

# Cautious Regret Minimization: Online Optimization with Long-Term Budget Constraints

Nikolaos Liakopoulos<sup>1,2,3</sup>, Apostolos Destounis<sup>1</sup>, Georgios S. Paschos<sup>1</sup>,  
Thrasyvoulos Spyropoulos<sup>2</sup>, Panayotis Mertikopoulos<sup>4</sup>

<sup>1</sup> France Research Center, Huawei Technologies Co. Ltd. , <sup>2</sup> Eurecom, Mobile Communications Department

<sup>3</sup>UPMC, Sorbonne Universities, <sup>4</sup>University of Grenoble Alpes, CNRS, Inria, LIG



# Motivating Example: Online Advertisement

# Motivating Example: Online Advertisement

- Select over  $d$  websites to place bids for advertisements  $\mathbf{x}_t = \{x_t^1, \dots, x_t^d\}$ , over  $T$  slots up to a budget of  $b_T$
- You select your bid and then pay the a priori unknown  $\mathbf{p}_t^\top \mathbf{x}_t$ 
  - Depends on: competitor fraud, website popularity change, customer behavior
- Long term budget constraint  $\sum_{t=1}^T \mathbf{p}_t^\top \mathbf{x}_t \leq b_t$
- Return profit is revealed after the decision  $\mathbf{r}_t^\top \mathbf{x}_t$

# Motivating Example: Online Advertisement

- Select over  $d$  websites to place bids for advertisements  $\mathbf{x}_t = \{x_t^1, \dots, x_t^d\}$ , over  $T$  slots up to a budget of  $b_T$
- You select your bid and then pay the a priori unknown  $\mathbf{p}_t^\top \mathbf{x}_t$ 
  - Depends on: competitor fraud, website popularity change, customer behavior
- Long term budget constraint  $\sum_{t=1}^T \mathbf{p}_t^\top \mathbf{x}_t \leq b_T$
- Return profit is revealed after the decision  $\mathbf{r}_t^\top \mathbf{x}_t$

**Target Online  
Optimization**

$$\begin{aligned} & \text{maximize} && \sum_{t=1}^T \mathbf{r}_t^\top \mathbf{x}_t \\ & \text{subject to} && \sum_{t=1}^T \mathbf{p}_t^\top \mathbf{x}_t \leq b_T \\ & && \mathbf{x}_t \in \mathcal{X}, \forall t \in \{1, \dots, T\} \end{aligned}$$

# Prior Work

*Cautious Regret Minimization: Online Optimization with Long-Term Budget Constraints, ICML 2019, Long Beach*  
*N. Liakopoulos, A. Destounis, G. Paschos, T. Spyropoulos, P. Mertikopoulos*

# Prior Work

**Online Convex Optimization with  
Long Term Adversarial Constraints**

$$\begin{aligned} &\text{maximize} && \sum_{t=1}^T f_t(\mathbf{x}_t) \\ &\text{subject to} && \sum_{t=1}^T g_t(\mathbf{x}_t) \\ &&& \mathbf{x}_t \in \mathcal{X}, \forall t \in \{1, \dots, T\} \end{aligned}$$

# Prior Work

## Online Convex Optimization with Long Term Adversarial Constraints

$$\begin{aligned} & \text{maximize} && \sum_{t=1}^T f_t(\mathbf{x}_t) \\ & \text{subject to} && \sum_{t=1}^T g_t(\mathbf{x}_t) \\ & && \mathbf{x}_t \in \mathcal{X}, \forall t \in \{1, \dots, T\} \end{aligned}$$

Paper	Type	Bench.	Regret	Residual
Yuan et al. 2018	Fix.	$T$	$\mathcal{O}(\sqrt{T} + \frac{T}{V})$	$\mathcal{O}(\sqrt{VT})$
Yu et al. 2017	Stoc.	$T$	$\mathcal{O}(\sqrt{T})^*$	$\mathcal{O}(\sqrt{T})^*$
Neely & Yu 2017	Adv.	1	$\mathcal{O}(\sqrt{T})^\dagger$	$\mathcal{O}(\sqrt{T})^\dagger$
Sun et al. 2017	Adv.	1	$\mathcal{O}(\sqrt{T})$	$\mathcal{O}(T^{3/4})$
Mannor et al. 2009	Adv.	$T$	$\Omega(T)$	$o(T)$
<b>Us</b>	<b>Adv.</b>	<b><math>K</math></b>	<b><math>\mathcal{O}(\sqrt{T} + \frac{KT}{V})</math></b>	<b><math>\mathcal{O}(\sqrt{VT})</math></b>

“No Regret” against the weakest benchmark

Impossibility against the strongest benchmark

# Cautious Online Lagrangian Descent



# Cautious Online Lagrangian Descent

## K-Benchmark

$$x_*^K \in \operatorname{argmin}_{x \in \mathcal{X}_K} \sum_{t=1}^T f_t(x)$$

$$\mathcal{X}_K \triangleq \left\{ x \in \mathcal{X} \mid \sum_{\tau=t}^{t+K-1} g_\tau(x) \leq 0, \forall t \in \{1, \dots, T-K+1\} \right\}$$

# Cautious Online Lagrangian Descent

## K-Benchmark

$$x_*^K \in \operatorname{argmin}_{x \in \mathcal{X}_K} \sum_{t=1}^T f_t(x)$$

$$\mathcal{X}_K \triangleq \left\{ x \in \mathcal{X} \mid \sum_{\tau=t}^{t+K-1} g_\tau(x) \leq 0, \forall t \in \{1, \dots, T-K+1\} \right\}$$

## COLD

$$\mathbf{x}_t = \Pi_{\mathcal{X}} \left[ \mathbf{x}_{t-1} - \frac{V \mathbf{f}'_{t-1}(\mathbf{x}_{t-1}) + Q(t) \mathbf{g}'_{t-1}(\mathbf{x}_{t-1})}{2\alpha} \right]$$

$$Q(t+1) = [Q(t) + \hat{g}_t(x_t)]^+$$

# Cautious Online Lagrangian Descent

## K-Benchmark

$$x_*^K \in \operatorname{argmin}_{x \in \mathcal{X}_K} \sum_{t=1}^T f_t(x)$$

$$\mathcal{X}_K \triangleq \left\{ x \in \mathcal{X} \mid \sum_{\tau=t}^{t+K-1} g_\tau(x) \leq 0, \forall t \in \{1, \dots, T-K+1\} \right\}$$

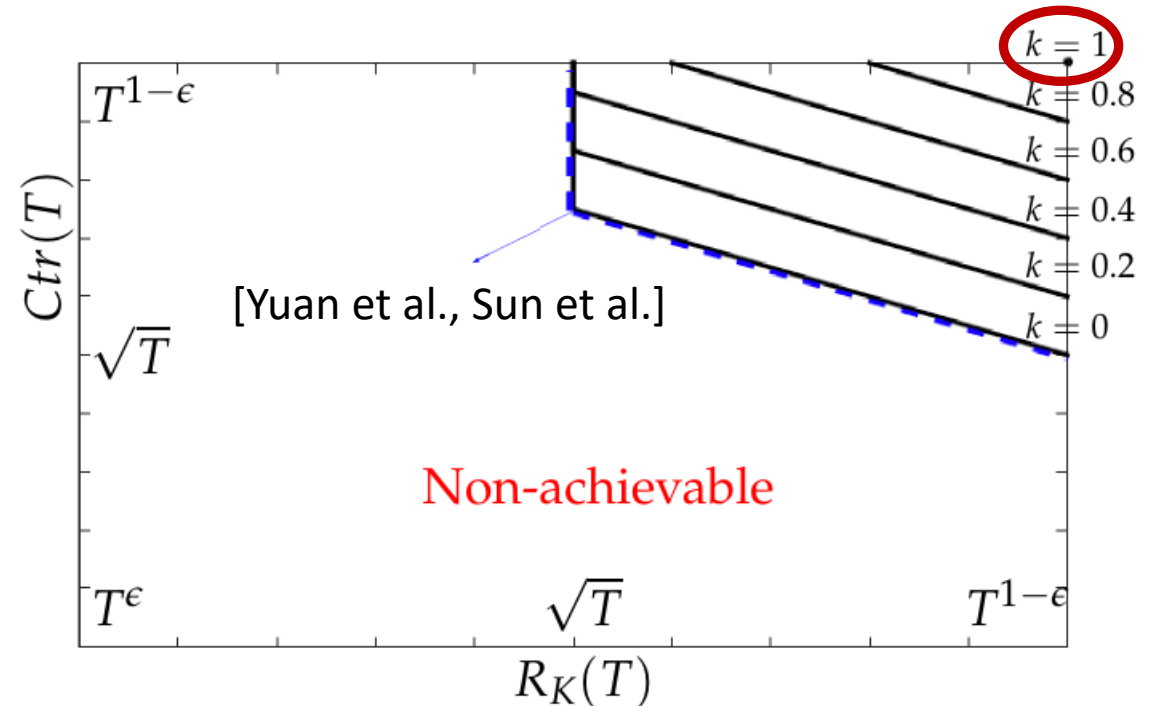
## COLD

$$\mathbf{x}_t = \Pi_{\mathcal{X}} \left[ \mathbf{x}_{t-1} - \frac{V \mathbf{f}'_{t-1}(\mathbf{x}_{t-1}) + Q(t) \mathbf{g}'_{t-1}(\mathbf{x}_{t-1})}{2\alpha} \right]$$

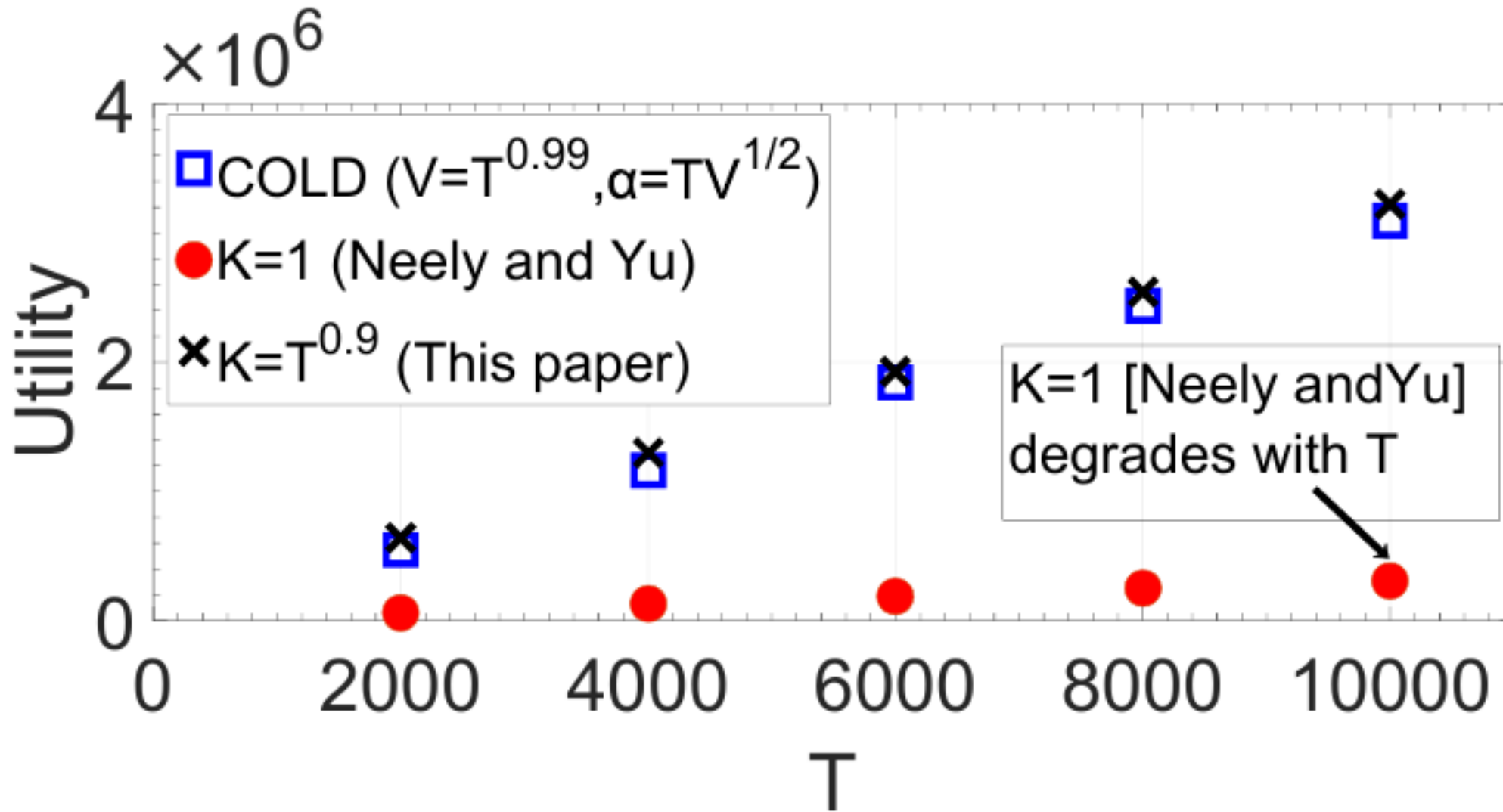
$$Q(t+1) = [Q(t) + \hat{g}_t(x_t)]^+$$

## Achievable Bounds

“No regret” against  $K = \mathcal{O}(T^{1-\epsilon})$



# Impact of Guarantee



# Cautious Regret Minimization: Online Optimization with Long-Term Budget Constraints

Nikolaos Liakopoulos<sup>1,2</sup> Apostolos Destounis<sup>1</sup> Georgios Paschos<sup>1</sup> Thrasyvoulos Spyropoulos<sup>2</sup> Panayiotis Mertikopoulos<sup>3</sup>

<sup>1</sup>Huawei Technologies, Paris Research Center <sup>2</sup>EURECOM, Sophia Antipolis <sup>3</sup>Univ. Grenoble Alpes, CNRS, Inria, LIG

Thanks!

See you at our poster!

**Objective**  
Provide an OCO framework with performance guarantees against *adversarial loss* and *adversarial long-term constraints*.

## Overview

- In this general setting prior work [1] has shown that “no regret” is impossible.
- We define a new benchmark for regret, the  $K$ -Benchmark.
- We prove “no regret” and asymptotic feasibility for any  $K = o(T)$ .

## Setup

We focus on the online problem

$$\begin{aligned} & \text{minimize} && \sum_{t=1}^T f_t(x_t) \\ & \text{subject to} && \sum_{t=1}^T g_t(x_t) \leq 0 \quad (\text{Opt}) \\ & && x_t \in \mathcal{X}, \forall t \in \{1, \dots, T\} \end{aligned}$$

Assumptions:

- (A1) Compact and Convex  $\mathcal{X}$  with diameter  $D$ .
- (A2) Convex  $f_t, g_t$  with  $\|f_t'\|_2 \leq G, \|g_t'\|_2 \leq G$ .
- (A3) For a given  $K \leq T$ ,  $\mathcal{X}_K$  is non-empty.

## K-Benchmark

$$x_*^K \in \operatorname{argmin}_{x \in \mathcal{X}_K} \sum_{t=1}^T f_t(x)$$

$$\mathcal{X}_K \triangleq \left\{ x \in \mathcal{X} \mid \sum_{t=1}^{t+K-1} g_t(x) \leq 0, \forall t \in \{1, \dots, T-K+1\} \right\}$$

- $K = T \rightarrow$  **Impossibility** [1].
- $K = 1 \rightarrow$  **“no regret”** [2].

## Prior Work

Paper	Type	Bench.	Regret	Residual
[3]	Fix.	$T$	$\mathcal{O}(\sqrt{T} + \frac{T}{V})$	$\mathcal{O}(\sqrt{VT})$
[4]	Stoc.	$T$	$\mathcal{O}(\sqrt{T})^*$	$\mathcal{O}(\sqrt{T})^*$
[2]	Adv.	1	$\mathcal{O}(\sqrt{T})^\dagger$	$\mathcal{O}(\sqrt{T})^\dagger$
[5]	Adv.	1	$\mathcal{O}(\sqrt{T})$	$\mathcal{O}(T^{3/4})$
[1]	Adv.	$T$	$\Omega(T)$	$o(T)$
Us	Adv.	$K$	$\mathcal{O}(\sqrt{T} + \frac{KT}{V})$	$\mathcal{O}(\sqrt{VT})$

Table 1: Comparison table of prior work handling long term constraints. Legend: \*Stochastic Slater, †Slater.

## Cautious Online Lagrangian Descent

### Algorithm 1 COLD

- for  $t \leftarrow 1$  to  $T$  do
  - $x_t = \Pi_{\mathcal{X}}[x_{t-1} - \frac{V f_{t-1}'(x_{t-1}) + Q(t) g_{t-1}'(x_{t-1})}{\alpha}]$
  - $Q(t+1) = [Q(t) + g_t(x_t)]^+$
  - end for
- $V$  cautiousness,  $\alpha$  strong convexity.
  - $\Pi$  is the euclidean projection.
  - $Q(t)$  is the predictor queue.
  - $g_t(x_t) = g_{t-1}(x_{t-1}) + g_{t-1}'(x_{t-1})(x_t - x_{t-1})$ .

## Main take-away

Using our  $K$ -benchmark idea, we can characterize the regret performance of online Lagrangian descent when faced with adversarial constraints (a case previously left open).

## Experiments

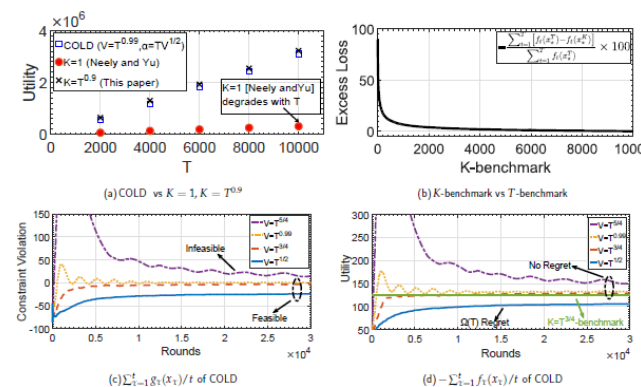


Figure 1: Online ad placement. (a) Utility comparison 1-benchmark vs  $T^{0.9}$ -benchmark. (b) Relative excess loss of K-benchmark to T-benchmark. (c) Constraint residual and (d) Utility performance of COLD for different values of  $V$ .

## Performance Bounds

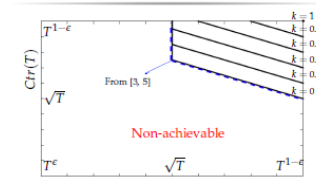


Figure 2: Achievable bounds for  $K = T^k, k = 0 : 1 : 0.2$ .

## Achievable Tradeoffs

Cases:

- $K = o(T), V \in (K, T), \alpha = \max\{T, V\sqrt{T}\}$   
Regret  $\mathcal{O}(KT/V + \sqrt{T})$ , Residual  $\mathcal{O}(\sqrt{VT})$
- $K = T^{1-\epsilon}, V = T^{1-\frac{\epsilon}{2}}, \alpha = V\sqrt{T}$   
Regret  $\mathcal{O}(T^{1-\frac{\epsilon}{2}})$ , Residual  $\mathcal{O}(T^{1-\frac{\epsilon}{2}})$

## References

- S. Mannor et al., “Online Learning with Sample Path Constraints,” *Journal of Machine Learning Research*, 2009.
- M. J. Neely et al., “Online Convex Optimization with Time-Varying Constraints,” *arXiv preprint arXiv:1702.04783*, 2017.
- J. Yuan et al., “Online Convex Optimization for Cumulative Constraints,” in *NeurIPS*, 2018.
- H. Yu et al., “Online Convex Optimization with Stochastic Constraints,” in *NeurIPS*, 2017.
- W. Sun et al., “Safety-Aware Algorithms for Adversarial Contextual Bandit,” in *ICML*, 2017.

## Contact Information

Email: nikolaos.liakopoulos@huawei.com