i} \in \mathcal{B}_R$ such that*

$$|\widetilde{g}_{t,i}(x_{t,i})| \leq \frac{1}{2}\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]R^2. \tag{S7}$$

Then, for $\alpha = L_{\mathcal{F}}/R$, step sizes $\eta_t = 1/(\alpha\sqrt{t+15})$ and $\mathcal{V}_\alpha = 7L_{\mathcal{F}}$, it holds for every $x \in \mathcal{B}_{4R}$ that

$$|g_{t+1,i}(x) - g_{t,i}(x)| \leq 2\eta_{t+1}^2\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2.$$

Proof. Using the inequality $\frac{1}{t+1} \leq \frac{17}{2}\frac{1}{t+16}$ for every $t \geq 1$ and $\eta_{t+1}^2 = 1/(\alpha^2(t+16))$, it follows by construction that

$$\begin{aligned}
|g_{t+1,i}(x) - g_{t,i}(x)| &= \left|\frac{1}{t+1}\widetilde{g}_{t+1,i}(x) + \frac{t}{t+1}g_{t,i}(x) - g_{t,i}(x)\right| \\
&= \frac{1}{t+1}\left|\widetilde{g}_{t+1,i}(x) - g_{t,i}(x)\right| \\
&= \frac{1}{t+1}\frac{1}{t}\left|\sum_{\ell=1}^{t}\widetilde{g}_{t+1,i}(x) - \widetilde{g}_{\ell,i}(x)\right| \\
&\leq \eta_{t+1}^2\frac{17}{2}\alpha^2 \cdot \frac{1}{t}\sum_{\ell=1}^{t}\left|\widetilde{g}_{t+1,i}(x) - \widetilde{g}_{\ell,i}(x)\right|. \qquad \text{(S8)}
\end{aligned}$$

By triangle inequality $|\widetilde{g}_{t+1,i}(x) - \widetilde{g}_{\ell,i}(x)| \leq |\widetilde{g}_{t+1,i}(x)| + |\widetilde{g}_{\ell,i}(x)|$ and thus it suffices to bound the term $|\widetilde{g}_{t,i}(x)|$ for every $t \geq 1$, $i \in \{1, \ldots, m\}$ and $x \in \mathcal{B}_{4R}$.

By assumption, $x \in \mathcal{B}_{4R}$ and there is $x_{t,i} \in \mathcal{B}_R$ satisfying inequality (S7). Further, $\widetilde{g}_{t,i}$ is concave, which implies

$$\widetilde{g}_{t,i}(x) - \widetilde{g}_{t,i}(x_{t,i}) \leq [\nabla\widetilde{g}_{t,i}(x_{t,i})]^\top[x - x_{t,i}] \leq 5L_{\mathcal{G}}R$$

and the fact that $\widetilde{g}_{t,i}$ is concave $\beta_{\mathcal{G}}$-smooth yields

$$\begin{aligned}
\widetilde{g}_{t,i}(x) - \widetilde{g}_{t,i}(x_{t,i}) &\geq [\nabla\widetilde{g}_{t,i}(x_{t,i})]^\top[x - x_{t,i}] - \frac{\beta_{\mathcal{G}}}{2}\|x_{t,i} - x\|^2 \\
&\geq -5\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]R^2.
\end{aligned}$$

Further, by combining $\left|\widetilde{g}_{t,i}(x) - \widetilde{g}_{t,i}(x_{t,i})\right| \leq 5\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]R^2$, triangle inequality and assumption (S7), we obtain for every $x \in \mathcal{B}_{4R}$ that

$$\begin{aligned}
|\widetilde{g}_{t,i}(x)| &= |\widetilde{g}_{t,i}(x) - \widetilde{g}_{t,i}(x_{t,i}) + \widetilde{g}_{t,i}(x_{t,i})| \\
&\leq |\widetilde{g}_{t,i}(x) - \widetilde{g}_{t,i}(x_{t,i})| + |\widetilde{g}_{t,i}(x_{t,i})| \\
&\leq \frac{11}{2}\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]R^2.
\end{aligned}$$

The statement follows by combining $\alpha = L_{\mathcal{F}}/R$, $\mathcal{V}_\alpha = 7L_{\mathcal{F}}$, (S8) and

$$\begin{aligned}
|g_{t+1,i}(x) - g_{t,i}(x)| &\leq \eta_{t+1}^2\frac{17}{2}\alpha^2 \cdot \frac{1}{t}\sum_{\ell=1}^{t}|\widetilde{g}_{t+1,i}(x) - \widetilde{g}_{\ell,i}(x)| \\
&\leq \eta_{t+1}^2\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]\frac{11}{2} \cdot 17\alpha^2R^2 \\
&< 2\eta_{t+1}^2\left[\frac{L_{\mathcal{G}}}{R} + 3\beta_{\mathcal{G}}\right]\mathcal{V}_\alpha^2.
\end{aligned}$$

□