# Online Optimization : Competing with Dynamic Comparators

**Ali Jadbabaie**
University of Pennsylvania

**Alexander Rakhlin**
University of Pennsylvania

**Shahin Shahrampour**
University of Pennsylvania

**Karthik Sridharan**
Cornell University

## Abstract

Recent literature on online learning has focused on developing adaptive algorithms that take advantage of a *regularity* of the sequence of observations, yet retain worst-case performance guarantees. A complementary direction is to develop prediction methods that perform well against complex benchmarks. In this paper, we address these two directions together. We present a fully adaptive method that competes with dynamic benchmarks in which regret guarantee scales with regularity of the sequence of cost functions and comparators. Notably, the regret bound adapts to the smaller complexity measure in the problem environment. Finally, we apply our results to drifting zero-sum, two-player games where both players achieve no regret guarantees against best sequences of actions in hindsight.

## 1 Introduction

The focus of this paper is an online optimization problem in which a *learner* plays against an *adversary* or *nature*. At each round $t \in \{1, \dots, T\}$, the learner chooses an action $x_t$ from some convex feasible set $\mathcal{X} \subseteq \mathbb{R}^d$. Then, nature reveals a convex function $f_t \in \mathcal{F}$ to the learner. As a result, the learner incurs the corresponding *loss* $f_t(x_t)$. A learner aims to minimize his *regret*, a comparison to a single best action in hindsight:

$$\mathbf{Reg}_T^s \triangleq \sum_{t=1}^T f_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^T f_t(x). \tag{1}$$

Let us refer to this as *static* regret in the sense that the comparator is *time-invariant*. In the literature, there are numerous algorithms that guarantee a static regret rate of $\mathcal{O}(\sqrt{T})$ (see e.g. [1–3]). Moreover, when the loss functions are strongly convex, a rate of $\mathcal{O}(\log T)$ could be achieved [4]. Furthermore, minimax optimality of algorithms with respect to the worst-case adversary has been established (see e.g. [5]).

There are two major directions in which the above-mentioned results can be strengthened: (1) by exhibiting algorithms that compete with non-static comparator sequences (that is, making the benchmark harder), and (2) by proving regret guarantees that take advantage of *niceness* of nature's sequence (that is, exploiting some non-adversarial quality of nature's moves). Both of these distinct directions are important avenues of investigation. In the present paper, we attempt to address these two aspects by developing a single, adaptive algorithm with a regret bound that shows the interplay between the difficulty of the comparison sequence and niceness of the sequence of nature's moves.

With respect to the first aspect, a more stringent benchmark is a *time-varying* comparator, a notion that can be termed *dynamic* regret [3, 6–8]:

$$\mathbf{Reg}_T^d \triangleq \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(x_t^*), \tag{2}$$

where $x_t^* \triangleq \operatorname{argmin}_{x \in \mathcal{X}} f_t(x)$. More generally, dynamic regret against a comparator sequence $\{u_t\}_{t=1}^T$ is

$$\mathbf{Reg}_T^d(u_1, \dots, u_T) \triangleq \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(u_t).$$

It is well-known that in the worst case, obtaining a bound on dynamic regret is not possible. However, it is possible to achieve worst-case bounds in terms of

$$C_T(u_1, \dots, u_T) \triangleq \sum_{t=1}^T \left\| u_t - u_{t-1} \right\|, \tag{3}$$

i.e., the *regularity* of the comparator sequence, interpolating between the static and dynamic regret notions. Furthermore, the authors in [9] introduce an algorithm which proposes a variant of $C_T$ involving a dynamical model.

In terms of the second direction, there are several ways of incorporating potential regularity of nature's sequence. The

Appearing in Proceedings of the $18^{th}$ International Conference on Artificial Intelligence and Statistics (AISTATS) 2015, San Diego, CA, USA. JMLR: W&CP volume 38. 

authors in [10, 11] bring forward the idea of predictable sequences – a generic way to incorporate some external knowledge about the gradients of the loss functions. Let $\{M_t\}_{t=1}^T$ be a *predictable* sequence computable by the learner at the beginning of round $t$. This sequence can then be used by an algorithm in order to achieve regret in terms of

$$D_T \triangleq \sum_{t=1}^{T} \left\|\nabla f_t(x_t) - M_t\right\|_*^2. \tag{4}$$

The framework of predictable sequences captures *variation* and *path-length* type regret bounds (see e.g. [12, 13]). Yet another way in which niceness of the adversarial sequence can be captured is through a notion of *temporal variability* studied in [14]:

$$V_T \triangleq \sum_{t=1}^{T} \sup_{x\in\mathcal{X}} \left|f_t(x) - f_{t-1}(x)\right|. \tag{5}$$

What is interesting—and intuitive—dynamic regret against the optimal sequence $\{x_t^*\}_{t=1}^T$ becomes a feasible objective when $V_T$ is *small*. When only noisy versions of gradients are revealed to the algorithm, Besbes et al. in [14] show that using a restarted *Online Gradient Descent (OGD)* [3] algorithm, one can get a bound of form $T^{2/3}(V_T+1)^{1/3}$ on the expected regret. However, the regret bounds attained in [14] are only valid when an upper bound on $V_T$ is known to the learner before the game begins. For the full information online convex optimization setting, when one receives exact gradients instead of noisy gradients, a bound of order $V_T$ is trivially obtained by simply playing (at each round) the minimum of the previous round.

The three quantities we just introduced — $C_T, D_T, V_T$ — measure distinct aspects of the online optimization problem, and their interplay is an interesting object of study. Our first contribution is to develop a fully adaptive method (without prior knowledge of these quantities) whose dynamic regret is given in terms of these three complexity measures. This is done for the full information online convex optimization setting, and augments the existing regret bounds in the literature which focus on only one of the three notions — $C_T, D_T, V_T$ — (and not all the three together). To establish a sub-linear bound on the dynamic regret, we utilize a variant of the *Optimistic Mirror Descent (OMD)* algorithm [10].

When noiseless gradients are available and we can calculate variations at each round, we not only establish a regret bound in terms of $V_T$ and $T$ (without a priori knowledge of a bound on $V_T$), but also show how the bound can in fact be improved when deviation $D_T$ is $o(T)$. We further also show how the bound can automatically adapt to $C_T$ the length of sequence of comparators. Importantly, this avoids suboptimal bounds derived only in terms of one of the quantities — $C_T, V_T$ — in an environment where the other one is small.

The second contribution of this paper is the technical analysis of the algorithm. The bound on the dynamic regret is derived by applying the *doubling trick* to a non-monotone quantity which results in a non-monotone step size sequence (which has not been investigated to the best of authors' knowledge).

We provide uncoupled strategies for two players playing a sequence of drifting zero sum games. We show how when the two players play the provided strategies, their pay offs converge to the average minimax value of the sequence of games (provided the games drift slowly). In this case, both players simultaneously enjoy no regret guarantees against best sequences of actions in hindsight that vary slowly. This is a generalization of the results by Daskalakis *et al.* [15], and Rakhlin *et al.* [11], both of which are for fixed games played repeatedly.

## 2 Preliminaries and Problem Formulation

### 2.1 Notation

Throughout the paper, we assume that for any action $x \in \mathcal{X} \subset \mathbb{R}^d$ at any time $t$, it holds that

$$|f_t(x)| \leq G. \tag{6}$$

We denote by $\|\cdot\|_*$ the dual norm of $\|\cdot\|$, by $[T]$ the set of natural numbers $\{1,\dots,T\}$, and by $f_{1:t}$ the shorthand of $f_1, ..., f_t$, respectively. Whenever $C_T$ is written without arguments, it will refer to regularity $C_T(x_1^*,\dots,x_T^*)$ of the sequence of minimizers of the loss functions. We point out that our initial statements hold for the regularity of any sequence of comparators. However, for upper bounds involving $\sqrt{C_T}$, one needs to choose a computable quantity to tune the step size, and hence our main results are stated for $C_T(x_1^*,\dots,x_T^*)$.

The quantity $D_T$ is defined with respect to an arbitrary predictable sequence $\{M_t\}_{t=1}^T$, but this dependence is omitted for brevity.

### 2.2 Comparing with existing regret bounds in the dynamic setting

We state and discuss relevant results from the literature on online learning in dynamic environments. For any comparator sequence $\{u_t\}_{t=1}^T$ and the specific minima sequence $\{x_t^*\}_{t=1}^T$ the following results are established in the literature:

| Reference | Regret Notion |
|---|---|
| | Regret Rate |
| [3] | $\sum_{t=1}^{T} f_t(x_t) - f_t(u_t)$ |
| [9] | $\mathcal{O}\left(\sqrt{T}(1 + C_T(u_1, \ldots, u_T))\right)$ |
| [14] | $\sum_{t=1}^{T} \mathbb{E}\left[f_t(x_t)\right] - f_t(x_t^*)$ |
| | $\mathcal{O}\left(T^{2/3}(1 + V_T)^{1/3}\right)$ |
| [11] | $\sum_{t=1}^{T} f_t(x_t) - f_t(u)$ |
| | $\mathcal{O}\left(\sqrt{D_T}\right)$ |
| Our work | $\sum_{t=1}^{T} f_t(x_t) - f_t(x_t^*)$ |
| | $\mathcal{O}\big(\sqrt{D_T + 1} + \min\left\{\sqrt{(D_T+1)C_T}, (D_T+1)^{1/3}T^{1/3}V_T^{1/3}\right\}\big)$ |

where $\tilde{\mathcal{O}}(\cdot)$ hides the $\log T$ factor. Lemma 1 below also yields a rate of $\mathcal{O}\left(\sqrt{D_T + 1}(1 + C_T(u_1, \ldots, u_T))\right)$ for any comparator sequence $\{u_t\}_{t=1}^T$. A detailed explanation of the bounds will be done after Theorem 3.

We remark that the authors in [14] consider a setting in which a *variation budget* (an upper bound on $V_T$) is known to the learner, but he/she only has noisy gradients available. Then, the restarted **OGD** guarantees the mentioned rate for convex functions; the rate is modified to $\sqrt{(V_T + 1)T}$ for strongly convex functions.

For the case of *noiseless* gradients, we first aim to show that our algorithm is adaptive in the sense that the learner needs not know an upper bound on $V_T$ in advance when he/she can calculate variations observed so far. Furthermore, we shall establish that our method recovers the known bounds for stationary settings (as well as cases where $V_T$ does not change gradually along the time horizon)

### 2.3 Comparison of Regularity and Variability

We now show that $V_T$ and $C_T$ are not comparable in general. To this end, we consider the classical problem of prediction with expert advice. In this setting, the learner deals with the linear loss $f_t(x) = \langle f_t, x\rangle$ on the $d$-dimensional probability simplex. Assume that for any $t \geq 1$, we have the vector sequence

$$f_t = \begin{cases} (-\frac{1}{T}, 0, 0, \ldots, 0) & \text{, if} \quad t \text{ even} \\ (0, -\frac{1}{T}, 0, \ldots, 0) & \text{, if} \quad t \text{ odd} \end{cases}.$$

Setting $u_t$, the comparator of round $t$, to be the minimizer of $f_t$, i.e. $u_t = x_t^*$, we have

$$C_T = \sum_{t=1}^{T} \|x_t^* - x_{t-1}^*\|_1 = \Theta(T)$$
$$V_T = \sum_{t=1}^{T} \|f_t - f_{t-1}\|_\infty = \mathcal{O}(1),$$

according to (3) and (5), respectively. We see that $V_T$ is considerably smaller than $C_T$ in this scenario. On the other hand, consider prediction with expert advice with two experts. Let $f_t = (-1/2, 0)$ on even rounds and $f_t = (0, 1/2)$ on odd rounds. Expert 1 remains to be the best throughout the game, and thus $C_T = \mathcal{O}(1)$, while variation $V_T = \Theta(T)$. Therefore, one can see that taking into account only one measure might lead us to suboptimal regret bounds. We show that both measures play a key role in our regret bound. Finally, we note that if $M_t = \nabla f_{t-1}(x_{t-1})$, the notion of $D_T$ can be related to $V_T$ in certain cases, yet we keep the predictable sequence arbitrary and thus as playing a role separate from $V_T$ and $C_T$.

## 3 Main Results

### 3.1 Optimistic Mirror Descent and Relation to Regularity

We now outline the **OMD** algorithm previously proposed in [10]. Let $\mathcal{R}$ be a 1-strongly convex function with respect to a norm $\|\cdot\|$, and $\mathcal{D}_{\mathcal{R}}(\cdot,\cdot)$ represent the Bregman divergence with respect to $\mathcal{R}$. Also, let $\mathcal{H}_t$ be the set containing all available information to the learner at the beginning of time $t$. Then, the learner can compute the vector $M_t : \mathcal{H}_t \to \mathbb{R}^d$, which we call the predictable process. Supposing that the learner has access to the side information $M_t \in \mathbb{R}^d$ from the outset of round $t$, the **OMD** algorithm is characterized via the following interleaved sequence,

$$x_t = \text{argmin}_{x\in\mathcal{X}}\left\{\eta_t\langle x, M_t\rangle + \mathcal{D}_{\mathcal{R}}(x, \hat{x}_{t-1})\right\} \tag{7}$$
$$\hat{x}_t = \text{argmin}_{x\in\mathcal{X}}\left\{\eta_t\langle x, \nabla_t\rangle + \mathcal{D}_{\mathcal{R}}(x, \hat{x}_{t-1})\right\}, \tag{8}$$

where $\nabla_t \triangleq \nabla f_t(x_t)$, and $\eta_t$ is the *step size* that can be chosen adaptively to attain low regret. One could observe that for $M_t = 0$, the **OMD** algorithm amounts to the well-known *Mirror Descent* algorithm [16, 17]. On the other hand, the special case of $M_t = \nabla_{t-1}$ recovers the scheme proposed in [13]. It is shown in [10] that the static regret satisfies

$$\mathbf{Reg}_T^s \leq 4R_{\max}\left(\sqrt{D_T} + 1\right),$$

using the step size

$$\eta_t = R_{\max}\min\left\{\left(\sqrt{D_{t-1}} + \sqrt{D_{t-2}}\right)^{-1}, 1\right\},$$

where $R_{\max}^2 \triangleq \sup_{x,y\in\mathcal{X}} \mathcal{D}_{\mathcal{R}}(x, y)$. The following lemma extends the result to arbitrary sequence of comparators $\{u_t\}_{t=1}^T$. Throughout, we assume that $\|\nabla_0 - M_0\|_*^2 = 1$ by convention.

**Lemma 1.** *Let $\mathcal{X}$ be a convex set in a Banach space $\mathcal{B}$. Let $\mathcal{R} : \mathcal{B} \mapsto \mathbb{R}$ be a 1-strongly convex function on $\mathcal{X}$ with*

*respect to a norm $\|\cdot\|$, and let $\|\cdot\|_*$ denote the dual norm. For any $L > 0$, employing the time-varying step size*

$$\eta_t = \frac{L}{\sqrt{\sum_{s=0}^{t-1} \|\nabla_s - M_s\|_*^2} + \sqrt{\sum_{s=0}^{t-2} \|\nabla_s - M_s\|_*^2}},$$

*and running the Optimistic Mirror Descent algorithm for any comparator sequence $\{u_t\}_{t=1}^T$, yields*

$$\mathbf{Reg}_T^d(u_1, \dots, u_T) \leq 2\sqrt{1 + D_T}L + 2\sqrt{1 + D_T}\frac{\gamma C_T(u_1, \dots, u_T) + 4R_{\max}^2}{L},$$

*so long as $\mathcal{D}_{\mathcal{R}}(x, z) - \mathcal{D}_{\mathcal{R}}(y, z) \leq \gamma\|x - y\|, \forall x, y, z \in \mathcal{X}$.*

Lemma 1 underscores the fact that one can get a tighter bound for regret once the learner advances a sequence of conjectures $\{M_t\}_{t=1}^T$ well-aligned with the gradients. Moreover, if the learner has prior knowledge of $C_T$ (or an upper bound on it), then the regret bound would be $\mathcal{O}\left(\sqrt{(D_T + 1)C_T}\right)$ by tuning $L$.

Note that when the function $\mathcal{R}$ is Lipschitz on $\mathcal{X}$, the Lipschitz condition on the Bregman divergence is automatically satisfied. For the particular case of KL divergence this can be achieved via mixing a uniform distribution to stay away from boundaries (see e.g. section 4.2 of the paper in this regard). In this case, the constant $\gamma$ is of $\mathcal{O}(\log T)$.

### 3.2 The Adaptive Optimistic Mirror Descent Algorithm

The main objective of the paper is to develop the *Adaptive Optimistic Mirror Descent (AOMD)* algorithm. The **AOMD** algorithm incorporates all notions of variation $D_T$, $C_T$ and $V_T$ to derive a comprehensive regret bound. The proposed method builds on the **OMD** algorithm with adaptive step size, combined with a *doubling trick* applied to a threshold growing non-monotonically (see e.g. [1, 10] for application of doubling trick on monotone quantities). The scheme is adaptive in the sense that no prior knowledge of $D_T$, $C_T$ or $V_T$ is necessary.

Observe that the prior knowledge of a variation budget (an upper bound on $V_T$) does not tell us how the changes between cost functions are distributed throughout the game. For instance, the variation can increase gradually along the time horizon, while it can also take place in the form of discrete switches. The learner does not have any information about the variation pattern. Therefore, she must adopt a flexible strategy that achieves low regret in the benign case of finite switches or shocks, while it is simultaneously able to compete with the worst-case of gradual change. Before describing the algorithm, let us first use Lemma 1 to bound the general dynamic regret in terms of $D_T$, $C_T$ and $V_T$.

**Lemma 2.** *Let $\mathcal{X}$ be a convex set in a Banach space $\mathcal{B}$. Let $\mathcal{R} : \mathcal{B} \mapsto \mathbb{R}$ be a 1-strongly convex function on $\mathcal{X}$ with respect to a norm $\|\cdot\|$. Run the Optimistic Mirror Descent algorithm with the step size given in the statement of Lemma 1. Letting the comparator sequence be $\{u_t\}_{t=1}^T$, for any $L > 2R_{\max}$ we have*

$$\mathbf{Reg}_T^d(u_1, \dots, u_T) \leq 4\sqrt{1 + D_T}L + \mathbf{1}\left\{\gamma C_T(u_1, \dots, u_T) > L^2 - 4R_{\max}^2\right\}\frac{4\gamma R_{\max} T V_T}{L^2 - 4R_{\max}^2},$$

*so long as $\mathcal{D}_{\mathcal{R}}(x, z) - \mathcal{D}_{\mathcal{R}}(y, z) \leq \gamma\|x - y\|, \forall x, y, z \in \mathcal{X}$.*

We now describe **AOMD** algorithm shown in table 1, and prove that it automatically adapts to $V_T$, $D_T$ and $C_T$. The algorithm can be cast as a repeated **OMD** using different step sizes. The learner sets the parameter $L = 3R_{\max}$ in Lemma 1, and runs the **OMD** algorithm. Along the process, the learner collects deviation, variation and regularity observed so far, and checks the doubling condition in table 1 after each round. Once the condition is satisfied, the learner doubles $L$, discards the accumulated deviation, variation and regularity, and runs a new **OMD** algorithm. Note importantly that the doubling condition results in a non-monotone sequence of step size during the learning process.

Notice that once we have completed running the algorithm, $N$ is the number of doubling epochs, $\Delta_i$ is the number of instances in epoch $i$, $k_i$ and $k_{i+1} - 1$ are the start and end points of epoch $i$, $\sum_{i=1} \Delta_i = T$ , $\sum_{i=1}^N C_{(i)} = C_T$, $\sum_{i=1}^N D_{(i)} = D_T + N$ and $\sum_{i=1}^N V_{(i)} = V_T$. Also, there is a technical reason for initialization choice of $L$ which shall become clear in the proof of Lemma 2. Theorem 3 shows the bound enjoyed by the proposed **AOMD** algorithm.

**Theorem 3.** *Assume that $\mathcal{D}_{\mathcal{R}}(x, z) - \mathcal{D}_{\mathcal{R}}(y, z) \leq \gamma\|x - y\|, \forall x, y, z \in \mathcal{X}$, and let $C_T = \sum_{t=1}^T \left\|x_t^* - x_{t-1}^*\right\|$. The* **AOMD** *algorithm enjoys the following bound on dynamic regret :*

$$\mathbf{Reg}_T^d \leq \tilde{\mathcal{O}}\left(\sqrt{D_T + 1}\right) + \tilde{\mathcal{O}}\left(\min\left\{\sqrt{(D_T + 1)C_T}, (D_T + 1)^{1/3}T^{1/3}V_T^{1/3}\right\}\right),$$

*where $\tilde{\mathcal{O}}(\cdot)$ hides a $\log T$ factor.*

Based on Theorem 3 we can obtain the following table that summarizes bounds on $\mathbf{Reg}_T^d$ for various cases (disregarding the first term $\tilde{\mathcal{O}}\left(\sqrt{D_T + 1}\right)$ in the bound above):

| Regime | Rate |
|---|---|
| $C_T \leq T^{2/3}(D_T + 1)^{-1/3}V_T^{2/3}$ | $\tilde{\mathcal{O}}\left(\sqrt{C_T(D_T + 1)}\right)$ |
| $V_T \leq D_T + 1$ | $\tilde{\mathcal{O}}\left((D_T + 1)^{2/3}T^{1/3}\right)$ |
| $D_T \leq V_T - 1$ | $\tilde{\mathcal{O}}\left(V_T^{2/3}T^{1/3}\right)$ |
| $D_T = \mathcal{O}(T)$ | $\tilde{\mathcal{O}}\left(T^{2/3}V_T^{1/3}\right)$ |

The following remarks are in order :

**Algorithm 1** Adaptive Optimistic Mirror Descent Algorithm

Parameter : $R_{\max}$, some arbitrary $x_0 \in \mathcal{X}$
Initialize $N = 1$, $C_{(1)} = V_{(1)} = 0$, $D_{(1)} = 1$, $x_1 = x_0$, $L_1 = 3R_{\max}$, $\Delta_1 = 0$ and $k_1 = 1$.
**for** $t = 1$ to $T$ **do**
  % check doubling condition
  **if** $L_N^2 < \gamma \min\left\{C_{(N)}\,,\, V_{(N)}^{2/3}\Delta_N^{2/3}D_{(N)}^{-1/3}\right\} + 4R_{\max}^2$ **then**
    % increment $N$ and double $L_N$
    $N = N + 1$
    $L_N = 3R_{\max}2^{N-1}$, $C_{(N)} = V_{(N)} = 0$, $D_{(N)} = 1$ and $\Delta_N = 0$
    $k_N = t$
  **end if**
  Play $x_t$ and suffer loss $f_t(x_t)$
  Calculate $M_{t+1}$ (predictable sequence) and gradient $\nabla_t = \nabla f_t(x_t)$
  % update $D_{(N)}, C_{(N)}, V_{(N)}$ and $\Delta_N$
  $D_{(N)} = D_{(N)} + \|\nabla_t - M_t\|_*^2$
  $C_{(N)} = C_{(N)} + \left\|x_t^* - x_{t-1}^*\right\|$
  $V_{(N)} = V_{(N)} + \sup_{x\in\mathcal{X}} |f_t(x) - f_{t-1}(x)|$
  $\Delta_N = \Delta_N + 1$
  % set step-size and perform optimistic mirror descent update

$$\eta_{t+1} = L_N\left(\sqrt{D_{(N)}} + \sqrt{D_{(N)} - \|\nabla_t - M_t\|_*^2}\right)^{-1}$$

$$\hat{x}_t = \operatorname*{argmin}_{x\in\mathcal{X}}\left\{\eta_t\langle x, \nabla_t\rangle + \mathcal{D}_{\mathcal{R}}(x, \hat{x}_{t-1})\right\}$$

$$x_{t+1} = \operatorname*{argmin}_{x\in\mathcal{X}}\left\{\eta_{t+1}\langle x, M_{t+1}\rangle + \mathcal{D}_{\mathcal{R}}(x, \hat{x}_t)\right\}$$

**end for**

- In all cases, given the condition $V_T = o(T)$, the regret is sub-linear. When the gradients are bounded, the regime $D_T = \mathcal{O}(T)$ always holds, guaranteeing the worst-case bound of $\tilde{\mathcal{O}}\big(T^{2/3}V_T^{1/3}\big)$.

- Theorem 3 allows us to recover $\tilde{\mathcal{O}}(1)$ regret for certain cases where $V_T = \mathcal{O}(1)$. Let nature divide the horizon into $B$ batches, and play a smooth convex function $f_i(x)$ on each batch $i \in [B]$, that is for some $H_i > 0$ it holds that

$$\|\nabla f_i(x) - \nabla f_i(y)\|_* \le H_i\|x - y\|, \qquad (9)$$

  $\forall i \in [B]$ and $\forall x, y \in \mathcal{X}$. Set $M_t = \nabla f_i(\hat{x}_{t-1})$ and note that the gradients are Lipschitz continuous. In this case, the **OMD** corresponding to each batch can be recognized as the *Mirror Prox* method [18], which results in $\tilde{\mathcal{O}}(1)$ regret during each period. Also, since $C_T = \mathcal{O}(1)$ the bound in Theorem 3 is of $\mathcal{O}(\log T)$.

# 4 Applications

## 4.1 Competing with Strategies

So far, we mainly considered dynamic regret $\mathbf{Reg}_T^d$ defined in Equation 2. However, in many scenarios one might want to consider regret against a more specific set of strategies, defined as follows :

$$\mathbf{Reg}_T^{\Pi} \triangleq \sum_{t=1}^{T} f_t(x_t) - \inf_{\pi\in\Pi}\sum_{t=1}^{T} f_t(\pi_t(f_{1:t-1})),$$

where each $\pi \in \Pi$ is a sequence of mappings $\pi = (\pi_1, \ldots, \pi_T)$ and $\pi_t : \mathcal{F}^{t-1} \to \mathcal{X}$. Notice that if $\Pi$ is the set of all mappings then $\mathbf{Reg}_T^{\Pi}$ corresponds to dynamic regret $\mathbf{Reg}_T^d$ and if $\Pi$ corresponds to set of constant history independent mappings, that is, each $\pi \in \Pi$ is indexed by some $x \in \mathcal{X}$ and $\pi_1^x(\cdot) = \ldots = \pi_T^x(\cdot) = x$, then $\mathbf{Reg}_T^{\Pi}$ corresponds to the static regret $\mathbf{Reg}_T^s$. We now define

$$C_T^{\Pi} = \sum_{t=1}^{T}\left\|\pi_t^*(f_{1:t-1}) - \pi_{t-1}^*(f_{1:t-2})\right\|,$$

where $\pi_t^* = \operatorname{arginf}_{\pi\in\Pi}\sum_{s=1}^{t} f_s(\pi_s(f_{1:s-1}))$. Assume that there exists sequence of mappings $\tilde{C}_1, \ldots, \tilde{C}_T$ where $\tilde{C}_t$ maps any $f_1, \ldots, f_t$ to reals and is such that for any $t$ and any $f_1, \ldots, f_{t-1}$,

$$\tilde{C}_{t-1}(f_{1:t-1}) \le \tilde{C}_t(f_{1:t}),$$

and further, for any $T$ and any $f_1, \ldots, f_T$,

$$\sum_{t=1}^{T}\left\|\pi_t^*(f_{1:t-1}) - \pi_{t-1}^*(f_{1:t-2})\right\| \le \tilde{C}_T(f_{1:T}).$$

In this case a simple modification of **AOMD** algorithm where $C_{(N)}$'s are replaced by $\tilde{C}_{\Delta_N}(f_{k_N:k_{N+1}-1})$ leads to the following corollary of Theorem 3.

**Corollary 4.** *Assume that $\mathcal{D}_{\mathcal{R}}(x, z) - \mathcal{D}_{\mathcal{R}}(y, z) \le \gamma\|x - y\|, \forall x, y, z \in \mathcal{X}$. The* **AOMD** *algorithm with the modification mentioned above achieves the following bound on regret*

$$\mathbf{Reg}_T^{\Pi} \le \tilde{\mathcal{O}}\left(\sqrt{D_T + 1}\right) + \tilde{\mathcal{O}}\left(\min\left\{\sqrt{(D_T+1)\tilde{C}_T(f_{1:T})}, (D_T+1)^{1/3}T^{1/3}V_T^{1/3}\right\}\right).$$

The corollary naturally interpolates between the static and dynamic regret. In other words, letting $\tilde{C}_T(f_{1:T}) = 0$ (which holds for constant mappings), we recover the result of [11] (up to logarithmic factors), whereas $\tilde{C}_T(f_{1:T}) = C_T$ simply recovers the regret bound in Theorem 3 corresponding to dynamic regret. The extra log factor is the cost of adaptivity of the algorithm as we assume no prior knowledge about the environment.

### 4.2 Switching Zero-sum Games with Uncoupled Dynamics

Consider two players playing $T$ zero sum games defined by matrices $A_t \in [-1,1]^{m\times n}$ for each $t \in [T]$. We would like to provide strategies for the two players such that, if both players honestly follow the prescribed strategies, the average payoffs of the players approach the average minimax value for the sequence of games at some fast rate. Furthermore, we would also like to guarantee that if one of the players (say the second) deviates from the prescribed strategy, then the first player still has small regret against sequence of actions that do not change drastically. To this end, one can use a simple modification of the **AOMD** algorithm for both players that uses KL divergence as $\mathcal{D}_{\mathcal{R}}$, and mixes in a bit of uniform distribution on each round, producing an algorithm similar to the one in [11] for unchanging uncoupled dynamic games. The following theorem provides bounds for when both players follow the strategy and bound on regret for player I when player II deviates from the strategy.

On round $t$, Player I performs

Play $x_t$ and observe $f_t^\top A_t$

Update

$\hat{x}_t(i) \propto \hat{x}'_{t-1}(i)\exp\{-\eta_t[f_t^\top A_t]_i\}$

$\hat{x}'_t = (1-\beta)\,\hat{x}_t + (\beta/n)\,\mathbf{1}_n$

$x_{t+1}(i) \propto \hat{x}'_t(i)\exp\{-\eta_{t+1}[f_t^\top A_t]_i\}$

and simultaneously Player II performs

Play $f_t$ and observe $A_t x_t$

Update

$\hat{f}_t(i) \propto \hat{f}'_{t-1}(i)\exp\{-\eta'_t[A_t x_t]_i\}$

$\hat{f}'_t = (1-\beta)\,\hat{f}_t + (\beta/m)\,\mathbf{1}_m$

$f_{t+1}(i) \propto \hat{f}'_t(i)\exp\{-\eta'_{t+1}[A_t x_t]_i\}$

Note that in the description of the algorithm as well as the following proposition and its proof, any letter with the prime symbol refers to Player II, and it is used to differentiate the letter from its counterpart for player I.

**Proposition 5.** *Let* $\mathscr{F}_t \triangleq \sum_{i=1}^t \left\|f_i^\top A_i - f_{i-1}^\top A_{i-1}\right\|_\infty^2$, *and set*

$$\eta_t = \min\left\{\log(T^2n)\frac{L}{\sqrt{\mathscr{F}_{t-1}}+\sqrt{\mathscr{F}_{t-2}}}, \frac{1}{32L}\right\}.$$

*Also define* $\mathscr{A}_t \triangleq \sum_{i=1}^t \|A_i x_i - A_{i-1}x_{i-1}\|_\infty^2$, *and let*

$$\eta'_t = \min\left\{\log(T^2m)\frac{L}{\sqrt{\mathscr{A}_{t-1}}+\sqrt{\mathscr{A}_{t-2}}}, \frac{1}{32L}\right\}.$$

*Let* $\beta = 1/T^2$, $M_t = f_{t-1}^\top A_{t-1}$, *and* $M'_t = A_{t-1}x_{t-1}$. *When Player I uses the prescribed strategy, irrespective of the actions of player II, the regret of Player I w.r.t. any sequence of actions* $u_1,\dots,u_T$ *is bounded as :*

$$\sum_{t=1}^T\left(f_t^\top A_t x_t - f_t^\top A_t u_t\right) \le \log(T^2n)\frac{L}{2}\sqrt{\mathscr{F}_T} + 2\log(T^2n)\left(C_T(u_1,\dots,u_T)+2\right)\left(32L + \frac{2\sqrt{\mathscr{F}_T}}{\log(T^2n)L}\right).$$

*Further if both players follow the prescribed strategies then, as long as*

$$2L^2 > \max\{C_T, C'_T\} + 3, \tag{10}$$

*we get,*

$$\sum_{t=1}^T \sup_{f_t\in\Delta_m} f_t^\top A_t x_t \le \sum_{t=1}^T \inf_{x_t\in\Delta_n}\sup_{f_t\in\Delta_m} f_t^\top A_t x_t + \frac{256L}{T} + \frac{1}{2L} + 4\sum_{t=1}^T\|A_{t-1}-A_t\|_\infty + 32L\left(\log(T^2n)C_T + \log(T^2m)C'_T + 2\log(T^4nm)\right) + \left(C_T + C'_T + 4\right)\frac{20 + 4\sqrt{\sum_{t=1}^T\|A_{t-1}-A_t\|_\infty^2}}{L}.$$

A simple consequence of the above proposition is that if for instance the game matrix $A_t$ changes at most $K$ times over the $T$ rounds, and we knew this fact a priori, then by letting $L = \frac{1}{\sqrt{\log(T^2n)}}$, we get that regret for Player I w.r.t. any sequence of actions that switches at most $K$ times even when Player II deviates from the prescribed strategy is $\mathcal{O}\left((K+2)\sqrt{\log(T^2n)T}\right)$.

At the same time if both players follow the strategy, then average payoffs of the players converge to the average minimax equilibrium at the rate of $\mathcal{O}\left(L\,(K+2)\log(T^4nm)\right)$ under the condition on $L$ given in (10). This shows that if the game matrix only changes/switches a constant number of times, then players get $\sqrt{\log(T)T}$ regret bound against arbitrary sequences and comparator actions that switch at most $K$ times while simultaneously get a convergence rate of $\mathcal{O}(\log(T))$ to average equilibrium when both players are honest.

Note that when we let $K = 0$ and set $L$ to some constant, the proposition recovers the rate in static setting [11] where the matrix sequence is time-invariant.

## 5 Conclusion

In this paper, we proposed an online learning algorithm for dynamic environments. We considered time-varying comparators to measure the dynamic regret of the algorithm. Our proposed method is fully adaptive in the sense that the learner needs no prior knowledge of the environment. We

derive a comprehensive upper bound on the dynamic regret capturing the interplay of regularity in the function sequence versus the comparator sequence. Interestingly, the regret bound adapts to the smaller quantity among the two, and selects the best of both worlds. As an instance of dynamic regret, we considered drifting zero-sum, two-player games, and characterized the convergence rate to the average minimax equilibrium in terms of variability in the sequence of payoff matrices.

## Acknowledgements

We gratefully acknowledge the support of ONR BRC Program on Decentralized, Online Optimization, NSF under grants CAREER DMS-0954737 and CCF-1116928, as well as Dean's Research Fund.

## Appendix

***Proof of Theorem 3***. For the sake of clarity in presentation, we stick to the following notation for the proof

$$\begin{aligned}
\underline{D}_{(i)} &\triangleq D_{(i)} - \|\nabla_{k_{i+1}-1} - M_{k_{i+1}-1}\|_*^2 \\
\underline{C}_{(i)} &\triangleq C_{(i)} - \|x^*_{k_{i+1}-1} - x^*_{k_{i+1}-2}\| \\
\underline{V}_{(i)} &\triangleq V_{(i)} - \sup_{x\in\mathcal{X}} \left|f_{k_{i+1}-1}(x) - f_{k_{i+1}-2}(x)\right| \\
\underline{\Delta}_{(i)} &\triangleq \Delta_i - 1,
\end{aligned}$$

for any doubling epoch $i = 1, ..., N$, where we recall that $k_{i+1} - 1$ is the last instance of epoch $i$. Therefore, any symbol with lower bar refers to its corresponding quantity removing only the value of the last instance of that interval.

Let the **AOMD** algorithm run with the step size given by Lemma 1 in the following form

$$\eta_t = \frac{L_i}{\sqrt{\sum_{s=0}^{t-1}\|\nabla_s - M_s\|_*^2} + \sqrt{\sum_{s=0}^{t-2}\|\nabla_s - M_s\|_*^2}},$$

and let $L_i$ be tuned with a doubling condition explained in the algorithm. Once the condition stated in the algorithm fails, the following pair of identities must hold

$$\begin{aligned}
&\gamma \min\{\underline{C}_{(i)}\,,\, \underline{\Delta}_i^{2/3}\underline{V}_{(i)}^{2/3}\underline{D}_{(i)}^{-1/3}\} + 4R_{\max}^2 \le L_i^2 \\
&\gamma \min\{C_{(i)}\,,\, \Delta_i^{2/3}V_{(i)}^{2/3}D_{(i)}^{-1/3}\} + 4R_{\max}^2 > L_i^2. \quad (11)
\end{aligned}$$

Observe that the algorithm doubles $L_i$ only after the condition fails, so at violation points we suffer at most $2G$ by boundedness (6). Then, under purview of Lemma 2, it holds that

$$\begin{aligned}
\mathbf{Reg}_T^d &\le 2NG + \sum_{i=1}^N 4\sqrt{\underline{D}_{(i)}}L_i \\
&\quad + \sum_{i=1}^N \mathbf{1}\left\{\gamma\underline{C}_{(i)} > L_i^2 - 4R_{\max}^2\right\}\frac{4\gamma R_{\max}\underline{\Delta}_i\underline{V}_{(i)}}{L_i^2 - 4R_{\max}^2} \\
&\le 2NG + \sum_{i=1}^N 4\sqrt{D_{(i)}}L_i \\
&\quad + \sum_{i=1}^N \mathbf{1}\left\{\underline{C}_{(i)} > \underline{\Delta}_i^{2/3}\underline{V}_{(i)}^{2/3}\underline{D}_{(i)}^{-1/3}\right\}\frac{4\gamma R_{\max}\underline{\Delta}_i\underline{V}_{(i)}}{L_i^2 - 4R_{\max}^2}, \quad (12)
\end{aligned}$$

where the last step follows directly from (11) and the fact that $\underline{D}_{(i)} \le D_{(i)}$.

Bounding $\sqrt{D_{(i)}}L_i$ in above, using the second inequality in (11), we get

$$\begin{aligned}
\sqrt{D_{(i)}}L_i &\le \sqrt{\gamma\min\left\{D_{(i)}C_{(i)}\,,\, \Delta_i^{2/3}V_{(i)}^{2/3}D_{(i)}^{2/3}\right\} + 4R_{\max}^2 D_{(i)}} \\
&\le 2R_{\max}\sqrt{D_{(i)}} \\
&\quad + \sqrt{\gamma}\min\left\{\sqrt{D_{(i)}C_{(i)}}\,,\, \Delta_i^{1/3}V_{(i)}^{1/3}D_{(i)}^{1/3}\right\},
\end{aligned}$$

by the simple inequality

$$\sqrt{a+b} \le \sqrt{a} + \sqrt{b}.$$

Plugging the bound above into (12) and noting that

$$\begin{aligned}
\sum_{i=1}^N \sqrt{D_{(i)}} &= N\sum_{i=1}^N \frac{1}{N}\sqrt{D_{(i)}} \\
&\le N\sqrt{\frac{1}{N}\sum_{i=1}^N D_{(i)}} = \sqrt{ND_T + N},
\end{aligned}$$

by Jensen's inequality, we obtain

$$\begin{aligned}
\mathbf{Reg}_T^d &\le 2NG + 8R_{\max}\sqrt{ND_T + N} \\
&\quad + 4\sqrt{\gamma}\sum_{i=1}^N \min\left\{\sqrt{D_{(i)}C_{(i)}}\,,\, D_{(i)}^{1/3}\Delta_i^{1/3}V_{(i)}^{1/3}\right\} \\
&\quad + \sum_{i=1}^N \frac{\mathbf{1}\left\{\underline{C}_{(i)} > \underline{\Delta}_i^{2/3}\underline{V}_{(i)}^{2/3}\underline{D}_{(i)}^{-1/3}\right\}4R_{\max}\underline{\Delta}_i\underline{V}_{(i)}}{\min\left\{\underline{C}_{(i)}, \underline{\Delta}_i^{2/3}\underline{V}_{(i)}^{2/3}\underline{D}_{(i)}^{-1/3}\right\}},
\end{aligned}$$

where we used the first inequality in (11) to bound the last term. Given the condition in the indicator function $\mathbf{1}\{\cdot\}$,

we can simplify above to derive,

$$\begin{aligned}
\mathbf{Reg}_T^d &\leq 2NG + 8R_{\max}\sqrt{ND_T + N} \\
&+ 4\sqrt{\gamma}\sum_{i=1}^{N} \min\left\{\sqrt{D_{(i)}C_{(i)}}\,,\; D_{(i)}^{1/3}\Delta_i^{1/3}V_{(i)}^{1/3}\right\} \\
&+ 4R_{\max}\sum_{i=1}^{N} \mathbf{1}\left\{\underline{C}_{(i)} > \underline{\Delta}_i^{2/3}\underline{V}_{(i)}^{2/3}\underline{D}_{(i)}^{-1/3}\right\}\underline{D}_{(i)}^{1/3}\underline{V}_{(i)}^{1/3}\underline{\Delta}_i^{1/3} \\
&= 2NG + 8R_{\max}\sqrt{ND_T + N} \\
&+ 4\sqrt{\gamma}\sum_{i=1}^{N} \min\left\{\sqrt{D_{(i)}C_{(i)}}\,,\; D_{(i)}^{1/3}\Delta_i^{1/3}V_{(i)}^{1/3}\right\} \\
&+ 4R_{\max}\sum_{i=1}^{N} \mathbf{1}\left\{\sqrt{\underline{D}_{(i)}\underline{C}_{(i)}} > \underline{\Delta}_i^{1/3}\underline{V}_{(i)}^{1/3}\underline{D}_{(i)}^{1/3}\right\}\underline{D}_{(i)}^{1/3}\underline{V}_{(i)}^{1/3}\underline{\Delta}_i^{1/3} \\
&\leq 2NG + 8R_{\max}\sqrt{ND_T + N} \\
&+ 4\sqrt{\gamma}\sum_{i=1}^{N} \min\left\{\sqrt{D_{(i)}C_{(i)}}\,,\; D_{(i)}^{1/3}\Delta_i^{1/3}V_{(i)}^{1/3}\right\} \\
&+ 4R_{\max}\sum_{i=1}^{N} \min\left\{\sqrt{\underline{D}_{(i)}\underline{C}_{(i)}}, \underline{D}_{(i)}^{1/3}\underline{V}_{(i)}^{1/3}\underline{\Delta}_i^{1/3}\right\}. \qquad (13)
\end{aligned}$$

Given the fact that removing the last instance along the interval only reduces variation, we get

$$\begin{aligned}
\underline{C}_{(i)} &\leq C_{(i)} \qquad & \underline{D}_{(i)} &\leq D_{(i)} \\
\underline{V}_{(i)} &\leq V_{(i)} \qquad & \underline{\Delta}_i &\leq \Delta_i,
\end{aligned}$$

and return to (13) to derive

$$\begin{aligned}
\mathbf{Reg}_T^d &\leq 2NG + 8R_{\max}\sqrt{ND_T + N} \\
&+ (4\sqrt{\gamma} + 4R_{\max})\sum_{i=1}^{N} \min\left\{\sqrt{D_{(i)}C_{(i)}}\,,\; D_{(i)}^{1/3}\Delta_i^{1/3}V_{(i)}^{1/3}\right\} \\
&\leq 2NG + 8R_{\max}\sqrt{ND_T + N} \\
&+ (4\sqrt{\gamma} + 4R_{\max})\min\left\{\sum_{i=1}^{N}\sqrt{D_{(i)}C_{(i)}}, \sum_{i=1}^{N} D_{(i)}^{1/3}\Delta_i^{1/3}V_{(i)}^{1/3}\right\} \\
&\leq 2N\left(G + 4R_{\max}\sqrt{D_T + 1}\right) \\
&+ 4N(\sqrt{\gamma} + R_{\max})\min\left\{\sqrt{(D_T+1)C_T}, (D_T+1)^{1/3}T^{1/3}V_T^{1/3}\right\}. \qquad (14)
\end{aligned}$$

where we bounded the sums using the following fact about the summands

$$\begin{aligned}
C_{(i)} &\leq C_T \qquad & D_{(i)} &\leq D_T + 1 \\
V_{(i)} &\leq V_T \qquad & \Delta_i &\leq T.
\end{aligned}$$

To bound the number of batches $N$, we recall from the description of the **AOMD** algorithm that

$$L_i = 3R_{\max}2^{i-1},$$

and use the second inequality in (11) to bound $L_{N-1}$ as follows

$$\begin{aligned}
N &= 2 + \log_2(2^{N-2}) \\
&= 2 + \log_2(L_{N-1}) - \log_2(3R_{\max}) \\
&\leq 2 + \frac{1}{2}\log_2\left(\gamma C_{(N-1)} + 4R_{\max}^2\right) - \log_2(3R_{\max}) \\
&\leq 2 + \frac{1}{2}\log_2\left(2\gamma R_{\max}T + 4R_{\max}^2\right) - \log_2(3R_{\max}).
\end{aligned}$$

In view of the preceding relation and (14), we have

$$\begin{aligned}
\mathbf{Reg}_T^d &\leq \kappa\big(G + 4R_{\max}\sqrt{D_T + 1}\big) \\
&+ 2\kappa\sqrt{\gamma}\min\left\{\sqrt{(D_T+1)C_T}, (D_T+1)^{1/3}T^{1/3}V_T^{1/3}\right\} \\
&+ 2\kappa R_{\max}\min\left\{\sqrt{(D_T+1)C_T}, (D_T+1)^{1/3}T^{1/3}V_T^{1/3}\right\},
\end{aligned}$$

where

$$\kappa \triangleq 4 + \log_2\left(2\gamma R_{\max}T + 4R_{\max}^2\right) - 2\log_2(3R_{\max}),$$

and thereby completing the proof. ■

## References


[1] N. Cesa-Bianchi, G. Lugosi *et al.*, *Prediction, learning, and games*. Cambridge University Press Cambridge, 2006, vol. 1, no. 1.1.

[2] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of computer and system sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[3] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *International Conference on Machine Learning*, 2003.

[4] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Machine Learning*, vol. 69, no. 2-3, pp. 169–192, 2007.

[5] J. Abernethy, P. L. Bartlett, A. Rakhlin, and A. Tewari, "Optimal strategies and minimax lower bounds for online convex games," in *Proceedings of the Nineteenth Annual Conference on Computational Learning Theory*, 2008.

[6] O. Bousquet and M. K. Warmuth, "Tracking a small set of experts by mixing past posteriors," *The Journal of Machine Learning Research*, vol. 3, pp. 363–396, 2003.

[7] N. Cesa-Bianchi, P. Gaillard, G. Lugosi, and G. Stoltz, "A new look at shifting regret," *CoRR abs/1202.3323*, 2012.

[8] N. Buchbinder, S. Chen, J. Naor, and O. Shamir, "Unified algorithms for online learning and competitive analysis." *Journal of Machine Learning Research-Proceedings Track*, vol. 23, pp. 5–1, 2012.

[9] E. C. Hall and R. M. Willett, "Online optimization in dynamic environments," *arXiv preprint arXiv:1307.5944*, 2013.

[10] A. Rakhlin and K. Sridharan, "Online learning with predictable sequences," in *Conference on Learning Theory*, 2013, pp. 993–1019.

[11] ——, "Optimization, learning, and games with predictable sequences," in *Advances in Neural Information Processing Systems*, 2013, pp. 3066–3074.

[12] E. Hazan and S. Kale, "Extracting certainty from uncertainty: Regret bounded by variation in costs," *Machine learning*, vol. 80, no. 2-3, pp. 165–188, 2010.

[13] C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu, "Online optimization with gradual variations," in *Conference on Learning Theory*, 2012.

[14] O. Besbes, Y. Gur, and A. Zeevi, "Non-stationary stochastic optimization," *arXiv preprint arXiv:1307.5449*, 2013.

[15] C. Daskalakis, A. Deckelbaum, and A. Kim, "Near-optimal no-regret algorithms for zero-sum games," *Games and Economic Behavior*, 2014.

[16] A. S. Nemirovski and D. B. Yudin, *Problem complexity and method efficiency in optimization*. Wiley (Chichester and New York), 1983.

[17] A. Beck and M. Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization," *Operations Research Letters*, vol. 31, no. 3, pp. 167–175, 2003.

[18] A. Nemirovski, "Prox-method with rate of convergence o (1/t) for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 229–251, 2004.