Online Convex Optimization with Ramp Constraints

Masoud Badiei[†], Na Li[†], Adam Wierman[*]
[†] Harvard University, Email:{mbadieik, nali}@seas.harvard.edu
[*] California Institute of Technology, Email: adamw@caltech.edu

***Abstract*—We study a novel variation of online convex optimization where the algorithm is subject to ramp constraints limiting the distance between consecutive actions. Our contribution is results providing asymptotically tight bounds on the worst-case performance, as measured by the competitive difference, of a variant of Model Predictive Control termed Averaging Fixed Horizon Control (AFHC). Additionally, we prove that AFHC achieves the asymptotically optimal achievable competitive difference within a general class of "forward looking" online algorithms. Furthermore, we illustrate that the performance of AFHC in practice is often much better than indicated by the (worst-case) competitive difference using a case study in the context of the economic dispatch problem.**

I. Introduction

In an online convex optimization (OCO) problem, an online learner interacts with an environment in a sequence of rounds. During each round t: (i) the learner must choose an action $\mathbf{x}_t = (x_t^1, x_t^2, \cdots, x_t^N)$ from a convex set $\mathscr{X}$; (ii) the environment reveals a cost loss function $c_t(\cdot) \in \mathscr{C}$; and (iii) the learner experiences cost $c_t(\mathbf{x}_t)$. The goal is to design learning algorithms that minimize the cost experienced over a (long) horizon of T rounds.

Online convex optimization has emerged as a fundamental tool in applications ranging from communication networks [1], [2] to cloud computing [3], [4] to portfolio optimization [5], [6]. As such, there is a large literature spanning machine learning, online algorithms, and operations research that seeks to develop and analyze algorithms for this setting [7], [8], [9]. Further, motivated by practical considerations, a large number of variations of online convex optimization have emerged, e.g., the addition of switching costs [3], the consideration of predictions [10], bandit formulations [11].

In this paper, we consider a novel variation of online convex optimization where the learner (algorithm) is subject to constraints bounding the distance between consecutive actions, i.e., is subject to *ramp constraints*. In particular, we consider a setting where the learner's actions during each round are additionally constrained by $-\underline{X}^i \leq x_t^i - x_{t-1}^i \leq \bar{X}^i$, for given ramp constraints $\underline{X}^i$ and $\bar{X}^i$.

Ramp constraints are common in a variety of practical settings. For example, in the economic dispatch problem, which plans the generation trajectories of power suppliers for the electricity grid, the ramping constraints of generators are of crucial importance [12], [13]. These limits are due to the thermal and mechanical inertia that has to be overtaken in order to ramp the output up or down, e.g., nuclear plants have strict ramping constraints that limit their use in providing response to fast timescale demand fluctuations. Another example in the energy domain is energy storage management, where the output power of a battery is subject to a hard constraint. Outside of energy, examples where ramp constraints are binding can be found in, e.g., portfolio optimization and network routing.

The addition of ramp constraints to the classical online convex optimization formulation adds considerable difficulty. In particular, without ramp constraints the current action and cost can be decoupled from future cost functions; however when ramp constraints are present it is necessary to ensure that the current action does not constrain future actions in a manner that leads to high costs.

Thus, one can expect that it is difficult to attain online algorithms with near-optimal performance without allowing the algorithm some access to predictions of future cost functions. Luckily, in many applications (e.g., economic dispatch and storage management) the learner has predictions/forecasts available.

Consequently, in this paper, we consider a situation where the learner has access to a limited number of future cost functions. Specifically, we allow the learner perfect lookahead for W time steps, i.e., when deciding $\mathbf{x}_t$ the learner has access to $c_t, \ldots, c_{t+W-1}$. This model is obviously overly optimistic about the quality of predictions available to the algorithm, but it is a common choice for studying the value of predictions in the online algorithms community, e.g., [14], [15], [3] so it is a natural starting point for the investigation of online convex optimization with ramp constraints. Note that it is typically straightforward to extend competitive analysis of online algorithms with perfect lookahead to the case of lookahead with bounded error.

Contributions of this paper: This paper initiates the study of online convex optimization with ramp constraints. Specifically, we focus our study on the analysis of a promising algorithm, named Averaging Fixed Horizon Control (AFHC), and derive (asymptotically) tight bounds on the competitive difference, i.e., the additive worst-case gap between the cost of obtained under AFHC and that obtained by the offline, static optimal solution.

Additionally, we introduce a general class of "forward looking" algorithms that include Model Predictive Control (MPC) and its variants, among other algorithms. We show that AFHC provides the asymptotically optimal competitive difference achievable by forward looking algorithms and, in fact, has competitive difference within a factor of 4 of any forward looking algorithm.

Finally, we illustrate that AFHC often performs much better than the (worst-case) bounds on competitive difference indicate. To do this, we perform a case study using the economic dispatch problem.

Related literature: Online convex optimization has been studied extensively, see [7] for a recent survey. While the addition of ramp constraints is novel, many variations that consider the addition of constraints of various forms to the optimizations have been studied previously. For example, [16] considers an online learning problem wherein the adversary reveals a constraint and a cost function at each step; [17] considers a variation in which the adversary is constrained in the actions that it can play at each round; and [18] illustrates that allowing the constraint violations results in a significant improvement in the achievable performance.

Perhaps the most closely related prior work is the literature studying online convex optimization with switching costs, which are often called "smoothed online convex optimization" problems [3], [19]. In such problems, there are no constraints, but there is an extra cost of $||\mathbf{x}_t - \mathbf{x}_{t-1}||$ incurred each round. Thus, like in the case of ramp constraints, when switching costs are present, the current action must consider future cost functions in order to avoid paying an avoidable switching cost in the next time step.

Clearly, the problems of online convex optimization with switching costs and with ramp constraints are related. But, note that the case of ramp constraints requires new analytic tools and leads to structurally different results compared to the case of switching costs. However, the literature studying switching costs provides us motivation for considering AFHC in the current paper. See the discussion in Section IV.

II. Problem Formulation

In this paper we consider a variation of online convex optimization where the algorithm is subject to ramp constraints that limit the maximal deviation in actions from one time step to the next. Clearly, one cannot hope for a online algorithm to perform well in this context if nothing is known about future cost functions, since the current choice places constraints on feasible choices in the next stage. Thus, we consider a situation where the algorithm has information about both the current cost function and a limited number of future cost functions. Specifically, the algorithm has perfect lookahead for W steps.

Formally, we consider online convex optimization over a finite time horizon $t \in [T]$. At each time step $t \in [T]$, a sequence of cost functions $(c_t, c_{t+1}, \cdots, c_{t+W-1})$ from a set $\mathscr{C}$ (finite or infinite) is revealed to the online algorithm. This W-lookahead window is "perfect" in the sense that these are the true cost functions the algorithm will experience in future time steps.

We assume that each cost function $c_t(\cdot) : \mathbb{R}^N \to \mathbb{R}$ in $\mathscr{C}$ is differentiable and convex. Furthermore, we suppose that the gradient of the cost function is bounded on the whole action space $\mathscr{X}$. In particular, we assume that there exists $Z_t \in \mathbb{R}_{\geq 0}$ for all $t \in [T]$ such that

$$\|\nabla c_t(\mathbf{x}_t)\|_\infty \leq Z_t, \quad \text{for all } \mathbf{x}_t \in \mathscr{X}, \tag{1}$$

where $\nabla c_t(\mathbf{x}_t) \triangleq (\partial c_t(\mathbf{x}_t)/\partial x_t^i)_{i=1}^N$. These assumptions are common in the literature, e.g., see [7].

When the cost functions further satisfy the strong convexity criterion, stronger results can be obtained.

Definition 1. *A cost function $c_t(\cdot)$ is strongly convex, if there exists $\sigma_t > 0$ such that for all the vectors $\mathbf{x}_t, \mathbf{y}_t \in \mathscr{X}$ it satisfies*

$$c_t(\mathbf{y}_t) - c_t(\mathbf{x}_t) \geq \nabla c_t(\mathbf{x}) \cdot (\mathbf{y}_t - \mathbf{x}_t)^T + \frac{\sigma_t}{2}\|\mathbf{y}_t - \mathbf{x}_t\|_2^2. \tag{2}$$

In the following, we define $\tilde{\mathscr{C}} \subset \mathscr{C}$ as the set of all strongly convex cost functions, i.e., for all $c_t(\cdot) \in \tilde{\mathscr{C}}$, the condition in (2) holds for some $\sigma_t > 0$.

At each time t, the online-decision maker takes an action $\mathbf{x}_t \triangleq (x_t^1, \ldots, x_t^N) \in \mathbb{R}^N$ from the action set $\mathscr{X} \subset \mathbb{R}^N$ after observing the cost functions $(c_t, \ldots, c_{t+W-1})$, i.e.,

$$\mathbf{x}_t \in \mathscr{X}, \tag{3}$$

where $\mathscr{X}$ is assumed to be convex. In the case that the action space $\mathscr{X}$ is bounded, we let $r > 0$ to be the radius of the largest ℓ_1-ball that contains it, i.e., $\mathscr{X} \subseteq \mathscr{B}_1(r) \triangleq \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\|_1 \leq r\}$.

We impose ramp constraints on consecutive actions. These take the following form

$$-\underline{X}^i \leq x_t^i - x_{t-1}^i \leq \bar{X}^i, \tag{4}$$

where, the initial condition for (5) is given by $\mathbf{x}_t = \mathbf{x}_0 \in \mathscr{X}$ for $t < 1$. Moreover, when $t > T$, we assume $\mathbf{x}_t$ is constant and equal to the value of $\mathbf{x}_T$.

The objective is to find a sequence of online actions $(\mathbf{x}_t)_{t=1}^T$ that yields the minimum cost subject to (3), (4). In particular, we define the online cost as

$$C_{\text{online}} = \sum_{t=1}^{T} c_t(\mathbf{x}_t), \tag{5}$$

where $(\mathbf{x}_t)_{t=1}^T$ are the actions that are taken by an online decision maker and satisfy (4). Further, we define $c_t(\cdot) \equiv 0$ for $t < 1, t > T$.

A. Performance measures

In the absence of the ramp rate constraint imposed on the actions, a greedy algorithm that optimizes the current cost function at each time step is optimal, in the sense that it achieves the optimal offline cost. However, in the presence of ramp constraints, taking an action that has immediate payoff can limit the action set for future decisions, which can result in poor performance. The focus of this paper is to understand the achievable performance of online algorithms relative to the cost of the optimal offline solution. To that end, we define the following optimal offline cost as our performance benchmark.

$$C_{\text{offline}} \triangleq \min_{\mathbf{x}_1, \cdots, \mathbf{x}_T} \sum_{t=1}^{T} c_t(\mathbf{x}_t), \quad \text{s.t. : (3), (4).} \tag{6}$$

Throughout, $(\mathbf{x}_t^\star)_{t=1}^T$ specifies the optimal offline solution.

As is standard in the online algorithms community, we use the competitive ratio in order to quantify the inefficiency of an online algorithm relative to the optimal offline solution.

Definition 2. *An online algorithm is α_T-competitive, if there exists a $\alpha_T \in \mathbb{R}_{\geq 0}$ such that*

$$\sup_{c_1, \cdots, c_t \in \mathscr{C}} \frac{C_{\textit{online}}}{C_{\textit{offline}}} \leq \alpha_T. \tag{7}$$

We achieve bounds on the competitive ratio through the derivation of bounds on the competitive difference.

Definition 3. *An online algorithm has a competitive difference of $\beta_T \in \mathbb{R}_{\geq 0}$ if*

$$\sup_{c_1,\cdots,c_t\in\mathscr{C}} (C_{online} - C_{offline}) \leq \beta_T. \tag{8}$$

Though the competitive difference is not as commonly used as the competitive ratio, it is more convenient for our context and it is straightforward to translate one into the other. Further, the competitive difference is more comparable to "regret", which is another common measure used to evaluate algorithms for online convex optimization, e.g., see [7]. Thus, we state our results in terms of competitive difference throughout this paper and leave it to the reader to interpret the results in terms of the competitive ratio. Note that, typically, the offline optimal cost is $\Omega(T)$, e.g., when there exists an e_0 such that $c_t(0) > e_0$ for all t as is imposed in [3].

III. An online algorithm

A wide variety of algorithms have been proposed for online convex optimization problems. The key differentiators of the context we consider in this paper are (i) ramp constraints and (ii) lookahead. Given these features of the model, perhaps the most natural candidates for consideration are algorithms from the control community such as Model Predictive Control (MPC) and its variations.

MPC, in its standard form, has been shown to perform poorly in the worst-case in the related problem of "smoothed" online convex optimization, which does not include ramp constraints but, instead, adds a regularizer to the objective. Specifically, MPC has a competitive ratio that does not improve as the lookahead window grows in high-dimensional smoothed online convex optimization problems [3].

However, a variation of MPC termed Averaging Fixed Horizon Control (AFHC) has much stronger worst-case guarantees. In particular, the competitive ratio of AFHC with lookahead W is $1 + O(1/W)$.[1] Thus, AFHC is a promising algorithm with which to begin the exploration of online convex optimization with ramp constraints.

A. Defining Averaging Fixed Horizon Control (AFHC)

As the name implies, AFHC averages the choices made by Fixed Horizon Control (FHC) algorithms. In particular, the decision $\mathbf{x}_t$ at the time slot $[t_n, t_n + W - 1]$ is made by averaging different solutions of the FHC algorithms corresponding to different initializations, i.e. $t_0 \in \mathscr{W}$.

Specifically, in an FHC algorithm, the action in the n^{th} round is characterized as the solution of the following optimization

$$(\hat{\mathbf{x}}_t^{t_0})_{t=t_n}^{t_n+W-1} = \arg \min_{\mathbf{x}_{t_n},\cdots,\mathbf{x}_{t_n+W-1}} \sum_{t=t_n}^{t_n+W-1} c_t(\mathbf{x}_t) \tag{9}$$
$$\text{s.t.} : (3), (4),$$

[1]Note that this result assumes that the action set is bounded, i.e., $r < \infty$, and that there exists $e_0 > 0$ s.t. $c_t(0) \geq e_0$ for all t. The results we prove here do not make these assumptions.

where the superscript in $\hat{\mathbf{x}}_t^{t_0}$ indicates the choice of parameter t_0. The initial condition for solving (9) is given by $\hat{\mathbf{x}}_{t_n-1}^{t_0}$, which is determined recursively by solving the optimization problem for $[t_{n-1}, t_{n-1} + W - 1]$.

By taking the average of $\hat{\mathbf{x}}_t^{t_0}$ over different values of $t_0 \in \mathscr{W}$, we obtain the AFHC solution. More precisely,

$$\hat{\mathbf{x}}_t \triangleq \frac{1}{W} \sum_{t_0\in\mathscr{W}} \hat{\mathbf{x}}_t^{t_0}, \tag{10}$$

for all $t \in [T]$. Let C_{AFHC} denote the overall cost of using the AFHC solution,

$$C_{\text{AFHC}} \triangleq \sum_{t=1}^{T} c_t(\hat{\mathbf{x}}_t).$$

B. Analyzing Averaging Fixed Horizon Control (AFHC)

In this section, we seek to understand how the performance of AFHC compares to that of the optimal offline solution. To that end, we prove upper and lower bounds on the competitive difference of AFHC.

1) Upper bounds on the competitive difference: Our first bound is stated in terms of the KKT multipliers associated with the constraints (3) and (4) in (9), which are denoted by $\boldsymbol{\lambda}_t^{t_0} \triangleq (\bar{\boldsymbol{\lambda}}_t^{t_0}, \underline{\boldsymbol{\lambda}}_t^{t_0})$, $\boldsymbol{\mu}_t^{t_0} \triangleq (\bar{\boldsymbol{\mu}}_t^{t_0}, \underline{\boldsymbol{\mu}}_t^{t_0})$, respectively.

Theorem 1. *The competitive difference of AFHC is upper bounded by*

$$C_{AFHC} - C_{offline} \leq \gamma \left\lfloor \frac{T}{W} \right\rfloor \min\left\{2r, \frac{\lfloor T/W \rfloor + 1}{2} W \|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1\right\}, \tag{11}$$

where $\gamma = \max_{t_0\in\mathscr{W}} \max\{\bar{\boldsymbol{\mu}}_t^{t_0}, \underline{\boldsymbol{\mu}}_t^{t_0}\}_{t=1}^T$. Furthermore, when $c_t(\cdot) \in \tilde{\mathscr{C}}$, the competitive difference is upper bounded by

$$C_{AFHC} - C_{offline} \tag{12}$$
$$\leq \gamma \min\left\{\frac{2NT\gamma}{W^2\sigma_{min}}, 2\left\lfloor \frac{T}{W} \right\rfloor r, \frac{\lfloor T/W \rfloor(\lfloor T/W \rfloor + 1)}{2} W \|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1\right\}.$$

A few remarks on Theorem 1 are in order.

First, if we assume that the optimal offline solution has cost $\Omega(T)$ and we consider the a bounded action space, i.e., $r < \infty$, then the general upper bound in (11) yields a bound on the competitive difference that scales as $O(T/W)$ and on the competitive ratio that scales as $O(1/W)$. Interestingly, this is the same scaling that AFHC achieves for the competitive ratio under these assumptions in the related problem of smoothed online convex optimization [3], [19]. Furthermore, if strong convexity of the cost functions is assumed, the upper bound improves to $O(1/W^2)$. In contrast, if the action space is unbounded, $r = \infty$ and then the bound on the competitive difference scales as $O(T^2/W)$.

Second, note that when the ramp constraints are loosened, i.e., $\bar{X}^i, \underline{X}^i \to \infty, \forall i \in [N]$, $\gamma \to 0$. Therefore, $C_{\text{AFHC}} = C_{\text{offline}}$. This highlights the tightness of the bound since, without the coupling between the actions provided by ramp constraints, the optimization problem in (6) is separable, and the decision vector $\mathbf{x}_t$ can be computed independently at each time $t \in [1, T]$ by optimizing $c_t(\mathbf{x}_t)$ with respect to the constraint (3).

Third, it is not immediately obvious whether the KKT multipliers of AFHC are bounded. In fact, without an upper bound on γ the competitive difference bounds in (11) and (12) can potentially be very loose. In the next theorem, we demonstrate that γ is indeed bounded, and the upper bound depends on the window size W as well as the parameter Z_t in (1). To state the next theorem, we consider a specific set of constraints on the action space, i.e.,

$$\mathscr{X} \triangleq \{\mathbf{x} \in \mathbb{R}^N : -\underline{x}^i \le x_t^i - x_{t-1}^i \le \bar{x}^i\}. \quad (13)$$

Theorem 2. *For all $t \in [t_n, t_n + W - 1]$, the KKT multipliers of the optimization problem in* (9) *with the set of constraints in* (13) *are bounded as follows*

$$0 \preceq \bar{\boldsymbol{\lambda}}_t^{t_0} \preceq Z_t \mathbf{I}, \quad 0 \preceq \underline{\boldsymbol{\lambda}}_t^{t_0} \preceq Z_t \mathbf{I},$$
$$0 \preceq \bar{\boldsymbol{\mu}}_t^{t_0} \preceq \sum_{\tau=t}^{t_n+W-1} Z_\tau \mathbf{I}, \quad 0 \preceq \underline{\boldsymbol{\mu}}_t^{t_0} \preceq \sum_{\tau=t}^{t_n+W-1} Z_\tau \mathbf{I}.$$

The consequence of the above theorem is that

$$\gamma = \max_{t_0 \in \mathscr{W}} \max_{t_1,\cdots,t_M} \max \Big\{ \sum_{\tau=t_n}^{t_n+W-1} Z_\tau \Big\}.$$

2) Lower bounds on the competitive difference: We now turn out attention from upper bounds to lower bounds and address the question of whether the upper bounds we have obtained are tight.

Theorem 3. *Suppose the action space is unbounded, i.e., $\mathscr{X} \in \mathbb{R}^N$. Then, the competitive difference of AFHC has a lower bound as*

$$\gamma \frac{\lfloor T/W \rfloor (\lfloor T/W \rfloor + 1)}{2} \frac{W-1}{2} \|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1 \le \sup_{c_1,\cdots,c_T \in \mathscr{C}} (C_{AFHC} - C_{offline}).$$

Moreover, in the special case of $W = 1$, the lower bound takes the following form

$$\gamma \lfloor T/2 \rfloor (\lfloor T/2 \rfloor + 1) \min\{\|\bar{\mathbf{X}}\|_1, \|\underline{\mathbf{X}}\|_1\} \le \sup_{c_1,\cdots,c_T \in \mathscr{C}} (C_{AFHC} - C_{offline}).$$

Note that the lower bound assumes that the action space is unbounded, i.e., $r = \infty$. Thus, compared to the upper bound in (11), the lower bound is tight to within a constant factor of $1/2$. This highlights that the asymptotic scaling of the upper bound is correct. Interestingly, to obtain this theorem we construct a sequence of *linear* cost functions that exploits the structure of AFHC to enforce a large cost on the algorithm (cf. Appendix for details). Thus, the cost functions in the construction are not strictly convex.

C. Beyond Ramp Constraints

While our focus in this paper is on online convex optimization with ramp constraints, our analysis can extend to more general forms of constraints. In particular, the upper bound in (11) can be extended to general constraint sets.

Suppose the decision vectors satisfy inequality constraints of the form of

$$f^k(\mathbf{x}_t) \le 0, \quad k \in [J], \quad (14)$$

where $f^k(\cdot) : \mathbb{R}^N \to \mathbb{R}$ are convex functions. Also, suppose that the domain $\mathscr{X} \triangleq \{\mathbf{x} \in \mathbb{R}^N : f^k(\mathbf{x}) \le 0, k \in [J]\}$ defined by the inequality constraints in (14) is bounded, and let r to be the radius of the largest ball that contains $\mathscr{X}$, i.e., $\mathscr{X} \subseteq \mathscr{B}_2(r)$.[2] Moreover, instead of the ramp limits given in (4), consider a more general setup,

$$g_t(\mathbf{x}_t, \mathbf{x}_{t-1}) \le 0, \quad (15)$$

where $g_t(\mathbf{x}_t, \cdot)$ is Lipschitz continuous in the sense of ℓ_1-norm, i.e., $\forall \mathbf{x}_t, \mathbf{x}_{t-1}, \mathbf{y}_{t-1} \in \mathscr{X}$, there exists $L_t > 0$ such that

$$|g_t(\mathbf{x}_t, \mathbf{x}_{t-1}) - g_t(\mathbf{x}_t, \mathbf{y}_{t-1})| \le L_t \|\mathbf{x}_{t-1} - \mathbf{y}_{t-1}\|_1. \quad (16)$$

We denote the Lagrange multiplier associated with (15) by μ_t. In this setting, Theorem 1 can be extended as follows.

Theorem 4. *For all sequences of the convex cost functions, under the constraints in* (14), (15), *and* (16), *AFHC has a competitive difference upper bounded by*

$$C_{AFHC} - C_{offline} \le \frac{2T\gamma}{W} rL\sqrt{N},$$

where L is such that $\frac{1}{T}\sum_{t=1}^T L_t \le L$. In addition, if $c_t(\cdot) \in \tilde{\mathscr{C}}, \forall t \in [T]$, we have

$$C_{AFHC} - C_{offline} \le \min \Big\{ \frac{2T\gamma}{W} \sqrt{N} rL, \frac{2T\gamma^2}{W^2 \sigma_{min}} NL \Big\}, \quad (17)$$

where $\gamma = \max_{t_0 \in \mathscr{W}} \max_{t \in [T]} \mu_t^{t_0}$.

Note that, similar to Theorem 1, further regularity conditions on the gradients of $f(\mathbf{x}_t)$ and $g(\mathbf{x}_t, \mathbf{x}_{t-1})$ are required to guarantee that γ is bounded.

IV. A General Lower Bound

Our goal in this section is to understand whether there are online algorithms that can significantly outperform AFHC. With this motivation, we seek a general lower bound on the competitive difference of any online algorithm. While we cannot achieve this goal in full, we what we can achieve is a lower bound on the competitive difference of a broad class of "forward looking" algorithms that includes, among others, MPC and its many variants.

A. A class of forward looking algorithms

In general, at each time t, an algorithm has information about the entire set of reveled cost functions, i.e., $(c_1, c_2, \cdots, c_{t+W-1})$. However, the class of forward looking algorithms we consider is limited to using only a subset of these cost functions in its decision making. In particular, at each time t, the algorithm may only use cost functions $\{c_\tau\}_{\tau=t-M}^{t+W-1}$. Thus, forward looking algorithms make use of only a limited number M of past cost functions.

[2]Without loss of generality, we assume that the ball is centered at the origin. If this is not the case, we can always translate $\mathscr{X}$ so that it includes the origin.

The informational constraint above is the key piece of the definition of forward looking algorithms. However, for technical reasons, we need to additionally impose a minor constraint on how the algorithms may use the cost functions.

Let $\mathbf{x}_t$ to denote the actions a forward looking algorithm. Let a be any time slot such that $t-M \leq a \leq t$, b be any time slot such that $t \leq b \leq t+W-1$, $\tilde{\mathbf{x}}_{a-1}$ be any action in $\mathscr{X}$, $\tilde{c}_a, \tilde{c}_{a+1}, \ldots, \tilde{c}_b$ be a series of cost functions from $\{c_\tau\}_{\tau=t-M}^{t+W-1}$ for all $\tau \in [a, b]$. Then consider the following optimization problem,

$$\min_{\tilde{\mathbf{x}}_\tau, \tau=a,\ldots,b} \quad \tilde{c}_t(\tilde{\mathbf{x}}_\tau) \qquad \text{s.t.} \quad (3), (4). \tag{18}$$

If the optimal solutions for all $a, b, \tilde{\mathbf{x}}_{a-1}, \tilde{c}$ have the same Δ_t, then the decision maker will take action $\mathbf{x}_t = \mathbf{x}_{t-1} + \Delta_t$.

This constraint ensures that, no matter how the algorithm uses the cost function $\{c_\tau\}_{\tau=t-M}^{t+W-1}$, i.e., either by permuting or throwing away cost function, if the "optimal" actions have the same ramping decision, i.e., $\Delta_t \triangleq \mathbf{x}_t - \mathbf{x}_{t-1}$, then the decision maker will take $\mathbf{x}_t = \mathbf{x}_{t-1} + \Delta_t$. Note that this condition does not constrain the class of online algorithms dramatically since it is very unlikely to be binding. In particular, it only applies if all the optimal solutions for (18) have the same Δ_t.

B. A lower bound on the competitive difference of forward looking algorithms

Though the class of forward looking algorithms is broad, e.g., including MPC and its variants, we can prove a bound on the competitive difference for all such algorithms that asymptotically matches the upper bound in Theorem 1 on the competitive difference of AFHC.

Theorem 5. *For any forward looking online algorithm defined on $\mathscr{X} = \mathbb{R}^N$, the competitive difference is lower bounded by:*

$$\rho\Big(\frac{MT}{2(M+W)} + \frac{T}{4}(1 + \frac{T}{2(M+W)}\Big)\|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1 \leq \sup_{c_1,\ldots,c_T \in \mathscr{C}} (C_{online} - C_{offline}), \tag{19}$$

where $\rho \triangleq \max_{t\in[T]} Z_t$.

Importantly, this general lower bound is within 1/4 of the upper bound in Theorem 1, which highlights that no forward looking algorithm can significantly outperform AFHC. Moreover, in most situations the bound is even tighter. In particular, $\rho = \max_{t\in[T]} Z_t$ while γ is the maximum sum of Z_t over a window of size W, which when divided by the extra factor of $1/W$ in (11), results in the maximum average of Z_t. Consequently, γ is potentially much smaller than ρ when Z_t has a large variation with respect to t and/or W is very large.

V. A Case Study: Economic Dispatch

In order to ground the theoretical work in this paper, we consider a case study with which to illustrate the performance of AFHC in practice. In particular, we use the economic dispatch problem as an illustrative example.

The objective of the economic dispatch problem is to schedule the committed generating units to supply the load demand with minimum cost while satisfying the operational constraints, see [20], [21]. The challenge in derives from the fact that generators have hard ramp constraints which limit the change in generation feasible between epochs.

Note that the focus of this paper is *not* on the economic dispatch problem, rather it is on the study of general online convex optimization with ramp constraints. Thus, the goal of this case study is (i) to emphasize the importance of ramp constraints in real-world example and (ii) to illustrate that AFHC often performs much better than the worst-case bound on the competitive difference suggest. Consequently, we consider a relatively simple variation of the economic dispatch problem and ignore many issues that one would need to consider before actually applying the ideas here in practice.

A. Experimental Setup

We consider a power network that has a wind power generator as a supplement to three hydro-thermal power generators ($N = 3$). We denote the power outputs of the hydro-thermal units by $x_t^i, i \in \{1, 2, 3\}$. Table I shows the bounds on each power station. The initial values $x_0^i, i \in \{1, 2, 3\}$ are chosen in a way that the load demand at $t = 1$ can be satisfied.

In line with the other works on the economic dispatch problem [12], [22], [23] we choose a polynomial cost function with time independent coefficients for the power generation units, i.e., $c_t(\mathbf{x}_t) = \sum_{i=1}^N \left(a^i (x_t^i)^2 + b^i x_t^i + c^i\right)$. The values for the coefficients a^i, b^i and c^i for each power generation unit are specified in Table I (cf. [12]).

We use d_t to denote the aggregated load demand at time $t \in [T]$. Fig. 1(a) shows the real data for d_t over the course of 5 days, where the dispatch is performed at every 5 minutes. We denote the net amount of power generated by wind power stations by r_t. Fig. 1(b) shows the amount of wind power that is generated over the same duration of 5 days. The data for wind power data is interpolated to obtain the amount of wind power generation for every 5 minutes.

The objective is to supply sufficient power to meet the demand at all the times, i.e.,

$$\sum_{i=1}^N x_t^i + r_t = d_t, \tag{20}$$

where $x_t^i, i \in [N]$ are bounded variables as in (3) and satisfy the ramp constraints as in (4).

Since r_t and d_t are both random processes, we can combine their net effect into a unified process, called the residual demand process $\hat{d}_t \triangleq \max\{0, d_t - r_t\}$. Condition (20) then simplifies to

$$\sum_{i=1}^N x_t^i = \hat{d}_t. \tag{21}$$

To satisfy (21), we consider the following optimization problem

$$\min_{\mathbf{x}_1,\cdots,\mathbf{x}_T} \sum_{t=1}^T \sum_{i=1}^N \left(a^i (x_t^i)^2 + b^i x_t^i + c^i\right) + \sum_{t=1}^T \eta_t \left(\hat{d}_t - \sum_{i=1}^N x_t^i\right)^2$$

s.t. : (3), (4),

Table I: The hypothetical cost and operational capacities of the hydro-thermal units.

Dispatcher	$\bar{x}^i$(MW)	$-\underline{x}^i$(MW)	a^i (\$ MW^{-2})	b^i (\$ MW^{-1})	c^i(\$)
Gas: x_t^1	5×10^3	1×10^3	4.87×10^{-5}	10.7583	143.5972
Hydro: x_t^2	5×10^3	3×10^3	2.63×10^{-5}	23.2	260.17
Nuclear: x_t^3	1.2×10^4	1×10^4	1.95×10^{-7}	7.5031	311.9102

where $\eta_t \in \mathbb{R}_{\geq 0}$ is a factor that determines the cost of violating (21) at time $t \in [T]$. For simplicity of analysis, we set $\eta_t = \eta$ for all $t \in [T]$.

Note that in formulating the objective we assumed that when the generated power $\sum_{i=1}^{N} x_t^i$ exceeds the residual demand $\hat{d}_t$ the regularizer does not vanish. This is justified by the fact that if the excess power is not exported, it must be stored which in turn bears the storage cost. In practice, it is more satisfactory to produce the power slightly above the demand level. This can however be included in our model readily by adding a constant margin to the residual demand $\hat{d}_t$.

B. Numerical Results

For our numerical results, we apply AFHC to the problem specified in the previous section.

Our experimental results consider the case of symmetric ramp rate limits, i.e., the case when increasing or decreasing the power output is subject to the same ramp rate limits $\bar{X}^i = \underline{X}^i = X^i$. For this scenario, we define the ratio of the ramp constraint to the range of each variable by $R^i \triangleq X^i/(\bar{x}^i + \underline{x}^i)$.

In the case of $R^1 = R^2 = R^3 = R$, Fig. 1(c) shows the ratio of $C_{\text{AFHC}}/C_{\text{offline}}$ as a function of R for all $0 \leq R \leq 0.1$. For all values of R that are larger than 0.1, the ratio is unity. To plot the figure, we assumed that AFHC has a prediction window size of $W = 1$ hour. It is evident from Fig. 1(c) that for stringent ramp rate limits, the performance of AFHC deteriorates. This can be explained by the fact that when the ramp rate limit is strict, it becomes more difficult for any online algorithm to adjust its trajectory after taking a suboptimal action.

We additionally note that in the special case of $R = 0$, both online and offline solutions are constant and equal to their initial value. As a result $C_{\text{AFHC}}/C_{\text{offline}} = 1$ when $R = 0$.

Lastly, Fig. 1(d) shows the trajectories of each algorithm to satisfy the residual demand $\hat{d}_t$ for the worst case value of the ramp constraints (i.e., $R = 0.009$). The similarity of the two trajectories is immediately evident.

VI. Conclusion

In this paper we have introduced a variation of online convex optimization where the actions are constrained by limits on the ramp rate and where the online algorithm has a W-step (perfect) lookahead window. This problem is motivated by, among other applications, the economic dispatch problem, for which ramp constraints are a fundamental challenge.

Our main results focus on a relatively new algorithm, Averaging Fixed Horizon Control (AFHC). We prove upper and lower bounds on the competitive difference of AFHC that are tight up to a factor of 2. Additionally, we prove a general lower bound on "forward looking" algorithms which shows that no such algorithm can have a competitive difference more than a factor of 4 better than AFHC, and that in most cases this factor is even smaller. Further, we have illustrated the performance of AFHC in the context of the economic dispatch problem, which highlights that the typical performance of AFHC is much better than indicated by the (worst-case) competitive difference.

Our results are a first step toward the design of algorithms for online convex optimization problems that are subject to ramp constraints, and there are many open questions that remain. For example, we have adopted the model of perfect lookahead here due to its popularity in the online algorithms community but, of course, this model misses many important features of predictions. Considering a more realistic prediction model, such as [10], is an important next step. Additionally, while AFHC has a competitive difference that is asymptotically tight among "forward looking" algorithms, there may be other algorithms that make use of historical cost functions to improve on the competitive difference and there may be forward looking algorithms that can improve on AFHC by a factor of (up to) four.

References

[1] N. Bansal, A. Blum, S. Chawla, and A. Meyerson, "Online oblivious routing," in *Proceedings of the fifteenth annual ACM symposium on Parallel algorithms and architectures*. ACM, 2003, pp. 44–49.

[2] B. Awerbuch and R. Kleinberg, "Online linear optimization and adaptive routing," *Journal of Computer and System Sciences*, vol. 74, no. 1, pp. 97–114, 2008.

[3] M. Lin, Z. Liu, A. Wierman, and L. L. Andrew, "Online algorithms for geographical load balancing," in *Green Computing Conference (IGCC), 2012 International*. IEEE, 2012, pp. 1–10.

[4] M. Lin, A. Wierman, L. L. Andrew, and E. Thereska, "Dynamic right-sizing for power-proportional data centers," *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 5, pp. 1378–1391, 2013.

[5] T. M. Cover, "Universal portfolios," *Mathematical finance*, vol. 1, no. 1, pp. 1–29, 1991.

[6] A. Blum, V. Kumar, A. Rudra, and F. Wu, "Online learning in online auctions," *Theoretical Computer Science*, vol. 324, no. 2, pp. 137–146, 2004.

[7] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2011.

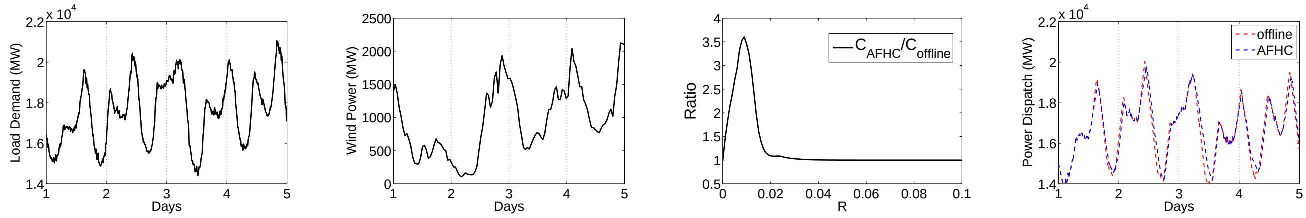
[8] E. Hazan and S. Kale, "Beyond the regret minimization barrier: Optimal algorithms for stochastic strongly-convex optimization," *Journal of Machine Learning Research*, vol. 15, pp. 2489–2512, 2014. [Online]. Available: http://jmlr.org/papers/v15/hazan14a.html

[9] E. Hazan, "10 the convex optimization approach to regret minimization," *Optimization for machine learning*, p. 287, 2012.

[10] N. Chen, A. Agarwal, A. Wierman, S. Barman, and A. Lachlan, "Online convex optimization using predictions," in *Proceeding of ACM SIGMETRICS Performance Evaluation Review*. ACM, 2015.

[11] S. Bubeck and N. Cesa-Bianchi, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[12] C. Wang and S. Shahidehpour, "Effects of ramp-rate limits on unit commitment and economic dispatch," *Power Systems, IEEE Transactions on*, vol. 8, no. 3, pp. 1341–1350, 1993.



(a) (b) (c) (d)

Figure 1: (a) The load demand profile of Ontario, Canada during March 1st to March 5th [24] (b): The wind power generation of Ontario, Canada during March 1st to March 5th [24] (c): The ratio $C_{\text{AFHC}}/C_{\text{offline}}$ for different values of the ramp constraint, i.e., different values of R (d): The trajectories of the offline and AFHC solutions.

[13] A. Frangioni, C. Gentile, and F. Lacalandra, "Solving unit commitment problems with general ramp constraints," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 5, pp. 316–326, 2008.

[14] A. Borodin, N. Linial, and M. E. Saks, "An optimal on-line algorithm for metrical task system," *Journal of the ACM (JACM)*, vol. 39, no. 4, pp. 745–763, 1992.

[15] J. Camacho, Y. Zhang, M. Chen, and D. M. Chiu, "Balance your bids before your bits: The economics of geographic load-balancing," in *Proceedings of the 5th international conference on Future energy systems*. ACM, 2014, pp. 75–85.

[16] B. Kveton, J. Y. Yu, G. Theocharous, and S. Mannor, "A lazy approach to online learning with constraints." in *ISAIM*. Citeseer, 2008.

[17] S. Mannor, J. N. Tsitsiklis, and J. Y. Yu, "Online learning with sample path constraints," *The Journal of Machine Learning Research*, vol. 10, pp. 569–590, 2009.

[18] M. Mahdavi, R. Jin, and T. Yang, "Trading regret for efficiency: online convex optimization with long term constraints," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 2503–2528, 2012.

[19] L. Andrew, S. Barman, K. Ligett, M. Lin, A. Meyerson, A. Roytman, and A. Wierman, "A tale of two metrics: simultaneous bounds on competitiveness and regret," in *Proceedings of COLT*, 2014.

[20] B. H. Chowdhury and S. Rahman, "A review of recent advances in economic dispatch." Institute of Electrical and Electronics Engineers, 1990.

[21] H. Happ, "Optimal power dispatch: A comprehensive survey," *Power Apparatus and Systems, IEEE Transactions on*, vol. 96, no. 3, pp. 841–854, 1977.

[22] S. Mei, Y. Wang, and Z. Sun, "Robust economic dispatch considering renewable generation," in *Innovative Smart Grid Technologies Asia (ISGT), 2011 IEEE PES*. IEEE, 2011, pp. 1–5.

[23] L. dos Santos Coelho and C.-S. Lee, "Solving economic load dispatch problems in power systems using chaotic and gaussian particle swarm optimization approaches," *International Journal of Electrical Power and Energy Systems*, vol. 30, no. 5, pp. 297–307, 2008.

[24] [online]. Available: http://www.ieso.ca.

[25] M. A. Epelman. (2012) Barrier methods for constrained optimization. [Online]. Available: http://www-personal.umich.edu/ mepelman/teaching/IOE511/Handouts/511notes07-11.pdf

VII. APPENDIX

A. Proof of Theorem 1

First we state a technical lemma. The proof is omitted due to space constraints.

Lemma 1. *Lets, $u, v, x, y, z \in \mathbb{R}^+$. Then, we have the following inequality*

$$(x-y-z)^+ + (y-x-u)^+ \le (v-y-z)^+ + (y-v-u)^+ + |x-v|,$$

where $x^+ \triangleq \min\{x, 0\}$.

Next, now define the cost function $C^{t_0}_{\text{FHC}}(n) = \sum_{t=t_n}^{t_n+W-1} c_t(\hat{\mathbf{x}}^{t_0}_t)$, where $\hat{x}^{i,t_0}_t$ is the solution given by (9). We use the penalty function method to define a relaxed version of the optimization problem in (9) as below

$$C^{t_0}_{\text{FHC}-\gamma^{t_0}_n}(n) = \min_{\mathbf{x}_{t_n},\cdots,\mathbf{x}_{t_n+W-1}} \sum_{t=t_n}^{t_n+W-1} \big(c_t(\mathbf{x}_t) + \gamma^{t_0}_n p(\mathbf{x}_t)\big), \quad (22)$$

where $\gamma^{t_0}_n \in \mathbb{R}_{\ge 0}$, and

$$p(\mathbf{x}_t) \triangleq \sum_{i=1}^{N} (x^i_t - x^i_{t-1} - \bar{X}^i)^+ + \sum_{i=1}^{N} (x^i_{t-1} - x^i_t - \underline{X}^i)^+.$$

The following lemma is shown in [25, p. 63]:

Lemma 2. *Suppose $\gamma^{t_0}_n \ge \max\{\bar{\boldsymbol{\mu}}^{t_0}_t, \underline{\boldsymbol{\mu}}^{t_0}_t\}_{t=t_n}^{t_n+W-1}$. Then, with the same initial condition, the set of optimal solutions of* (9) *and* (22) *coincide. Specifically, we have*

$$C^{t_0}_{FHC}(n) = C^{t_0}_{FHC-\gamma^{t_0}_n}(n).$$

Based on Lemma 2 we now obtain that

$$C^{t_0}_{\text{FHC}}(n) = C^{t_0}_{\text{FHC}-\gamma^{t_0}_n}(n) \quad (23)$$

$$\le \sum_{t=t_n}^{t_n+W-1} c_t(\mathbf{x}^\star_t) + \gamma^{t_0}_n \sum_{t=t_n+1}^{t_n+W-1} p(\mathbf{x}^\star_t) \quad (24)$$

$$+ \gamma^{t_0}_n \sum_{i=1}^{N} \big((x^{i,\star}_{t_n} - \hat{x}^{i,t_0}_{t_n-1} - \bar{X}^i)^+ + (\hat{x}^{i,t_0}_{t_n-1} - x^{i,\star}_{t_n} - \underline{X}^i)^+\big),$$

where the inequality follows from the fact that the offline trajectory $(\mathbf{x}^\star_t)_{t=1}^T$ for the optimization problem in (9) with the initial condition $\hat{\mathbf{x}}^{t_0}_{t_n-1}$ is sub-optimal. Using Lemma 1, we further derive

$$\begin{aligned}&(x^{i,\star}_{t_n} - \hat{x}^{i,t_0}_{t_n-1} - \bar{X}^i)^+ + (\hat{x}^{i,t_0}_{t_n-1} - x^{i,\star}_{t_n} - \underline{X}^i)^+\\ &\le (x^{i,\star}_{t_n} - x^{i,\star}_{t_n-1} - \bar{X}^i)^+ + (x^{i,\star}_{t_n-1} - x^{i,\star}_{t_n} - \underline{X}^i)^+\\ &\qquad + |\hat{x}^{i,t_0}_{t_n-1} - x^{i,\star}_{t_n-1}|. \quad (25)\end{aligned}$$

In the case of $n = 0$, we have $\hat{x}^{i,t_0}_{t_0-1} = x^{i,\star}_{t_0-1} = x^i_0$ and thus, the above inequality can be replaced by the equality without the third term.

We substitute (25) into (24) to obtain,

$$C^{t_0}_{\text{FHC}}(n) \le \sum_{t=t_n}^{t_n+W-1} c_t(\mathbf{x}^\star_t) + \gamma^{t_0}_n p(\mathbf{x}^\star_t) + \gamma^{t_0}_n \|\hat{\mathbf{x}}^{t_0}_{t_n-1} - \mathbf{x}^\star_{t_n-1}\|_1$$

$$\overset{(a)}{=} \sum_{t=t_n}^{t_n+W-1} c_t(\mathbf{x}^\star_t) + \gamma^{t_0}_n \|\hat{\mathbf{x}}^{t_0}_{t_n-1} - \mathbf{x}^\star_{t_n-1}\|_1,$$

where in (a) comes from the fact that the penalty function for the offline solution is zero, i.e. $p(\mathbf{x}_t^\star) = 0$.

We use the two upper bounds to bound $\|\hat{\mathbf{x}}_{t_n-1}^{t_0} - \mathbf{x}_{t_n-1}^\star\|_1$. The first bound is trivially given by

$$\|\hat{\mathbf{x}}_{t_n-1}^{t_0} - \mathbf{x}_{t_n-1}^\star\|_1 \leq 2r, \tag{26}$$

and the second bound is

$$\|\hat{\mathbf{x}}_{t_n-1}^{t_0} - \mathbf{x}_{t_n-1}^\star\|_1 \leq nW\|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1. \tag{27}$$

Lets define $C_{\text{FHC}}^{t_0} \triangleq \sum_{n=0}^{M} C_{\text{FHC}}^{t_0}(n)$. By using the first bound in (26), we can proceed as below

$$\begin{aligned} C_{\text{AFHC}} &\overset{\text{(c)}}{\leq} \frac{1}{W} \sum_{t_0 \in \mathscr{W}} C_{\text{FHC}}^{t_0} \\ &= \frac{1}{W} \sum_{t_0 \in \mathscr{W}} \sum_{n=0}^{M} C_{\text{FHC}}^{t_0}(n) \qquad (28) \\ &\overset{\text{(d)}}{\leq} C_{\text{offline}} + \frac{1}{W} \sum_{t_0 \in \mathscr{W}} \sum_{n=1}^{M} \gamma_n^{t_0} \|\hat{\mathbf{x}}_{t_n-1}^{t_0} - \mathbf{x}_{t_n-1}^\star\|_1 \qquad (29) \\ &\overset{\text{(e)}}{=} C_{\text{offline}} + 2\gamma M r, \qquad (30) \end{aligned}$$

where (c) follows from the convexity of the cost function $c_t(\cdot)$ as well as the remark after (25), (d) is due to the following inequality

$$\sum_{n=0}^{M} \sum_{t=t_n}^{t_n+W-1} c_t(\mathbf{x}_t^\star) \leq \sum_{t=1}^{T} c_t(\mathbf{x}_t^\star) = C_{\text{offline}},$$

and (e) follows by defining $\gamma \triangleq \max_{t_0 \in \mathscr{W}} \max \left\{\bar{\boldsymbol{\mu}}_t^{t_0}, \underline{\boldsymbol{\mu}}_t^{t_0}\right\}_{t=1}^{T}$.

Using the alternative upper bound in (27) and following a similar argument as above gives

$$\begin{aligned} C_{\text{AFHC}} &\leq C_{\text{offline}} + \frac{\gamma}{W} \sum_{t_0 \in \mathscr{W}} \sum_{n=1}^{M} nW\|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1 \\ &= C_{\text{offline}} + \gamma \frac{(M+1)M}{2} W\|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1. \qquad (31) \end{aligned}$$

The final bound is the minimum of (30), (31).

We omit the argument for the case of strong convexity due to space constraints.

B. Proof of Theorem 3

The proof of Theorem 3 follows by constructing a specific sequence of linear cost functions and establishing the difference between AFHC cost and the optimal offline cost over this sequence. In particular, for all $W \geq 2$, we consider a set of linear cost functions with the following structure,

$$c_t(\mathbf{x}_t) = \begin{cases} \sum_{i=1}^{N} \alpha x_t^i & \text{for } t = nW \\ \sum_{i=1}^{N} \beta x_t^i & \text{for } t \in [T],\ t \neq nW, \end{cases} \tag{32}$$

where $n \in [M]$, and $\alpha > 0$ and $\beta < 0$ are two constants such that $|\beta| \ll \alpha$.

Since the cost functions are all linear, we have

$$\begin{aligned} C_{\text{AFHC}} &= \sum_{t=1}^{T} c_t(\hat{\mathbf{x}}_t) = \sum_{t=1}^{T} c_t \left(\frac{1}{W} \sum_{t_0 \in \mathscr{W}} \hat{\mathbf{x}}_t^{t_0}\right) \\ &= \frac{1}{W} \sum_{t_0 \in \mathscr{W}} \sum_{t=1}^{T} c_t(\hat{\mathbf{x}}_t^{t_0}). \end{aligned}$$

Hence,

$$\begin{aligned} C_{\text{AFHC}} - C_{\text{offline}} &= (1/W) \sum_{t_0 \in \mathscr{W}} \sum_{t=1}^{T} \left(c_t(\hat{\mathbf{x}}_t^{t_0}) - c_t(\mathbf{x}_t^\star)\right) \\ &\overset{\text{(a)}}{=} (1/W) \sum_{t_0 \in \mathscr{W}} \sum_{n=1}^{M} \sum_{i=1}^{N} \alpha\left(\hat{x}_{nW}^{i,t_0} - x_{nW}^{i,\star}\right) + \epsilon(\beta), \qquad (33) \end{aligned}$$

where in (a) we define

$$\epsilon(\beta) \triangleq (1/W) \sum_{t_0 \in \mathscr{W}} \sum_{\substack{t=1 \\ t \neq nW}}^{T} \sum_{i=1}^{N} \beta\left(\hat{x}_t^{i,t_0} - x_t^{i,\star}\right).$$

We now argue that from the structure of the given cost functions, an optimal offline solution is strictly decreasing over the time sequence of $t \in \{W, 2W, \cdots, MW\}$. This is due to the fact that α is significantly larger than $|\beta|$. With the given initial value of $\mathbf{x}_0$, we can formulate the solution of optimal offline problem in a closed form as below

$$x_t^{i,\star}|_{t=nW} = x_0^i - nW\underline{X}^i \tag{34}$$

for all $n \in [M]$, where we assumed that the action space is unbounded, i.e., $\mathscr{X} \in \mathbb{R}^N$.

For different values of $t_0 \in \mathscr{W}$, it can also be shown that each version of FHC has the following closed form solution

$$\hat{x}_t^{i,t_0}|_{t=nW} = x_0^i + nW\bar{X}^i - n(1-t_0)(\bar{X}^i + \underline{X}^i), \tag{35}$$

for all $n \in [M]$ and $t_0 \in [-W+1, 0]$[3]. From (34) and (35) we obtain that

$$\hat{x}_t^{i,t_0}|_{t=nW} - x_t^{i,\star}|_{t=nW} = n(W + t_0 - 1)(\bar{X}^i + \underline{X}^i), \tag{36}$$

where $n \in [M]$. Combining (33) and (36) gives

$$C_{\text{AFHC}} - C_{\text{offline}} = \alpha \frac{M(M+1)}{2} \frac{(W-1)}{2} \|\bar{\mathbf{X}} + \underline{\mathbf{X}}\|_1 + \epsilon(\beta).$$

Since for all sufficiently small values of $\beta < 0$ satisfying $|\beta| \ll \alpha$, (34) and (35) are valid, we can take the limit of $\beta \uparrow 0$, which results in $\epsilon(\beta) \to 0$.

The special case of $W = 1$ can be handled separately using the following sequence of the cost functions

$$c_t(\mathbf{x}_t) = \begin{cases} \sum_{i=1}^{N} \alpha x_t^i & \text{if } t \text{ is even} \\ \sum_{i=1}^{N} \beta x_t^i & \text{if } t \text{ is odd,} \end{cases} \tag{37}$$

where α and β satisfy the same conditions as in (32). We omit the details due to space constraints.

[3]Note that $t_0 = 1$ and $t_0 = -W + 1$ result the same FHC optimization problem.