Learning-Augmented Online Convex Optimization with Adversarial Constraints

Anonymous authors
Paper under double-blind review

Abstract

We study online convex optimization with adversarial constraints, where both cost and constraint functions are chosen by an adaptive adversary and the learner must simultaneously achieve sublinear regret and sublinear cumulative constraint violation (CCV). For α-strongly convex costs, the sharpest known CCV bound is $O(\sqrt{T})$ in general; the $O(\log T)$ rate requires a non-negative regret assumption that can fail against adaptive adversaries. We propose LA-COCO, a learning-augmented framework that integrates constraint predictions into the Lyapunov-based primal-dual paradigm. By augmenting the surrogate cost with a prediction-dependent penalty, we introduce a non-negative bonus term in the Lyapunov drift that compensates for negative regret, replacing the non-negative regret assumption with a strictly weaker posterior condition—the prediction compensation event $\mathcal{E}$. Under $\mathcal{E}$, the algorithm attains $\mathrm{CCV}_T = O(G^2 \log T/\alpha + \eta_T)$, where η_T is the cumulative prediction error; with perfect predictions this recovers $O(\log T)$ without any regret-sign assumption. A Hedge-based meta-algorithm that dynamically combines the prediction-augmented policy with a prediction-free baseline guarantees an unconditional $O(\sqrt{T})$ fallback regardless of prediction quality, establishing a consistency–robustness tradeoff. For general convex costs, we prove an unconditional $\mathrm{CCV}_T = O(GD\sqrt{T}\log T)$ bound. Experiments across 120 configurations with rotating adversarial constraints confirm that $\mathcal{E}$ holds with probability one under stochastic adversaries and that all empirical CCV values remain below the theoretical bounds.

1 Introduction

Sequential decision-making under uncertainty pervades modern engineering systems. In cloud computing, a scheduler must allocate resources across virtual machines while respecting time-varying capacity limits that shift with demand spikes. In online portfolio selection, an investor rebalances assets each day to minimize cumulative loss subject to risk constraints that depend on realized market volatility. In real-time control of power grids, an operator dispatches generation to track fluctuating renewable output while satisfying safety margins that change with grid topology. These problems share a common mathematical structure: a learner repeatedly chooses an action from a convex set, observes a convex cost function and a convex constraint function selected by an adversary, and seeks to minimize cumulative cost while keeping cumulative constraint violation (CCV) small. Online convex optimization (OCO) (Zinkevich, 2003; Hazan, 2016; Shalev-Shwartz, 2012) provides the canonical framework for the cost-minimization aspect, but the simultaneous control of adversarial constraints introduces challenges that go well beyond standard regret analysis.

The study of constrained OCO has evolved through several stages. Early work by Mannor et al. (2009) and Mahdavi et al. (2012) considered long-term constraints that are fixed across rounds, achieving $O(\sqrt{T})$ CCV through primal-dual methods that maintain a running average of constraint violations. Jenatton et al. (2016) extended this line with adaptive step-size schemes, and Yu et al. (2020) obtained $O(1)$ constraint violation for stochastic long-term constraints by exploiting Slater-type feasibility conditions. A qualitative shift occurred when the community turned to fully adversarial, time-varying constraints, where the constraint function g_t may change arbitrarily at each round. Neely (2017) and Liakopoulos et al. (2019) developed

Table 1: Comparison of CCV bounds in constrained online convex optimization. The top block lists prior results under various constraint models; the bottom block presents our results. Here $\eta_T = \sum_{t=1}^{T} \varepsilon_t$ denotes cumulative prediction error. The prediction compensation event $\mathcal{E}$ is strictly weaker than non-negative regret (see Theorem 3). "SC" = α-strongly convex costs; "Cvx" = general convex costs.

	Constraints	CCV bound	Condition	Predictions
Mahdavi et al. (2012)	Fixed	$O(T^{3/4})$	Cvx	No
Jenatton et al. (2016)	Fixed	$O(\sqrt{T})$	Cvx	No
Yu et al. (2020)	Stochastic	$O(1)$	Cvx + Slater	No
Neely (2017)	Adversarial	$O(\sqrt{T})$	Cvx	No
Yuan & Lamperski (2018)	Adversarial	$O(\sqrt{T})$	Cvx	No
Yi et al. (2023)	Adversarial	$O(\sqrt{T})$	Cvx	No
Sinha & Vaze (2024)	Adversarial	$O\left(\frac{G^2 \log T}{\alpha}\right)$	SC, $\text{Regret}_T \geq 0$	No
Sinha & Vaze (2024)	Adversarial	$O\left(\sqrt{\frac{G^3 D T \log T}{\alpha}}\right)$	SC	No
LA-COCO (this paper)	Adversarial	$O\left(\frac{G^2 \log T}{\alpha} + \eta_T\right)$	SC, $\mathcal{E}$ (weaker)	Yes
LA-COCO (this paper)	Adversarial	$O\left(\sqrt{\frac{G^3 D T \log T}{\alpha}}\right)$	SC	Yes

Lyapunov-based approaches for this harder setting, and Sun et al. (2017) and Yi et al. (2023) studied safety-aware and distributed variants. The state of the art was established by Sinha & Vaze (2024), who proved that under α-strongly convex costs, a Lyapunov-based primal-dual algorithm achieves $\text{CCV}_T = O(\sqrt{T})$ unconditionally, and this can be sharpened to $O(\log T)$ when the algorithm's regret is non-negative. The logarithmic rate represents a dramatic improvement, yet the non-negative regret condition is restrictive: it is not guaranteed against a fully adaptive adversary, and there exist adversarial instances that drive the regret to $-\Omega(T)$. In such cases, the $O(\log T)$ guarantee simply does not apply, and the algorithm falls back to the generic $O(\sqrt{T})$ rate. Meanwhile, the learning-augmented (or "algorithms with predictions") paradigm (Lykouris & Vassilvitskii, 2021; Purohit et al., 2018; Mitzenmacher & Vassilvitskii, 2022; Cao et al., 2024) has demonstrated that machine-learned predictions can improve worst-case online algorithms for caching, scheduling, ski rental, and unconstrained OCO with switching costs (Chen et al., 2018; Li et al., 2021; Christianson et al., 2023). Optimistic methods that exploit predictable sequences have also yielded tighter regret bounds in unconstrained OCO (Rakhlin & Sridharan, 2013; Chiang et al., 2012). However, none of these results address constraint violation, and to the best of our knowledge, the learning-augmented paradigm has not been systematically applied to online convex optimization with adversarial constraints.

In this paper, we introduce LA-COCO (Learning-Augmented Constrained Online Convex Optimization), a framework that incorporates constraint predictions into the Lyapunov-based primal-dual paradigm to close this gap. At each round t, the algorithm receives a prediction $\hat{g}_{t+1}$ of the next constraint function and augments the surrogate cost with a prediction-dependent penalty term $\mu_t \Phi'(Q(t)) \tilde{g}_t^{\text{pred}}(x)$ that anticipates future constraint violations. This predictive augmentation introduces a "bonus" term in the Lyapunov drift analysis that can compensate for negative regret, allowing us to replace the non-negative regret assumption with a strictly weaker posterior condition, namely the prediction compensation event $\mathcal{E} := \{V\text{Regret}_T + \text{Bonus}_T \geq 0\}$. Since non-negative regret implies $\mathcal{E}$ but not conversely, this is a genuine relaxation. To ensure robustness against poor predictions, we embed the prediction-augmented sub-policy alongside a prediction-free baseline within a Hedge meta-algorithm (Freund & Schapire, 1997) that dynamically allocates weight between the two, guaranteeing that the overall CCV never exceeds the baseline's $O(\sqrt{T})$ rate regardless of prediction quality. Table 1 summarizes the comparison with prior work.

Our main contributions are as follows.

- We propose LA-COCO, the first learning-augmented framework for online convex optimization with adversarial constraints. Under α-strongly convex costs, a Hedge meta-algorithm yields $\mathrm{CCV}_T = O(G^2 \log T/\alpha + \eta_T)$ when the prediction compensation event $\mathcal{E}$ holds (a condition strictly weaker than the non-negative regret assumption in prior work), and an unconditional $O(\sqrt{T})$ fallback otherwise (Theorem 1). For general convex costs, we establish $\mathrm{CCV}_T = O(GD\sqrt{T}\log T)$ unconditionally and a doubling-trick variant that makes the algorithm fully online (Proposition 1).
- We derive a prediction-augmented regret decomposition (Lemma 3) that extends the Lyapunov drift framework with non-negative bonus and penalty terms capturing the value and cost of constraint predictions. This inequality serves as the unified analytical engine behind all our results and provides a modular template of independent interest for future learning-augmented constrained optimization.
- We conduct experiments across 120 configurations with rotating adversarial constraints, confirming that $\mathcal{E}$ holds with probability 1 under stochastic adversaries, all empirical CCV values lie below the theoretical bounds, and the decomposition inequality is verified with non-negative gap throughout.

The remainder of this paper is organized as follows. Section 2 introduces the problem formulation and assumptions. Section 3 presents the LA-COCO algorithm. Section 4 states the main theoretical results. Section 5 reports experimental validation. Section 6 concludes. Related work is discussed in Appendix B, and additional experiments appear in Appendix D.

2 Problem Setup

We consider an online convex optimization problem in which a learner faces an adaptive adversary who controls both cost and constraint functions. The interaction proceeds over T rounds. At each round $t = 1, \ldots, T$, the learner selects an action x_t from a closed convex set $\mathcal{X} \subseteq \mathbb{R}^d$ with Euclidean diameter $D := \sup_{x,y\in\mathcal{X}} \|x - y\|$. The adversary then reveals a convex cost function $f_t : \mathcal{X} \to \mathbb{R}$ and a convex constraint function $g_t : \mathcal{X} \to \mathbb{R}$. The learner incurs cost $f_t(x_t)$ and constraint violation $(g_t(x_t))^+ := \max(g_t(x_t), 0)$. The learner's goal is to simultaneously minimize cumulative cost relative to the best fixed action in hindsight and keep the cumulative constraint violation small.

We measure performance through two standard metrics. The static regret against a comparator $x^\star \in \mathcal{X}$ is

$$\mathrm{Regret}_T(x^\star) := \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x^\star), \tag{1}$$

and the cumulative constraint violation is

$$\mathrm{CCV}_T := \sum_{t=1}^{T} (g_t(x_t))^+. \tag{2}$$

The comparator set is the collection of actions that are feasible for every constraint function in the sequence: $\mathcal{X}^\star := \{x \in \mathcal{X} : g_t(x) \le 0,\ \forall\, t = 1, \ldots, T\}$. We seek algorithms for which both $\mathrm{Regret}_T(x^\star)$ and CCV_T grow sublinearly in T for all $x^\star \in \mathcal{X}^\star$.

In addition to the adversarial cost and constraint functions, our framework incorporates a prediction oracle that provides, at the end of each round t, a prediction $\hat{g}_{t+1} : \mathcal{X} \to \mathbb{R}$ of the next constraint function g_{t+1}. The quality of this prediction is measured by the uniform approximation error $\varepsilon_{t+1} := \sup_{x\in\mathcal{X}} |g_{t+1}(x) - \hat{g}_{t+1}(x)|$, which is determined in hindsight once g_{t+1} is revealed. The cumulative prediction error over the horizon is $\eta_T := \sum_{t=1}^{T} \varepsilon_t$. No distributional or boundedness assumptions are imposed on η_T; the algorithm must perform well regardless of prediction quality.

Assumptions. We state the assumptions under which our results hold. Assumptions 1–2 are standard in the constrained OCO literature (Sinha & Vaze, 2024; Neely, 2017; Yi et al., 2023). Assumption 3 is invoked only for the strongly convex results (Theorem 1). Assumption 4 is the only additional requirement introduced by our framework.

Assumption 1 (Convexity and Lipschitz continuity)**.** *For each $t = 1, \ldots, T$, the cost function $f_t : \mathcal{X} \to \mathbb{R}$ and the constraint function $g_t : \mathcal{X} \to \mathbb{R}$ are convex and G-Lipschitz continuous with respect to the Euclidean norm: $|f_t(x) - f_t(y)| \leq G\|x - y\|$ and $|g_t(x) - g_t(y)| \leq G\|x - y\|$ for all $x, y \in \mathcal{X}$.*

The Lipschitz constant G bounds the magnitude of subgradients and ensures that per-round cost and constraint variation are controlled by the diameter D. This is a minimal regularity condition satisfied by virtually all OCO formulations.

Assumption 2 (Feasibility)**.** *The comparator set $\mathcal{X}^\star := \{x \in \mathcal{X} : g_t(x) \leq 0,\ \forall\, t\}$ is non-empty.*

Feasibility of the comparator set is necessary for the CCV metric to be meaningful: without a benchmark that satisfies all constraints, one cannot meaningfully measure how much the learner's actions violate them. This assumption does not require the learner to know $\mathcal{X}^\star$ or to play feasible actions at each round.

Assumption 3 (Strong convexity)**.** *Each cost function f_t is α-strongly convex over $\mathcal{X}$: $f_t(y) \geq f_t(x) + \nabla f_t(x)^\top (y - x) + \frac{\alpha}{2}\|y - x\|^2$ for all $x, y \in \mathcal{X}$.*

Strong convexity enables logarithmic regret for the surrogate cost minimization subroutine and is the key structural property that allows the CCV to be reduced from $O(\sqrt{T})$ to $O(\log T)$. This assumption is standard in the strongly convex OCO literature (Hazan, 2016; Neely, 2017) and is satisfied, for example, by quadratic cost functions arising in online regression and resource allocation.

Assumption 4 (Prediction availability)**.** *At the end of each round t, the learner receives a prediction $\hat{g}_{t+1} : \mathcal{X} \to \mathbb{R}$ that is convex and G-Lipschitz continuous. No assumption is made on the cumulative prediction error η_T.*

The convexity and Lipschitz requirements on $\hat{g}_{t+1}$ ensure that the prediction-augmented surrogate cost remains convex and has bounded subgradients, which is needed for the OGD subroutine. The absence of any prior bound on η_T is a deliberate design choice: the algorithm's robustness guarantee (the $O(\sqrt{T})$ unconditional fallback) must hold even when predictions are adversarially bad, while the consistency guarantee (the $O(\log T)$ conditional bound) improves gracefully as η_T decreases.

Preprocessing and Lyapunov tracking. Following standard practice in the constrained OCO literature (Neely, 2017), we introduce a scaling parameter $\beta > 0$ and define the preprocessed functions $\tilde{f}_t := \beta f_t$, $\tilde{g}_t := \beta (g_t)^+$, and the prediction-based term $\tilde{g}_t^{\text{pred}} := \beta(\hat{g}_{t+1})^+$. The subscript t on $\tilde{g}_t^{\text{pred}}$ indicates that this quantity is used in the round-t surrogate cost, even though it is constructed from the prediction $\hat{g}_{t+1}$ of the next constraint. The cumulative constraint violation is tracked by a virtual queue $Q(t) := Q(t-1) + \tilde{g}_t(x_t)$ with $Q(0) = 0$, so that $\text{CCV}_T = \beta^{-1} Q(T)$. A Lyapunov function $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\Phi(0) = 0$ and Φ convex and non-decreasing is applied to $Q(t)$ to construct the drift-plus-penalty surrogate. In the strongly convex setting we use $\Phi(x) = x^2$; in the general convex setting we use $\Phi(x) = e^{\lambda x} - 1$ for a parameter $\lambda > 0$.

Optimization objective. The prediction-augmented surrogate cost that the learner minimizes at each round is

$$\hat{f}_t(x) := V \tilde{f}_t(x) + \Phi'(Q(t))\, \tilde{g}_t(x) + \mu_t\, \Phi'(Q(t))\, \tilde{g}_t^{\text{pred}}(x), \tag{3}$$

where $V > 0$ is a trade-off parameter that balances cost minimization against constraint satisfaction, and $\mu_t \in [0, 1]$ is a trust parameter that controls the influence of the prediction. Setting $\mu_t = 0$ recovers the prediction-free baseline; setting $\mu_t = 1$ fully incorporates the prediction. The surrogate equation 3 inherits convexity from its three constituent terms and is α-strongly convex (scaled by $V\beta$) when f_t is α-strongly convex. The learner applies online gradient descent to equation 3 with step size $\eta_t = 1/(V\beta\alpha t)$ in the strongly convex setting and an AdaGrad-style step size in the general convex setting. The third term in equation 3, the prediction-dependent penalty, is the key departure from prior work: it anticipates future constraint violations through $\hat{g}_{t+1}$ and introduces a non-negative bonus into the Lyapunov drift analysis that can compensate for negative regret, enabling the replacement of the non-negative regret assumption with the weaker prediction compensation event $\mathcal{E}$ defined in Section 4.

3 Algorithm

This section presents the LA-COCO framework in full detail. We begin with the two sub-policies that form the core of the algorithm, then describe the Hedge meta-algorithm that combines them, and finally discuss a doubling-trick variant for the general convex setting where the oracle parameter λ depends on the unknown cumulative prediction error η_T.

3.1 Sub-policies

LA-COCO maintains two (or three, in the convex setting) independent sub-policies, each running its own instance of online gradient descent on a surrogate cost function and maintaining its own virtual queue.

Sub-policy A (prediction-free baseline). Sub-policy A sets $\mu_t = 0$ for all t, discarding the prediction entirely. Its surrogate cost reduces to

$$\hat{f}_t^{(A)}(x) = V\tilde{f}_t(x) + \Phi'(Q^{(A)}(t))\,\tilde{g}_t(x), \tag{4}$$

which coincides with the standard Lyapunov-based surrogate. The virtual queue $Q^{(A)}(t) = Q^{(A)}(t-1) + \tilde{g}_t(x_t^{(A)})$ tracks the constraint violations of sub-policy A's own action sequence. In the strongly convex setting, OGD is applied to equation 4 with step size $\eta_t = 1/(V\alpha t)$; in the convex setting, an AdaGrad-style step size $\eta_t = D/\sqrt{\sum_{s=1}^{t} \|\nabla \hat{f}_s^{(A)}(x_s^{(A)})\|^2}$ is used. Sub-policy A provides the unconditional $O(\sqrt{T})$ CCV guarantee that serves as the robustness baseline.

Sub-policy B (prediction-augmented). Sub-policy B sets $\mu_t = 1$ for all t, fully incorporating the prediction. Its surrogate cost is the complete expression equation 3:

$$\hat{f}_t^{(B)}(x) = V\tilde{f}_t(x) + \Phi'(Q^{(B)}(t))\,\tilde{g}_t(x) + \Phi'(Q^{(B)}(t))\,\tilde{g}_t^{\text{pred}}(x), \tag{5}$$

with its own virtual queue $Q^{(B)}(t) = Q^{(B)}(t-1) + \tilde{g}_t(x_t^{(B)})$. The third term $\Phi'(Q^{(B)}(t))\,\tilde{g}_t^{\text{pred}}(x)$ penalizes actions that the predicted next constraint $\hat{g}_{t+1}$ deems infeasible, steering the learner away from anticipated violations before they occur. The step size schedule is identical to that of sub-policy A. When predictions are accurate, sub-policy B achieves $O(\log T)$ CCV under the prediction compensation event $\mathcal{E}$; when predictions are poor, its CCV may exceed that of sub-policy A, but the Hedge layer ensures this does not degrade the overall guarantee.

Sub-policy C (regret control, convex setting only). In the general convex setting, a third sub-policy is introduced to control the regret of the Hedge output. Sub-policy C runs standard OGD directly on the original cost functions f_t with step size $\eta = D/(G\sqrt{T})$, ignoring constraints entirely. Its regret satisfies $\text{Regret}_T^{(C)}(x^\star) \le GD\sqrt{T}$ by the classical OGD guarantee. Sub-policy C is not needed in the strongly convex setting, where the surrogate-based sub-policies already achieve $O(\log T)$ regret.

3.2 Hedge meta-algorithm

The sub-policies are combined via the Hedge algorithm (Freund & Schapire, 1997), which maintains a probability distribution over the N sub-policies and updates it based on their per-round constraint violations. The complete procedure is given in Algorithm 1.

At each round, the algorithm plays the convex combination $x_t = \sum_i w_i(t)\, x_t^{(i)}$, which lies in $\mathcal{X}$ by convexity. Each sub-policy then independently updates its own state: it evaluates the revealed functions at its own action $x_t^{(i)}$ (not at the mixed action x_t), updates its virtual queue, computes the surrogate gradient, and performs a projected gradient step. The Hedge weights are updated using the normalized constraint violations $\ell_i(t) = (g_t(x_t^{(i)}))^+/(GD) \in [0,1]$ as expert losses. This ensures that sub-policies with lower constraint violations receive increasing weight over time.

Algorithm 1 LA-COCO (Learning-Augmented Constrained Online Convex Optimization)

Require: Convex set $\mathcal{X} \subseteq \mathbb{R}^d$, horizon T, parameters V, β, α (or λ), prediction oracle
  STRONGLY CONVEX: $N = 2$, $V = 64G^2 \ln(Te)/\alpha$, $\beta = 1$, $\Phi(x) = x^2$
  CONVEX: $N = 3$, $V = 1$, $\beta = (2GD)^{-1}$, $\Phi(x) = e^{\lambda x} - 1$
Initialize $x_1^{(i)} \in \mathcal{X}$ for each sub-policy $i \in \{1, \dots, N\}$
Initialize virtual queues $Q^{(i)}(0) = 0$ for each i
Initialize Hedge weights $w_i(1) = 1/N$, learning rate $\eta_H = \sqrt{\ln N/T}$
for $t = 1, \dots, T$ **do**
  Compute mixed action $x_t = \sum_{i=1}^{N} w_i(t)\, x_t^{(i)}$
  Play x_t; observe cost f_t and constraint g_t from adversary
  Receive prediction $\hat{g}_{t+1}$ from oracle
  for $i = 1, \dots, N$ **do** ▷ Update each sub-policy independently
    Evaluate $f_t(x_t^{(i)})$, $g_t(x_t^{(i)})$, and (for sub-policy B) $\hat{g}_{t+1}(x_t^{(i)})$
    $Q^{(i)}(t) \leftarrow Q^{(i)}(t-1) + \beta\,(g_t(x_t^{(i)}))^+$
    Compute $\nabla \hat{f}_t^{(i)}(x_t^{(i)})$ per equation 4 or equation 5
    if strongly convex **then**
      $\eta_t^{(i)} \leftarrow 1/(V\beta\alpha\, t)$
    else
      $\eta_t^{(i)} \leftarrow D/\sqrt{\sum_{s=1}^{t} \|\nabla \hat{f}_s^{(i)}(x_s^{(i)})\|^2}$
    end if
    $x_{t+1}^{(i)} \leftarrow \Pi_{\mathcal{X}}\left(x_t^{(i)} - \eta_t^{(i)}\, \nabla \hat{f}_t^{(i)}(x_t^{(i)})\right)$
  end for
  for $i = 1, \dots, N$ **do** ▷ Update Hedge weights
    $\ell_i(t) \leftarrow (g_t(x_t^{(i)}))^+/(GD)$
  end for
  $w_i(t+1) \propto w_i(t) \exp(-\eta_H\, \ell_i(t))$ for each i, normalized so $\sum_i w_i(t+1) = 1$
end for

The decoupling between the played action x_t and the sub-policy updates is a standard feature of the experts framework: each sub-policy operates as if it were the sole decision-maker, and the Hedge layer aggregates their outputs. The CCV of the mixed action is controlled by the convexity of g_t and the $(\cdot)^+$ operator, which together yield $(g_t(x_t))^+ \leq \sum_i w_i(t)(g_t(x_t^{(i)}))^+$. Combined with the Hedge regret guarantee, this gives $\mathrm{CCV}_T^{\mathrm{Hedge}} \leq \min_i \mathrm{CCV}_T^{(i)} + 2GD\sqrt{T \ln N}$ (Lemma 7 in Section 4).

3.3 Parameter selection and online realizability

In the strongly convex setting, all parameters of Algorithm 1 are determined by the known quantities T, G, D, and α: the trade-off parameter $V = 64G^2 \ln(Te)/\alpha$, the scaling $\beta = 1$, the Lyapunov function $\Phi(x) = x^2$, and the trust parameter $\mu_t = 1$ (for sub-policy B) or $\mu_t = 0$ (for sub-policy A). The algorithm is therefore fully online and requires no knowledge of the adversary's strategy or the prediction quality η_T.

In the general convex setting, the analysis-optimal parameter $\lambda = 1/(2(\sqrt{T} + \beta\eta_T))$ for the exponential Lyapunov function $\Phi(x) = e^{\lambda x} - 1$ depends on the cumulative prediction error η_T, which is unknown before the algorithm terminates. Theorem 2 in Section 4 is therefore an oracle theorem: it characterizes the performance achievable with the best parameter choice in hindsight. To obtain a fully online algorithm, we employ a standard doubling trick.

Algorithm 2 replaces sub-policy B in Algorithm 1 for the convex setting. It maintains a running estimate $\hat{\eta}_k$ of the cumulative prediction error, initialized to 1 and doubled whenever the actual accumulated error exceeds the current estimate. At each doubling event, the Lyapunov parameter λ_k is recalculated and the virtual queue $Q^{(B)}$ is reset to zero, effectively restarting the Lyapunov analysis for a new phase. Sub-policy

Algorithm 2 LA-COCO with Doubling Trick (convex setting, sub-policy B only)

Require: Convex set $\mathcal{X}$, horizon T, parameters $V = 1$, $\beta = (2GD)^{-1}$
$k \leftarrow 0$, $\hat{\eta}_0 \leftarrow 1$, $t_0 \leftarrow 1$, $\eta_{\text{accum}} \leftarrow 0$
$\lambda_0 \leftarrow 1/(2(\sqrt{T} + \beta\hat{\eta}_0))$
Initialize sub-policy B with $\Phi_0(x) = e^{\lambda_0 x} - 1$, $Q^{(B)} = 0$
for $t = 1, \ldots, T$ **do**
  Run sub-policy B for round t (output $x_t^{(B)}$, update $Q^{(B)}$)
  Observe g_t; compute $\varepsilon_t = \sup_{x \in \mathcal{X}} |g_t(x) - \hat{g}_t(x)|$
  $\eta_{\text{accum}} \leftarrow \eta_{\text{accum}} + \varepsilon_t$
  if $\eta_{\text{accum}} > \hat{\eta}_k$ **then**
    $k \leftarrow k + 1$, $\hat{\eta}_k \leftarrow 2\hat{\eta}_{k-1}$, $t_k \leftarrow t + 1$
    $\lambda_k \leftarrow 1/(2(\sqrt{T} + \beta\hat{\eta}_k))$
    Reset: $Q^{(B)} \leftarrow 0$, $\eta_{\text{accum}} \leftarrow 0$
    Reinitialize sub-policy B with $\Phi_k(x) = e^{\lambda_k x} - 1$
  end if
end for

A and the Hedge layer are unaffected by the restarts: from the Hedge perspective, sub-policy B remains a single expert that outputs an action $x_t^{(B)} \in \mathcal{X}$ at every round.

The number of phases is at most $K + 1 \leq \lceil \log_2 \eta_T \rceil + 2$, since the geometric doubling ensures $\hat{\eta}_K \leq 2\eta_T$. The CCV of sub-policy B across all phases satisfies $\text{CCV}_T^{(B)} = \sum_{k=0}^{K} \text{CCV}_k$, where each phase contributes $\text{CCV}_k \leq O((GD\sqrt{T} + \beta\hat{\eta}_k) \log T)$ by applying the single-phase analysis of Theorem 2. Summing over phases, the total CCV acquires an additional $\log \eta_T$ factor on the $GD\sqrt{T}$ term compared to the oracle bound, while the η_T term is preserved up to a constant (Proposition 1 in Section 4).

3.4 Computational complexity

Each round of Algorithm 1 requires $O(Nd)$ arithmetic operations: each of the $N \leq 3$ sub-policies computes one surrogate gradient ($O(d)$), performs one projected gradient step ($O(d)$ for projection onto common convex sets such as ℓ_2 balls or simplices), and evaluates the constraint violation ($O(1)$). The Hedge weight update costs $O(N)$. The prediction oracle is treated as a black box; its computational cost depends on the specific prediction mechanism and is not counted. Storage is $O(Nd)$ for the sub-policy iterates plus $O(N)$ for the virtual queues and Hedge weights. The per-round cost is therefore dominated by the N gradient evaluations, which is the same order as running N independent instances of OGD.

4 Theoretical Results

This section presents the complete theoretical analysis of LA-COCO. We first establish the supporting lemmas (§4.1), then state and prove the main theorems for the strongly convex setting (§4.2), the general convex setting (§4.3), and the sufficient conditions for the prediction compensation event (§4.4). Corollaries and the online-realizable variant are collected in §4.5.

4.1 Supporting lemmas

We begin with the Lyapunov drift inequality, which is the foundation of the entire analysis.

Lemma 1 (Drift inequality)**.** *For all $t \geq 1$,*

$$\Phi(Q(t)) - \Phi(Q(t-1)) \leq \Phi'(Q(t))\, \tilde{g}_t(x_t). \tag{6}$$

The proof is given in Appendix A.1.

Lemma 2 (Gradient norm bound)**.** *The surrogate gradient norm satisfies*

$$G_t := \|\nabla \hat{f}_t(x_t)\| \le \beta G(V + (1+\mu_t)\Phi'(Q(t))). \tag{7}$$

The proof is given in Appendix A.2.

The next lemma is the central analytical tool of this paper. It extends the standard Lyapunov drift analysis with bonus and penalty terms that capture the value and cost of predictions.

Lemma 3 (Prediction-augmented regret decomposition)**.** *For any $x^\star \in \mathcal{X}^\star$ and any $t \ge 1$,*

$$\Phi(Q(t)) + V\beta\,\mathrm{Regret}_t(x^\star) + \mathrm{Bonus}_t \le \mathrm{Regret}'_t(x^\star) + \mathrm{Penalty}_t, \tag{8}$$

where $\mathrm{Regret}'_t(x^\star) := \sum_{\tau=1}^t [\hat{f}_\tau(x_\tau) - \hat{f}_\tau(x^\star)]$ *is the surrogate regret, and*

$$\mathrm{Bonus}_t := \sum_{\tau=1}^t \mu_\tau \Phi'(Q(\tau))\,\tilde{g}_\tau^{\mathrm{pred}}(x_\tau) \ge 0, \tag{9}$$

$$\mathrm{Penalty}_t := \sum_{\tau=1}^t \mu_\tau \Phi'(Q(\tau))\,\beta\varepsilon_{\tau+1} \ge 0. \tag{10}$$

The proof is given in Appendix A.3.

Lemma 4 (Bonus decomposition under quadratic Lyapunov)**.** *Let $\Phi(x) = x^2$, $\mu_\tau = 1$, $\beta = 1$, and $\Delta_{\tau+1} := (g_{\tau+1}(x_{\tau+1}))^+ \ge 0$. Under perfect prediction ($\varepsilon_{t+1} = 0$),*

(a) $\mathrm{Bonus}_t = \sum_{\tau=1}^t 2Q(\tau)(g_{\tau+1}(x_\tau))^+$*;*

(b) $\sum_{\tau=1}^t 2Q(\tau)\Delta_{\tau+1} = Q^2(t+1) - Q^2(1) - \sum_{\tau=1}^t \Delta_{\tau+1}^2$*;*

(c) $\Delta_{\tau+1} \le GD$ *implies* $\sum_{\tau=1}^t \Delta_{\tau+1}^2 \le GD\big(Q(t+1) - Q(1)\big)$*;*

(d) combining (a)–(c) with the Lipschitz bound $(g_{\tau+1}(x_\tau))^+ \ge \Delta_{\tau+1} - G\|x_\tau - x_{\tau+1}\|$*,*

$$\mathrm{Bonus}_t \ge Q^2(t+1) - Q^2(1) - GD\big(Q(t+1) - Q(1)\big) - 2GQ(T)\sum_{\tau=1}^t \|x_\tau - x_{\tau+1}\|. \tag{11}$$

The gap between $(g_{\tau+1}(x_\tau))^+$ in part (a) and $\Delta_{\tau+1} = (g_{\tau+1}(x_{\tau+1}))^+$ is bridged by the non-expansiveness of $(\cdot)^+$ and G-Lipschitz continuity of $g_{\tau+1}$; the resulting path-length term is controlled by Lemma 5. The proof is given in Appendix A.4.

Lemma 5 (Path length bound)**.** *Let $G_{\max} := \max_{1\le t\le T} G_t$.*

(a) Strongly convex setting (step size $\eta_t = 1/(V\alpha t)$):

$$\sum_{t=1}^T \|x_t - x_{t+1}\| \le D + \frac{G_{\max}}{V\alpha}\ln T. \tag{12}$$

(b) Convex setting (AdaGrad step size $\eta_t = D/\sqrt{\sum_{s=1}^t G_s^2}$):

$$\sum_{t=1}^T \|x_t - x_{t+1}\| \le 2D\sqrt{\sum_{t=1}^T G_t^2} \le 2DG_{\max}\sqrt{T}. \tag{13}$$

The proof is given in Appendix A.5.

In the convex setting, $G_{\max}$ depends on $Q(T)$ through Lemma 2, so equation 13 cannot serve as an independent path-length control; the proof of Theorem 2 proceeds directly via Lemma 3. In the strongly convex setting, the circular dependence through $G_{\max}/(V\alpha)$ is absorbed by the leading term.

Lemma 6 (Strongly convex surrogate regret)**.** *When $\hat{f}_t$ is $V\alpha$-strongly convex and OGD is run with step size $\eta_t = 1/(V\alpha t)$,*

$$\mathrm{Regret}'_T(x^\star) \le \frac{1}{2}\sum_{t=1}^{T} \frac{G_t^2}{V\alpha\, t}. \tag{14}$$

The proof is given in Appendix A.6.

Lemma 7 (Hedge guarantees)**.** *Let the Hedge algorithm run with N sub-policies, mixed action $x_t = \sum_i w_i(t) x_t^{(i)}$, and learning rate $\eta_H = \sqrt{\ln N/T}$.*

(i) CCV control:

$$\mathrm{CCV}_T^{\mathrm{Hedge}} \le \min_i \mathrm{CCV}_T^{(i)} + 2GD\sqrt{T\ln N}. \tag{15}$$

(ii) Regret control:

$$\mathrm{Regret}_T^{\mathrm{Hedge}}(x^\star) \le \min_i \mathrm{Regret}_T^{(i)}(x^\star) + 2GD\sqrt{T\ln N}. \tag{16}$$

The proof is given in Appendix A.7.

Lemma 8 (Baseline CCV bounds for sub-policy A)**.** *(a) Strongly convex setting ($\Phi(x) = x^2$, $V = 64G^2\ln(Te)/\alpha$, $\beta = 1$):*

$$\mathrm{CCV}_T^{(A)} \le \sqrt{\frac{16}{15}\left(\frac{64G^4\ln^2(Te)}{\alpha^2} + \frac{64G^3 D\ln(Te)}{\alpha}\,T\right)} = O\left(\sqrt{\frac{G^3 D\ln(Te)\cdot T}{\alpha}}\right). \tag{17}$$

(b) Convex setting ($\Phi(x) = e^{\lambda x} - 1$, $V = 1$, $\beta = (2GD)^{-1}$):

$$\mathrm{CCV}_T^{(A)} = O(GD\sqrt{T}\log T). \tag{18}$$

The proof is given in Appendix A.8.

If one additionally assumes $\mathrm{Regret}_T^{(A)} \ge 0$ (the non-negative regret condition), then equation 17 in part (a) sharpens to $Q(T) = O(G^2\log T/\alpha)$, recovering the logarithmic rate directly. The unconditional bound in part (a) arises from the worst-case linear control $|\mathrm{Regret}_T^{(A)}| \le GDT$ and serves as the mathematical foundation for the Hedge unconditional guarantee (Theorem 1(iii)).

Lemma 9 (Quadratic inequality)**.** *If $q \ge 0$, $a > 0$, $A > 0$, $\eta \ge 0$ satisfy $aq^2 \le A + 2q\eta$, then*

$$q \le \frac{2\eta}{a} + \sqrt{\frac{A}{a}}. \tag{19}$$

The proof is given in Appendix A.9.

4.2 Theorem 1: Strongly convex costs with predictions

Theorem 1 (Strongly convex setting)**.** *Under Assumptions 1–4, let sub-policy B run with $\beta = 1$, $V = 64G^2\ln(Te)/\alpha$, $\Phi(x) = x^2$, $\mu_t = 1$, and let sub-policy A run with the same parameters but $\mu_t = 0$. Then:*

(i) Regret of sub-policy B *(unconditional).*

$$V\mathrm{Regret}_T^{(B)}(x^\star) \le \frac{VG^2\ln(Te)}{\alpha} + 2Q(T)(\eta_T + 2GD). \tag{20}$$

(ii) CCV of sub-policy B *(conditional on $\mathcal{E}$).*

$$\mathrm{CCV}_T^{(B)} \le \frac{16G^2\ln(Te)}{\alpha} + 4\eta_T + 6GD. \tag{21}$$

(iii) Hedge CCV *(unconditional).*

$$\mathrm{CCV}_T^{\mathrm{Hedge}} \leq O\left(\sqrt{\frac{G^3 D \ln(Te) \cdot T}{\alpha}}\right) + O(GD\sqrt{T}). \tag{22}$$

(iv) Hedge CCV *(conditional on $\mathcal{E}$).*

$$\mathrm{CCV}_T^{\mathrm{Hedge}} \leq \min\left(\frac{16G^2 \ln(Te)}{\alpha} + 4\eta_T + 6GD,\ O\left(\sqrt{\frac{G^3 D \ln(Te) \cdot T}{\alpha}}\right)\right) + O(GD\sqrt{T}). \tag{23}$$

The proof is given in Appendix A.10.

Remark 1 (Special cases of Theorem 1)**.** *Several useful bounds follow directly from the theorem. Setting $\eta_T = 0$ in equation 20 and dividing by V gives $\mathrm{Regret}_T^{(B)} \leq G^2 \ln(Te)/\alpha + 4GDQ(T)/V$; under $\mathcal{E}$, the CCV bound equation 21 implies $Q(T) = O(G^2 \log T/\alpha)$, so the second term is $O(GD)$ and is absorbed by the leading term. More generally, substituting equation 21 into equation 20 under $\mathcal{E}$ yields*

$$\mathrm{Regret}_T^{(B)}(x^\star) \leq O\left(\frac{G^2 \log T}{\alpha} + \frac{\eta_T}{\alpha} + \frac{\alpha \eta_T^2}{G^2 \log T}\right). \tag{24}$$

Setting $\eta_T = 0$ in equation 21 gives $\mathrm{CCV}_T^{(B)} = O(G^2 \log T/\alpha)$, recovering the logarithmic rate under perfect prediction.

4.3 Theorem 2: General convex costs with predictions

Theorem 2 (Convex setting)**.** *Under Assumptions 1, 2, and 4, let sub-policy B use $\Phi(x) = e^{\lambda x} - 1$, $V = 1$, $\beta = (2GD)^{-1}$, $\mu = 1$; sub-policy A use the same parameters with $\mu = 0$; and sub-policy C run standard OGD on f_t. With the analysis-optimal choice $\lambda = 1/(2(\sqrt{T} + \beta\eta_T))$, the following bounds hold unconditionally:*

(i) Hedge regret. $\mathrm{Regret}_T^{\mathrm{Hedge}} \leq O(GD\sqrt{T})$.

(ii) Hedge CCV. $\mathrm{CCV}_T^{\mathrm{Hedge}} \leq O(GD\sqrt{T}\log T)$.

(iii) CCV of sub-policy B. $\mathrm{CCV}_T^{(B)} \leq O((GD\sqrt{T} + \eta_T)\log T)$.

(iv) Hedge CCV (detailed).

$$\mathrm{CCV}_T^{\mathrm{Hedge}} = O\Big(\min\Big((GD\sqrt{T} + \eta_T)\log T,\ GD\sqrt{T}\log T\Big) + GD\sqrt{T}\Big). \tag{25}$$

The proof is given in Appendix A.11.

Remark 2. *Since $\eta_T \geq 0$, the* min *in equation 25 is always achieved by the sub-policy A branch. Thus, in the convex setting, predictions do not improve the asymptotic CCV of the Hedge output. The value of predictions is confined to sub-policy B's own CCV bound (iii). This limitation stems from the $O(\sqrt{T})$ surrogate regret bottleneck inherent to convex (non-strongly-convex) losses.*

4.4 Theorem 3: Sufficient conditions for $\mathcal{E}$

We now formalize the prediction compensation event and establish conditions under which it holds.

Definition 1 (Prediction compensation event)**.** *The prediction compensation event is*

$$\mathcal{E} := \{V\mathrm{Regret}_T(x^\star) + \mathrm{Bonus}_T \geq 0, \quad \forall x^\star \in \mathcal{X}^\star\},$$

where $\mathrm{Bonus}_T = \sum_{\tau=1}^{T} \mu_\tau \Phi'(Q(\tau))\tilde{g}_\tau^{\mathrm{pred}}(x_\tau) \geq 0$.

Theorem 3 (Sufficient conditions for $\mathcal{E}$). *(a) If* $\mathrm{Regret}_T(x^\star) \geq 0$ *for all* $x^\star \in \mathcal{X}^\star$*, then* $\mathcal{E}$ *holds.*

(b) If $\eta_T = 0$ *and* $\sum_{t=1}^{T} Q(t)(g_{t+1}(x_t))^+ \geq \frac{V}{2}\sum_{t=1}^{T} |f_t(x_t) - f_t(x^\star)|$ *for all* $x^\star \in \mathcal{X}^\star$*, then* $\mathcal{E}$ *holds.*

The proof is given in Appendix A.12.

Remark 3. *The event* $\mathcal{E}$ *is a posterior condition whose role is to keep the left-hand side of equation 30 non-negative, enabling a sublinear bound on* $Q(T)$*. Non-negative regret implies* $\mathcal{E}$ *(part (a)), but* $\mathcal{E}$ *can also hold when regret is negative provided the bonus compensates, making it strictly weaker than the non-negative regret assumption in prior work. Part (b) identifies a complementary regime where frequently active constraints ensure compensation even under negative regret. Neither condition is guaranteed against a fully adaptive adversary, and when* $\mathcal{E}$ *fails the Hedge layer provides the unconditional* $O(\sqrt{T})$ *fallback of Theorem 1(iii).*

4.5 Corollaries and online realizability

Corollary 1 (Prediction quality threshold). *Define* $\eta_T^\star := O(\sqrt{G^3 DT \log T/\alpha})$*. When* $\mathcal{E}$ *holds and* $\eta_T \leq \eta_T^\star$*, the prediction-augmented bound equation 21 is tighter than the baseline equation 22.*

The proof is given in Appendix A.13.

Corollary 2 (Improvement factor under perfect prediction). *When* $\eta_T = 0$ *and* $\mathcal{E}$ *holds, the CCV improves from* $O(\sqrt{G^3 DT \log T/\alpha})$ *to* $O(G^2 \log T/\alpha)$*, an improvement factor of* $\Theta(\sqrt{GDT/\log T})$*.*

Corollary 3 (Online convex set chasing). *When* $f_t = 0$ *for all* t*,* $\mathrm{Regret}_T = 0$*, so* $\mathcal{E}$ *holds unconditionally by Theorem 3(a). The logarithmic CCV bound equation 21 applies without any posterior condition.*

Proposition 1 (Online-realizable variant for the convex setting). *Algorithm 2 achieves*

$$\mathrm{CCV}_T^{(B)} \leq O\Big((GD\sqrt{T}\log \eta_T + \eta_T)\log T\Big). \tag{26}$$

The additional cost compared to the oracle bound of Theorem 2(iii) is a $\log \eta_T$ *factor on the* $GD\sqrt{T}$ *term. The Hedge guarantees (i)–(ii) remain unchanged.*

The proof is given in Appendix A.14.

5 Experiments

We evaluate LA-COCO through simulations on constrained online optimization problems with rotating adversarial constraints. These experiments are designed to confirm that the empirical CCV of each algorithm component remains below its corresponding theoretical upper bound. The simulations also validate the structural predictions of our theory, specifically the prediction compensation event $\mathcal{E}$ and the regret decomposition inequality of Lemma 3. Alongside these theoretical validations, we compare the CCV growth rates of LA-COCO against standard baselines. The detailed experimental setup, including adversary design, algorithm specifications, and parameter choices, is provided in Appendix C.

5.1 CCV growth rate comparison

Figure 1 compares the CCV growth rates of all five methods as T ranges from 50 to 10,000 under the stochastic adversary with perfect predictions ($\eta_T = 0$). Panel (a) shows the absolute CCV on a log-log scale. Naive OGD, which ignores constraints entirely, exhibits near-linear CCV growth. Primal-Dual achieves $\mathrm{CCV}_T \sim T^{0.52}$, consistent with its $O(\sqrt{T})$ theoretical guarantee. Both Lyapunov-based sub-policies (A and B) grow substantially slower, with fitted exponents $\sim T^{0.08}$ and $\sim T^{0.06}$ respectively—well below $\sqrt{T}$ and consistent with the logarithmic regime predicted by the theory when $\mathcal{E}$ holds. The Hedge output tracks the better of the two sub-policies with a small additive overhead. Panel (b) shows the normalized CCV/T, confirming that all Lyapunov-based methods converge to zero at a rate faster than $1/\sqrt{T}$, while Primal-Dual converges at the expected $1/\sqrt{T}$ rate.

Additional experiments, including theoretical upper bound verification (Appendix D.1), regret decomposition verification (Appendix D.2), and online convex set chasing (Appendix D.3), confirm that all empirical CCV

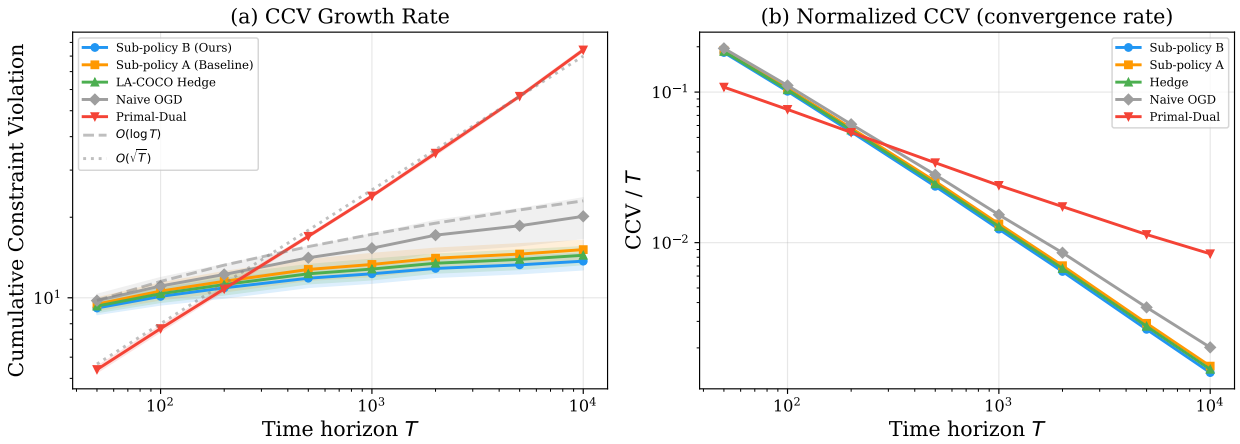

Figure 1: CCV growth rate comparison under the stochastic adversary with perfect predictions ($d = 5$, $G = D = \alpha = 1$, 10 seeds per T). (a) Absolute CCV on log-log scale. Reference lines show $O(\log T)$ and $O(\sqrt{T})$ growth. (b) Normalized CCV$/T$ showing convergence rates.

values lie strictly below the theoretical bounds. Furthermore, the decomposition inequality of Lemma 3 holds with a non-negative gap across all 120 configurations, and $\Pr[\mathcal{E}] = 1.0$ in the OCS setting as predicted by Corollary 3.

6 Conclusion

We introduced LA-COCO, a framework that integrates one-step constraint predictions into the Lyapunov-based primal-dual paradigm for online convex optimization with adversarial constraints. Its core analytical device, the prediction-augmented regret decomposition (Lemma 3), extends the standard drift-plus-penalty analysis with bonus and penalty terms that quantify the value and cost of predictions in a single inequality. This formulation replaces the non-negative regret assumption of prior work with the strictly weaker prediction compensation event $\mathcal{E}$, recovering $\mathrm{CCV}_T = O(G^2 \log T/\alpha + \eta_T)$ under strongly convex costs when predictions are accurate. Alongside this, a Hedge meta-algorithm provides an unconditional $O(\sqrt{T})$ fallback regardless of prediction quality. The most pressing open question is to establish high-probability guarantees for $\mathcal{E}$ under natural stochastic or ergodic adversary models. Extending the improved CCV rate to the general convex setting, perhaps via optimistic FTRL or multi-step predictions, is another promising direction.

References

Keerti Anand, Rong Ge, and Debmalya Panigrahi. Customizing ML predictions for online algorithms. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 303–313, 2020.

Dheeraj Baby and Yu-Xiang Wang. Optimal dynamic regret in exp-concave online learning. In *Proceedings of the 35th Conference on Learning Theory*, pp. 2394–2442, 2022.

Aditya Bhaskara, Ashok Cutkosky, Ravi Kumar, and Manish Purohit. Online optimization with predictions and non-convex losses. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 869–879, 2020.

Zhiyi Cao, Alexander Lindermayr, and Nicole Megow. A survey on learning-augmented algorithms. *arXiv preprint arXiv:2407.13338*, 2024.

Niangjun Chen, Joshua Comden, Zhenhua Liu, Anshul Gandhi, and Adam Wierman. Using predictions in online optimization with switching costs: A fast algorithm and a fundamental limit. In *Proceedings of the American Control Conference*, pp. 3008–3013, 2018.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Annual Conference on Learning Theory*, pp. 6.1–6.20, 2012.

Nicolas Christianson, Tinashe Handina, and Adam Wierman. Optimal robustness-consistency tradeoffs for learning-augmented online algorithms. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, pp. 8042–8053, 2023.

Ashok Cutkosky. Combining online learning guarantees. In *Proceedings of the 32nd Conference on Learning Theory*, pp. 895–913, 2019.

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

Hao Guo, Xin Liu, Luo Chen, and Chenghao Liao. Online convex optimization with hard constraints: Towards the best of two worlds and beyond. *Advances in Neural Information Processing Systems*, 35, 2022.

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

Ali Jadbabaie, Alexander Rakhlin, Shahin Shahrampour, and Karthik Sridharan. Online optimization: Competing with dynamic comparators. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics*, pp. 398–406, 2015.

Rodolphe Jenatton, Jim Huang, and Cédric Archambeau. Adaptive algorithms for online convex optimization with long-term constraints. In *Proceedings of the 33rd International Conference on Machine Learning*, pp. 402–411, 2016.

Silvio Lattanzi, Thomas Lavastida, Benjamin Moseley, and Sergei Vassilvitskii. Online scheduling via learned weights. In *Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1859–1877, 2020.

Yingying Li, Guannan Li, and Guannan Qu. Online optimization with predictions and switching costs: Fast algorithms and the fundamental limit. In *IEEE Transactions on Automatic Control*, volume 67, pp. 5199–5213, 2021.

Nikolaos Liakopoulos, Apostolos Destounis, Georgios Paschos, Thrasyvoulos Spyropoulos, and Panayotis Mertikopoulos. Cautious regret minimization: Online optimization with long-term budget constraints. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 3944–3952, 2019.

Thodoris Lykouris and Sergei Vassilvitskii. Competitive caching with machine learned advice. In *Journal of the ACM*, volume 68, pp. 1–25, 2021.

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

Shie Mannor, John N Tsitsiklis, and Jia Yuan Yu. Online learning with knapsacks. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pp. 1–8, 2009.

Michael Mitzenmacher and Sergei Vassilvitskii. Algorithms with predictions. In *Communications of the ACM*, volume 65, pp. 33–35, 2022.

Mehryar Mohri and Scott Yang. Accelerating online convex optimization via adaptive prediction. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pp. 848–856, 2016.

Michael J Neely. Stochastic network optimization with application to communication and queueing systems. *Synthesis Lectures on Communication Networks*, 3(1):1–211, 2010.

Michael J Neely. Online convex optimization with time-varying constraints. *arXiv preprint arXiv:1702.04783*, 2017.

Manish Purohit, Zoya Svitkina, and Ravi Kumar. Improving online algorithms via ML predictions. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *Proceedings of the 26th Annual Conference on Learning Theory*, pp. 993–1019, 2013.

Dhruv Rohatgi. Near-optimal bounds for online caching with machine learned advice. In *Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1834–1845, 2020.

Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012.

Abhishek Sinha and Rahul Vaze. Optimal algorithms for online convex optimization with adversarial constraints. *Advances in Neural Information Processing Systems*, 37, 2024.

Wen Sun, Debadeepta Dey, and Ashish Kapoor. Safety-aware algorithms for adversarial contextual bandit. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 3280–3288, 2017.

Chen-Yu Wei, Haipeng Luo, and Alekh Agarwal. Taking a hint: How to leverage loss predictors in contextual bandits. In *Proceedings of the 33rd Conference on Learning Theory*, pp. 3583–3634, 2020.

Xinlei Yi, Xiuxian Li, Tao Yang, Lihua Xie, Tianyou Chai, and Karl H Johansson. Distributed online convex optimization with adversarial constraints. *IEEE Transactions on Automatic Control*, 68(5):2722–2737, 2023.

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

Jianjun Yuan and Andrew Lamperski. Online convex optimization for cumulative constraints. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning*, pp. 928–936, 2003.

A Proofs

A.1 Proof of Lemma 1 (Drift inequality)

Proof. Let $a = Q(t-1)$ and $b = Q(t) = a + \tilde{g}_t(x_t)$. Since $\tilde{g}_t(x_t) \geq 0$, we have $b \geq a$. By convexity of Φ, $\Phi(a) \geq \Phi(b) + \Phi'(b)(a-b)$, which rearranges to $\Phi(b) - \Phi(a) \leq \Phi'(b)(b-a) = \Phi'(Q(t))\,\tilde{g}_t(x_t)$. □

A.2 Proof of Lemma 2 (Gradient norm bound)

Proof. By the triangle inequality applied to the gradient of equation 3, $G_t \leq V\|\nabla \tilde{f}_t\| + \Phi'(Q(t))\|\nabla \tilde{g}_t\| + \mu_t \Phi'(Q(t))\|\nabla \tilde{g}_t^{\text{pred}}\|$. Each preprocessed function $\tilde{f}_t = \beta f_t$, $\tilde{g}_t = \beta(g_t)^+$, $\tilde{g}_t^{\text{pred}} = \beta(\hat{g}_{t+1})^+$ has subgradient norm at most βG (by Assumptions 1 and 4). The bound follows. □

A.3 Proof of Lemma 3 (Prediction-augmented regret decomposition)

Proof. For any round τ and feasible $x^\star$, expand the surrogate regret term:

$$\begin{aligned}\hat{f}_\tau(x_\tau) - \hat{f}_\tau(x^\star) = V(\tilde{f}_\tau(x_\tau) - \tilde{f}_\tau(x^\star)) + \Phi'(Q(\tau))(\tilde{g}_\tau(x_\tau) - \tilde{g}_\tau(x^\star)) \\ + \mu_\tau \Phi'(Q(\tau))(\tilde{g}_\tau^{\text{pred}}(x_\tau) - \tilde{g}_\tau^{\text{pred}}(x^\star)).\end{aligned}$$

Since $x^\star \in \mathcal{X}^\star$, we have $g_\tau(x^\star) \leq 0$, so $\tilde{g}_\tau(x^\star) = \beta(g_\tau(x^\star))^+ = 0$. For the prediction term, $x^\star \in \mathcal{X}^\star$ implies $g_{\tau+1}(x^\star) \leq 0$, hence $\hat{g}_{\tau+1}(x^\star) = \hat{g}_{\tau+1}(x^\star) - g_{\tau+1}(x^\star) + g_{\tau+1}(x^\star) \leq |\hat{g}_{\tau+1}(x^\star) - g_{\tau+1}(x^\star)| \leq \varepsilon_{\tau+1}$, and taking the positive part gives $(\hat{g}_{\tau+1}(x^\star))^+ \leq \varepsilon_{\tau+1}$, so $\tilde{g}_\tau^{\text{pred}}(x^\star) \leq \beta\varepsilon_{\tau+1}$. Therefore:

$$\begin{aligned}\hat{f}_\tau(x_\tau) - \hat{f}_\tau(x^\star) \geq V(\tilde{f}_\tau(x_\tau) - \tilde{f}_\tau(x^\star)) + \Phi'(Q(\tau))\tilde{g}_\tau(x_\tau) \\ + \mu_\tau \Phi'(Q(\tau))\tilde{g}_\tau^{\text{pred}}(x_\tau) - \mu_\tau \Phi'(Q(\tau))\beta\varepsilon_{\tau+1}.\end{aligned}$$

By Lemma 1, $\Phi'(Q(\tau))\tilde{g}_\tau(x_\tau) \geq \Phi(Q(\tau)) - \Phi(Q(\tau-1))$. Summing over $\tau = 1, \ldots, t$, the left-hand side becomes $\text{Regret}'_t(x^\star)$. On the right-hand side, the cost differences sum to $V\beta\,\text{Regret}_t(x^\star)$ (since $\tilde{f}_\tau = \beta f_\tau$), the drift terms telescope to $\Phi(Q(t)) - \Phi(0) = \Phi(Q(t))$, and the remaining terms give $\text{Bonus}_t - \text{Penalty}_t$. Rearranging yields equation 8. □

A.4 Proof of Lemma 4 (Bonus decomposition under quadratic Lyapunov)

Proof. Part (a) follows directly from $\Phi'(Q(\tau)) = 2Q(\tau)$ and $\tilde{g}_\tau^{\text{pred}}(x_\tau) = (\hat{g}_{\tau+1}(x_\tau))^+ = (g_{\tau+1}(x_\tau))^+$ under perfect prediction.

For part (b), the identity $2ab = (a+b)^2 - a^2 - b^2$ with $a = Q(\tau)$ and $b = \Delta_{\tau+1}$ gives $2Q(\tau)\Delta_{\tau+1} = (Q(\tau) + \Delta_{\tau+1})^2 - Q^2(\tau) - \Delta_{\tau+1}^2 = Q^2(\tau+1) - Q^2(\tau) - \Delta_{\tau+1}^2$. Summing over $\tau = 1, \ldots, t$ telescopes to $Q^2(t+1) - Q^2(1) - \sum_\tau \Delta_{\tau+1}^2$.

For part (c), $\Delta_{\tau+1} \leq GD$ implies $\sum_\tau \Delta_{\tau+1}^2 \leq GD \sum_\tau \Delta_{\tau+1} = GD(Q(t+1) - Q(1))$.

For part (d), by the non-expansiveness of the positive part and G-Lipschitz continuity of $g_{\tau+1}$, $(g_{\tau+1}(x_\tau))^+ \geq (g_{\tau+1}(x_{\tau+1}))^+ - G\|x_\tau - x_{\tau+1}\| = \Delta_{\tau+1} - G\|x_\tau - x_{\tau+1}\|$. Substituting into part (a) and applying parts (b)–(c) with $Q(\tau) \leq Q(T)$ yields equation 11. □

A.5 Proof of Lemma 5 (Path length bound)

Proof. For part (a), the projection non-expansiveness gives $\|x_{t+1} - x_t\| \leq \eta_t G_t \leq G_{\max}/(V\alpha t)$, and also $\|x_{t+1} - x_t\| \leq D$. Define $t^\star := \lfloor G_{\max}/(V\alpha D) \rfloor$. For $t \leq t^\star$, use the diameter bound D, contributing at most $t^\star D \leq G_{\max}/(V\alpha)$. For $t > t^\star$, use the step-size bound, contributing at most $(G_{\max}/(V\alpha)) \ln(T/t^\star)$. If $t^\star = 0$, the entire sum is bounded by $(G_{\max}/(V\alpha))(1 + \ln T)$. Combining gives equation 12.

For part (b), let $S_t = \sum_{s=1}^t G_s^2$. Then $\|x_{t+1} - x_t\| \leq DG_t/\sqrt{S_t}$. The auxiliary inequality $a/\sqrt{b+a^2} \leq 2(\sqrt{b+a^2} - \sqrt{b})$ (which follows from $\sqrt{b+a^2} - \sqrt{b} = a^2/(\sqrt{b+a^2} + \sqrt{b}) \geq a^2/(2\sqrt{b+a^2})$) with $a = G_t$ and $b = S_{t-1}$ gives $G_t/\sqrt{S_t} \leq 2(\sqrt{S_t} - \sqrt{S_{t-1}})$. Summing over $t = 1, \ldots, T$ telescopes to $2\sqrt{S_T}$. Multiplying by D and using $S_T \leq TG_{\max}^2$ yields equation 13. □

A.6 Proof of Lemma 6 (Strongly convex surrogate regret)

Proof. By σ-strong convexity of $\hat{f}_t$ with $\sigma = V\alpha$, $\hat{f}_t(x_t) - \hat{f}_t(x^\star) \le \nabla \hat{f}_t(x_t)^\top (x_t - x^\star) - (\sigma/2)\|x_t - x^\star\|^2$. The OGD update and projection non-expansiveness give $\nabla \hat{f}_t(x_t)^\top (x_t - x^\star) \le (\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2)/(2\eta_t) + \eta_t G_t^2/2$. Combining and writing $\delta_t := \|x_t - x^\star\|^2$:

$$\hat{f}_t(x_t) - \hat{f}_t(x^\star) \le \frac{1}{2}\left(\frac{1}{\eta_t} - \sigma\right)\delta_t - \frac{1}{2\eta_t}\delta_{t+1} + \frac{\eta_t G_t^2}{2}.$$

With $\eta_t = 1/(\sigma t)$, we have $1/\eta_t - \sigma = \sigma(t-1)$. Summing over $t = 1, \ldots, T$, the δ_t terms telescope: the $t = 1$ term contributes $\sigma \cdot 0 \cdot \delta_1/2 = 0$, and the final term $-\sigma T \delta_{T+1}/2 \le 0$ is discarded. The step-size terms sum to $\sum_{t=1}^{T} G_t^2/(2\sigma t)$. □

A.7 Proof of Lemma 7 (Hedge guarantees)

Proof. For part (i), convexity of g_t and $(\cdot)^+$ gives $(g_t(x_t))^+ \le \sum_i w_i(t)(g_t(x_t^{(i)}))^+$. Define $c_i(t) := (g_t(x_t^{(i)}))^+$ and $\ell_i(t) := c_i(t)/(GD) \in [0,1]$. Summing over t, $\mathrm{CCV}_T^{\mathrm{Hedge}} \le \sum_t \sum_i w_i(t) c_i(t) = GD \sum_t \sum_i w_i(t)\ell_i(t)$. The Hedge regret bound (Freund & Schapire, 1997) gives $\sum_t \sum_i w_i(t)\ell_i(t) \le \min_i \sum_t \ell_i(t) + 2\sqrt{T \ln N}$. Multiplying by GD yields equation 15.

For part (ii), convexity of f_t gives $f_t(x_t) \le \sum_i w_i(t) f_t(x_t^{(i)})$. Define $f_t^{\min} := \min_{x \in \mathcal{X}} f_t(x)$ and $\ell_i(t) := (f_t(x_t^{(i)}) - f_t^{\min})/(GD) \in [0,1]$. Then $\mathrm{Regret}_T^{\mathrm{Hedge}} \le GD \sum_t \sum_i w_i(t)\ell_i(t) - \sum_t (f_t(x^\star) - f_t^{\min})$. Applying the Hedge bound to $\ell_i(t)$ and noting that $GD \sum_t \ell_j(t) - \sum_t (f_t(x^\star) - f_t^{\min}) = \mathrm{Regret}_T^{(j)}(x^\star)$ (the $f_t^{\min}$ terms cancel exactly) gives equation 16. □

A.8 Proof of Lemma 8 (Baseline CCV bounds for sub-policy A)

Proof. For part (a), Lemma 3 with $\mu = 0$ (so Bonus = Penalty = 0) gives $Q^2(T) + V\mathrm{Regret}_T^{(A)} \le \mathrm{Regret}_T'^{(A)}$. By Lemma 2 with $\mu = 0$, $G_t^{(A)} \le G(V + 2Q^{(A)}(t))$. Expanding $(V + 2Q(t))^2 \le 2V^2 + 8Q^2(T)$ (using monotonicity of Q) and applying Lemma 6:

$$\mathrm{Regret}_T'^{(A)} \le \frac{G^2 \ln(Te)}{V\alpha}(V^2 + 4Q^2(T)).$$

Define $c := 4G^2 \ln(Te)/(V\alpha) = 1/16$. Then:

$$\frac{15}{16}Q^2(T) + V\mathrm{Regret}_T^{(A)} \le \frac{VG^2 \ln(Te)}{\alpha}.$$

Since $\mathrm{Regret}_T^{(A)}$ may be negative, we bound $|\mathrm{Regret}_T^{(A)}| \le GDT$ (each round contributes at most GD):

$$\frac{15}{16}Q^2(T) \le \frac{VG^2 \ln(Te)}{\alpha} + VGDT = \frac{64G^4 \ln^2(Te)}{\alpha^2} + \frac{64G^3 D \ln(Te)}{\alpha}T.$$

Taking square roots gives equation 17.

For part (b), Lemma 3 with $\mu = 0$, $V = 1$, $\Phi(x) = e^{\lambda x} - 1$ gives $e^{\lambda Q^{(A)}(T)} - 1 + \beta\mathrm{Regret}_T^{(A)} \le \mathrm{Regret}_T'^{(A)}$. By Lemma 2 with $\mu = 0$, $G_t^{(A)} \le (1/(2D))(1 + \lambda e^{\lambda Q^{(A)}(T)})$. The AdaGrad convex regret bound gives $\mathrm{Regret}_T'^{(A)} \le (\sqrt{T}/2)(1 + 2\lambda e^{\lambda Q^{(A)}(T)})$. Using $|\beta\mathrm{Regret}_T^{(A)}| \le T/2$:

$$e^{\lambda Q^{(A)}(T)}(1 - \lambda\sqrt{T}) \le T + 1.$$

Choosing $\lambda = 1/(2\sqrt{T})$ gives $e^{\lambda Q^{(A)}(T)} \le 2(T+1)$, hence $Q^{(A)}(T) \le 2\sqrt{T}\ln(2(T+1))$ and $\mathrm{CCV}_T^{(A)} = 2GD \cdot Q^{(A)}(T) = O(GD\sqrt{T}\log T)$. □

A.9 Proof of Lemma 9 (Quadratic inequality)

Proof. The inequality $aq^2 - 2\eta q - A \leq 0$ has discriminant $\Delta = 4\eta^2 + 4aA > 0$. The non-negative root satisfies $q \leq (\eta + \sqrt{\eta^2 + aA})/a$. By $\sqrt{\eta^2 + aA} \leq \eta + \sqrt{aA}$ (since $(\eta + \sqrt{aA})^2 = \eta^2 + aA + 2\eta\sqrt{aA} \geq \eta^2 + aA$), we obtain $q \leq 2\eta/a + \sqrt{A/a}$. □

A.10 Proof of Theorem 1 (Strongly convex setting)

Proof. By Lemma 2 with $\mu = 1$, $\beta = 1$, $\Phi'(Q) = 2Q$: $G_t^{(B)} \leq G(V + 4Q^{(B)}(t))$. Since $(V + 4Q)^2 \leq 2V^2 + 32Q^2(T)$ and $\sum_{t=1}^{T} 1/t \leq \ln(Te)$, Lemma 6 gives

$$\mathrm{Regret}_T'^{(B)} \leq \frac{VG^2 \ln(Te)}{\alpha} + \frac{1}{4}Q^2(T), \tag{27}$$

where $16G^2 \ln(Te)/(V\alpha) = 1/4$.

Lemma 3 with $\beta = 1$, $\mu = 1$ yields

$$Q^2(T) + V\mathrm{Regret}_T^{(B)} + \mathrm{Bonus}_T \leq \mathrm{Regret}_T'^{(B)} + \mathrm{Penalty}_T. \tag{28}$$

For the penalty, $Q(\tau) \leq Q(T)$ by monotonicity, so $\mathrm{Penalty}_T = \sum_{\tau=1}^{T} 2Q(\tau)\varepsilon_{\tau+1} \leq 2Q(T)\sum_{\tau=1}^{T} \varepsilon_{\tau+1}$. The index shift $\sum_{\tau=1}^{T} \varepsilon_{\tau+1} = \varepsilon_2 + \cdots + \varepsilon_{T+1}$ differs from $\eta_T = \varepsilon_1 + \cdots + \varepsilon_T$ by $\varepsilon_{T+1} - \varepsilon_1$, and each $\varepsilon_t \leq 2GD$, giving

$$\mathrm{Penalty}_T \leq 2Q(T)(\eta_T + 2GD). \tag{29}$$

Substituting equation 27 and equation 29 into equation 28:

$$\frac{3}{4}Q^2(T) + V\mathrm{Regret}_T^{(B)} + \mathrm{Bonus}_T \leq \frac{VG^2 \ln(Te)}{\alpha} + 2Q(T)\eta_T + 4GDQ(T). \tag{30}$$

For part (i), drop $\frac{3}{4}Q^2 \geq 0$ and Bonus ≥ 0 from the left side of equation 30 to obtain equation 20. Setting $\eta_T = 0$ in equation 20 gives $V\mathrm{Regret}_T^{(B)} \leq VG^2 \ln(Te)/\alpha + 4GDQ(T)$; dividing by V yields the regret bound stated in Remark 1.

For part (ii), assume $\mathcal{E}$, i.e., $V\mathrm{Regret}_T^{(B)} + \mathrm{Bonus}_T \geq 0$. Then equation 30 gives

$$\frac{3}{4}Q^2(T) \leq \frac{VG^2 \ln(Te)}{\alpha} + 2Q(T)(\eta_T + 2GD).$$

Apply Lemma 9 with $q = Q(T)$, $a = 3/4$, $A = VG^2 \ln(Te)/\alpha = 64G^4 \ln^2(Te)/\alpha^2$, $\eta = \eta_T + 2GD$:

$$Q(T) \leq \frac{8}{3}(\eta_T + 2GD) + \frac{2}{\sqrt{3}}\sqrt{A} = \frac{8}{3}(\eta_T + 2GD) + \frac{16G^2 \ln(Te)}{\sqrt{3}\,\alpha}.$$

Using $8/3 < 4$, $16/3 < 6$, $1/\sqrt{3} < 1$:

$$Q(T) \leq \frac{16G^2 \ln(Te)}{\alpha} + 4\eta_T + 6GD,$$

which is equation 21 since $\mathrm{CCV}_T^{(B)} = Q(T)$ when $\beta = 1$. Setting $\eta_T = 0$ gives $\mathrm{CCV}_T^{(B)} = O(G^2 \log T/\alpha)$ as noted in Remark 1.

For the refined regret bound equation 24, substitute the bound $Q(T) \leq 16G^2 \ln(Te)/\alpha + 4\eta_T + 6GD$ from equation 21 into equation 20 and divide by V:

$$\mathrm{Regret}_T^{(B)} \leq \frac{G^2 \ln(Te)}{\alpha} + \frac{2(\eta_T + 2GD)}{V}\left(\frac{16G^2 \ln(Te)}{\alpha} + 4\eta_T + 6GD\right) = O\left(\frac{G^2 \log T}{\alpha} + \frac{\eta_T}{\alpha} + \frac{\alpha\eta_T^2}{G^2 \log T}\right),$$

where the $O(GD)$ additive constants are absorbed by the leading terms as $T \to \infty$.

For parts (iii)–(iv), Lemma 7(i) gives $\mathrm{CCV}_T^{\mathrm{Hedge}} \leq \min(\mathrm{CCV}_T^{(A)}, \mathrm{CCV}_T^{(B)}) + O(GD\sqrt{T})$. Part (iii) follows from $\min \leq \mathrm{CCV}_T^{(A)}$ and Lemma 8(a). Part (iv) follows by taking the minimum of both branches when $\mathcal{E}$ holds. □

A.11 Proof of Theorem 2 (Convex setting)

Proof. By Lemma 2 with $\mu = 1$, $\beta = (2GD)^{-1}$: $G_t^{(B)} \leq (1/(2D))(1 + 2\lambda e^{\lambda Q(T)})$. The AdaGrad convex regret bound gives

$$\mathrm{Regret}_T'^{(B)} \leq \frac{\sqrt{T}}{2}(1 + 2\lambda e^{\lambda Q(T)}). \tag{31}$$

Lemma 3 with $V = 1$, $\beta = (2GD)^{-1}$, dropping Bonus ≥ 0:

$$e^{\lambda Q(T)} - 1 + \frac{1}{2GD}\mathrm{Regret}_T^{(B)} \leq \mathrm{Regret}_T'^{(B)} + \mathrm{Penalty}_T.$$

The penalty satisfies $\mathrm{Penalty}_T \leq \lambda e^{\lambda Q(T)} \cdot \beta\eta_T$. Substituting equation 31 and using $|(2GD)^{-1}\mathrm{Regret}_T^{(B)}| \leq T/2$:

$$e^{\lambda Q(T)}(1 - \lambda\sqrt{T} - \lambda\beta\eta_T) \leq T + 1.$$

Choosing $\lambda = 1/(2(\sqrt{T} + \beta\eta_T))$ makes $\lambda\sqrt{T} + \lambda\beta\eta_T = 1/2$, giving $e^{\lambda Q(T)} \leq 2(T + 1)$, hence

$$Q(T) \leq 2(\sqrt{T} + \beta\eta_T)\ln(2(T + 1)). \tag{32}$$

Since $\mathrm{CCV}_T^{(B)} = \beta^{-1}Q(T) = 2GD \cdot Q(T)$:

$$\mathrm{CCV}_T^{(B)} \leq 4GD\sqrt{T}\ln(2(T + 1)) + 2\eta_T \ln(2(T + 1)) = O((GD\sqrt{T} + \eta_T)\log T),$$

which is (iii).

For (i), sub-policy C runs standard OGD on f_t with step size $\eta = D/(G\sqrt{T})$, achieving $\mathrm{Regret}_T^{(C)} \leq GD\sqrt{T}$. By Lemma 7(ii), $\mathrm{Regret}_T^{\mathrm{Hedge}} \leq \min_i \mathrm{Regret}_T^{(i)} + O(GD\sqrt{T}) \leq GD\sqrt{T} + O(GD\sqrt{T}) = O(GD\sqrt{T})$.

For (ii) and (iv), Lemma 8(b) gives $\mathrm{CCV}_T^{(A)} = O(GD\sqrt{T}\log T)$ unconditionally. By Lemma 7(i), $\mathrm{CCV}_T^{\mathrm{Hedge}} \leq \min(\mathrm{CCV}_T^{(A)}, \mathrm{CCV}_T^{(B)}) + O(GD\sqrt{T})$. Since $\eta_T \geq 0$, the sub-policy A branch is never worse than sub-policy B, so (iv) simplifies to (ii). □

A.12 Proof of Theorem 3 (Sufficient conditions for $\mathcal{E}$)

Proof. For part (a), $V\mathrm{Regret}_T \geq 0$ and $\mathrm{Bonus}_T \geq 0$ immediately give $V\mathrm{Regret}_T + \mathrm{Bonus}_T \geq 0$.

For part (b), under $\eta_T = 0$ with $\Phi(x) = x^2$, $\mu_\tau = 1$, $\beta = 1$, Lemma 4(a) gives $\mathrm{Bonus}_T = \sum_{\tau=1}^{T} 2Q(\tau)(g_{\tau+1}(x_\tau))^+$. The hypothesis states $\sum_t Q(t)(g_{t+1}(x_t))^+ \geq (V/2)\sum_t |f_t(x_t) - f_t(x^\star)|$, so $\mathrm{Bonus}_T \geq V\sum_t |f_t(x_t) - f_t(x^\star)|$. Since $|f_t(x_t) - f_t(x^\star)| \geq f_t(x_t) - f_t(x^\star)$ for each t, summing gives $\sum_t |f_t(x_t) - f_t(x^\star)| \geq \mathrm{Regret}_T$. Moreover, $|a| \geq -a$ for all $a \in \mathbb{R}$, so $\mathrm{Bonus}_T \geq V|\mathrm{Regret}_T| \geq -V\mathrm{Regret}_T$, hence $V\mathrm{Regret}_T + \mathrm{Bonus}_T \geq 0$. □

A.13 Proof of Corollary 1 (Prediction quality threshold)

Proof. The first branch of equation 23 is smaller than the second when $16G^2\ln(Te)/\alpha + 4\eta_T + 6GD \leq O(\sqrt{G^3DT\log T/\alpha})$, which holds whenever $\eta_T \leq O(\sqrt{G^3DT\log T/\alpha})$. □

A.14 Proof of Proposition 1 (Online-realizable variant)

Proof. The time horizon $[1, T]$ is partitioned into at most $K+1$ phases by the doubling events, where phase k spans $[t_k, t_{k+1}-1]$ with length $T_k := t_{k+1} - t_k$ and accumulated prediction error $\eta^{(k)} := \sum_{s=t_k}^{t_{k+1}-1} \varepsilon_s \le \hat{\eta}_k$. Since CCV is additive, $\mathrm{CCV}_T^{(B)} = \sum_{k=0}^{K} \mathrm{CCV}_k$.

Within phase k, sub-policy B runs with $\lambda_k = 1/(2(\sqrt{T} + \beta\hat{\eta}_k))$ and $Q^{(B)}$ reset to zero. The analysis of Theorem 2 applies to each phase independently: the AdaGrad surrogate regret bound depends on the actual phase length $T_k \le T$, and the condition $\lambda_k\sqrt{T_k} + \lambda_k\beta\eta^{(k)} \le \lambda_k\sqrt{T} + \lambda_k\beta\hat{\eta}_k = 1/2$ is satisfied. This gives $e^{\lambda_k Q_k} \le 2(T_k+1) \le 2(T+1)$, hence

$$\mathrm{CCV}_k \le 4GD(\sqrt{T} + \beta\hat{\eta}_k)\ln(2(T+1)).$$

Summing over phases: the $\sqrt{T}$ terms contribute $(K+1)\cdot O(GD\sqrt{T}\log T)$, and the $\hat{\eta}_k$ terms contribute $O(GD\beta\sum_k \hat{\eta}_k \cdot \log T)$. The geometric doubling $\hat{\eta}_k = 2^k$ ensures $K \le \lceil \log_2 \eta_T \rceil + 1$ and $\sum_{k=0}^{K} \hat{\eta}_k = 2^{K+1} - 1 \le 4\eta_T$. Therefore:

$$\mathrm{CCV}_T^{(B)} \le O(\log \eta_T)\cdot O(GD\sqrt{T}\log T) + O(\eta_T \log T) = O((GD\sqrt{T}\log \eta_T + \eta_T)\log T).$$

The Hedge layer treats sub-policy B as a single expert throughout (its per-round output $x_t^{(B)} \in \mathcal{X}$ is well-defined regardless of internal restarts), so Lemma 7 applies over the full horizon. Sub-policy A requires no doubling (its $\lambda_A = 1/(2\sqrt{T})$ is independent of η_T), so (i)–(ii) are preserved. □

B Related Work

Online convex optimization with constraints. The classical OCO framework of Zinkevich (2003) and its subsequent refinements (Hazan, 2016; Shalev-Shwartz, 2012) focus on regret minimization without constraints. Constrained variants emerged from the observation that many practical problems impose time-varying feasibility requirements that cannot be folded into the action set $\mathcal{X}$. Mannor et al. (2009) initiated the study of online learning with knapsack-type constraints, establishing $O(\sqrt{T})$ regret and CCV bounds through primal-dual methods. Mahdavi et al. (2012) formalized the long-term constraint setting and showed that trading regret for efficiency yields $O(T^{3/4})$ CCV, later improved to $O(\sqrt{T})$ by Jenatton et al. (2016) via adaptive step sizes. Yu et al. (2020) achieved $O(1)$ constraint violation for stochastic long-term constraints by exploiting Slater-type feasibility conditions, and Yuan & Lamperski (2018) studied cumulative constraints with $O(\sqrt{T})$ regret-CCV tradeoffs.

The transition to fully adversarial, time-varying constraints—where g_t may change arbitrarily at each round—introduced fundamentally harder challenges. Neely (2017) adapted the Lyapunov drift-plus-penalty framework from stochastic network optimization (Neely, 2010) to this adversarial setting, achieving $O(\sqrt{T})$ CCV under convex costs. Liakopoulos et al. (2019) proposed cautious regret minimization with long-term budget constraints, and Yi et al. (2023) extended the adversarial constraint model to distributed multi-agent settings. Guo et al. (2022) studied hard constraints where the learner must satisfy $g_t(x_t) \le 0$ at every round, obtaining $O(\sqrt{T})$ regret under a cumulative Slater condition. The state-of-the-art results for the soft-constraint (CCV) metric were established by Sinha & Vaze (2024), who proved that under α-strongly convex costs, a Lyapunov-based primal-dual algorithm achieves $\mathrm{CCV}_T = O(\sqrt{T})$ unconditionally and $\mathrm{CCV}_T = O(\log T)$ when the algorithm's regret is non-negative. The logarithmic rate is a substantial improvement, yet the non-negative regret condition is not guaranteed against a fully adaptive adversary. Our work replaces the non-negative regret assumption with the strictly weaker prediction compensation event $\mathcal{E}$, while maintaining the same $O(\sqrt{T})$ unconditional fallback through a Hedge meta-algorithm.

Learning-augmented online algorithms. The learning-augmented (or "algorithms with predictions") paradigm was introduced by Lykouris & Vassilvitskii (2021) and Purohit et al. (2018), who showed that machine-learned predictions can improve the competitive ratio of online caching and ski rental algorithms while maintaining worst-case guarantees when predictions are inaccurate. This consistency-robustness trade-off has since become a central design principle in the field. Rohatgi (2020) obtained near-optimal bounds for

online caching with learned advice, and Lattanzi et al. (2020) applied learned weights to online scheduling. Anand et al. (2020) developed a general framework for customizing ML predictions for online algorithms, and Christianson et al. (2023) characterized the optimal robustness-consistency tradeoff for a broad class of problems. Comprehensive surveys of this rapidly growing area are provided by Mitzenmacher & Vassilvitskii (2022) and Cao et al. (2024).

In the context of online optimization, Chen et al. (2018) and Li et al. (2021) incorporated predictions into online optimization with switching costs, establishing fundamental limits and fast algorithms that interpolate between the with-prediction and without-prediction regimes. Bhaskara et al. (2020) studied online optimization with predictions under non-convex losses. These works focus exclusively on the cost (regret) dimension and do not address constraint violation. In contrast, our framework applies the learning-augmented paradigm to the constraint dimension: predictions of future constraint functions are used to reduce CCV, and the Hedge meta-algorithm provides the standard consistency-robustness guarantee by dynamically selecting between prediction-augmented and prediction-free sub-policies.

Optimistic and predictive methods in online convex optimization. A related but distinct line of work exploits predictability of the loss sequence to obtain tighter regret bounds in unconstrained OCO. Chiang et al. (2012) showed that when consecutive loss functions vary slowly (gradual variation), mirror descent achieves regret scaling with the path length of the loss sequence rather than $\sqrt{T}$. Rakhlin & Sridharan (2013) formalized optimistic online learning, where the learner receives a "hint" about the next loss function and achieves regret proportional to the cumulative prediction error. Mohri & Yang (2016) accelerated online convex optimization via adaptive prediction of gradients, and Jadbabaie et al. (2015) studied dynamic regret bounds that compete with time-varying comparators. Wei et al. (2020) extended the hint framework to contextual bandits, and Baby & Wang (2022) obtained optimal dynamic regret in exp-concave online learning. Cutkosky (2019) developed techniques for combining multiple online learning guarantees, which is conceptually related to our Hedge-based meta-algorithm.

These optimistic methods share with our approach the use of predictions to improve performance, but differ in two fundamental respects. First, they predict cost (loss) functions to reduce regret, whereas we predict constraint functions to reduce CCV—a qualitatively different objective that requires augmenting the Lyapunov drift analysis rather than the standard regret decomposition. Second, optimistic OCO methods operate in the unconstrained setting and do not provide any guarantee on constraint violation. Our framework is, to the best of our knowledge, the first to systematically integrate the learning-augmented paradigm into online convex optimization with adversarial constraints, addressing both the consistency-robustness tradeoff and the interplay between prediction quality and constraint violation.

C Experimental Setup

Adversary design. We use a rotating halfspace adversary that creates genuine cost-constraint tension. At each round t, the constraint function is $g_t(x) = a_t^\top x + b_t$, where the direction $a_t \in \mathbb{R}^d$ rotates at a controlled angular velocity ω with additive Gaussian perturbation. The offset $b_t \leq 0$ is chosen so that the origin $x^\star = 0$ is always feasible ($g_t(0) = b_t \leq 0$), but the constraint boundary nearly touches the origin, making violations easy to incur. The cost function is $f_t(x) = (\alpha/2)\|x - v_t\|^2$, where v_t is aligned with a_t—pushing the learner toward the constraint-violating direction. This creates a strong pull-push dynamic: the cost drives the learner toward v_t (which violates g_t), while the constraint pushes it back. The prediction oracle returns $\hat{g}_{t+1}(x) = a_{t+1}^\top x + b_{t+1}$ with optional additive noise of magnitude σ_{pred}, yielding per-round prediction error $\varepsilon_t \approx \sigma_{\text{pred}} D\sqrt{d}$.

We consider three adversary types. Under a stochastic adversary, the rotation angle receives i.i.d. Gaussian perturbations. In the adaptive setting, v_t is fixed in a direction orthogonal to the constraint rotation plane, creating strongly negative regret. The third configuration is online convex set chasing (OCS), where $f_t = 0$ for all t.

Algorithms. We compare five methods: Sub-policy A (prediction-free Lyapunov baseline, $\mu = 0$), Sub-policy B (prediction-augmented, $\mu = 1$), LA-COCO Hedge (the full meta-algorithm), Naive OGD (minimizes

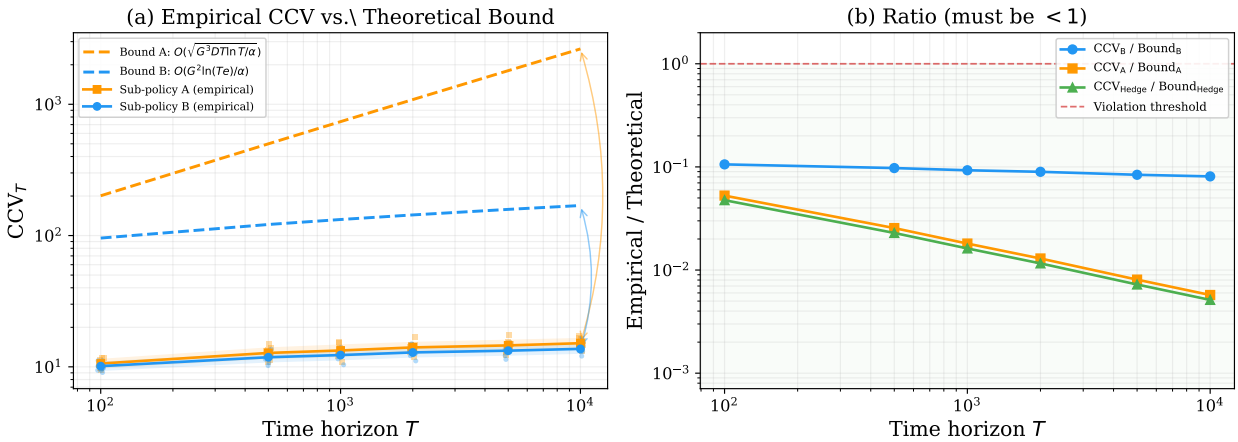

Figure 2: Theoretical upper bound verification ($d = 5$, $G = D = \alpha = 1$, $\eta_T = 0$, 10 seeds per T). (a) Empirical CCV vs. theoretical bounds from Lemma 8(a) and Theorem 1(ii). All 60 seed-level values lie below the bounds. (b) Ratio of empirical to theoretical CCV (must be < 1).

f_t only, ignores constraints), and Primal-Dual (standard constrained OCO with dual variable update). All methods use the ℓ_2-ball $\mathcal{X} = \{x \in \mathbb{R}^d : \|x\| \leq D\}$ as the action set.

Parameters. Unless otherwise stated, $d = 5$, $G = D = \alpha = 1$, rotation speed $\omega = 0.1$, and the strongly convex setting is used with $V = 64G^2 \ln(Te)/\alpha$, $\beta = 1$, $\Phi(x) = x^2$. Each configuration is repeated over multiple random seeds (5–20 depending on the experiment block), and we report means with standard deviations.

D Additional Experiments

D.1 Theoretical upper bound verification

A central claim of the paper is that the empirical CCV of each algorithm component lies strictly below the corresponding worst-case theoretical bound. Figure 2 validates this claim across $T \in \{100, 500, 1000, 2000, 5000, 10000\}$ with 10 seeds per configuration. Panel (a) overlays the empirical CCV of sub-policies A and B against their theoretical bounds from Lemma 8(a) and Theorem 1(ii). In every configuration and every seed, the empirical CCV lies strictly below the theoretical upper bound, confirming the correctness of the analysis. Panel (b) shows the ratio of empirical to theoretical CCV. The ratios range from 0.5% to 11%, indicating that the worst-case bounds are conservative—as expected for adversarial upper bounds evaluated on a specific (non-worst-case) adversary instance.

D.2 Regret decomposition verification

Lemma 3 is the core analytical tool of the paper. It asserts that $\Phi(Q(T)) + V\beta\,\mathrm{Regret}_T + \mathrm{Bonus}_T \leq \mathrm{Regret}'_T + \mathrm{Penalty}_T$, i.e., the gap RHS $-$ LHS ≥ 0 must hold for every run. Figure 3 verifies this across 12 configurations ($T \in \{100, 500, 1000, 5000\}$, $\sigma_{\mathrm{pred}} \in \{0, 0.1, 0.5\}$) with 10 seeds each. Panel (a) compares the LHS and RHS for the perfect-prediction case ($\sigma_{\mathrm{pred}} = 0$): the RHS exceeds the LHS for every T, with the gap increasing as T grows. Panel (b) shows the gap across all 120 configurations. The gap is non-negative in every case, with values ranging from near-zero (when $\sigma_{\mathrm{pred}} = 0$ and the Lyapunov drift inequality is nearly tight) to large positive values (when $\sigma_{\mathrm{pred}} > 0$ and the penalty term dominates). This confirms the mathematical validity of the decomposition.

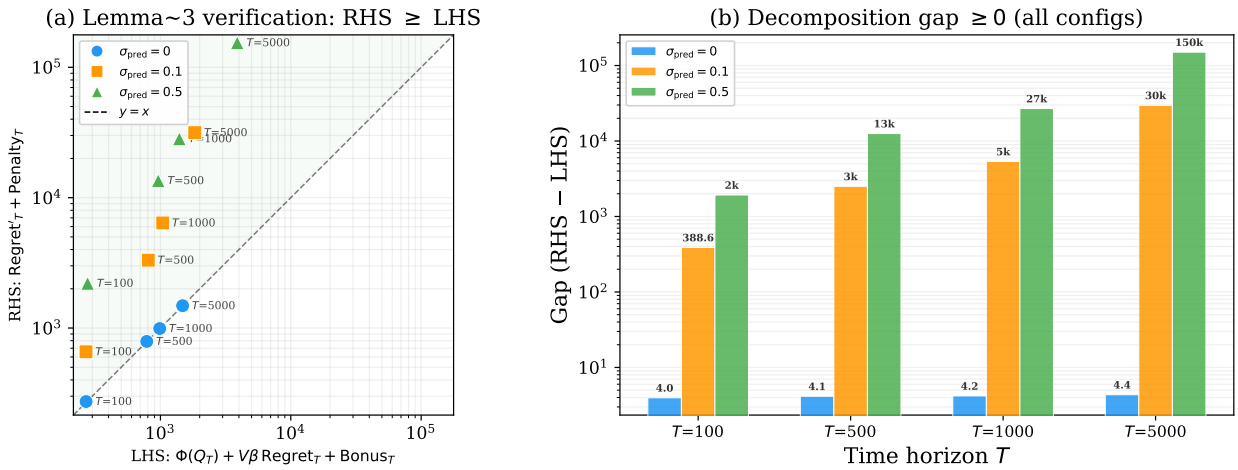

Figure 3: Lemma 3 verification. (a) LHS vs. RHS for $\sigma_{\text{pred}} = 0$. (b) Gap = RHS − LHS ≥ 0 across all 120 configurations (12 settings $\times$ 10 seeds per T).

D.3 Online convex set chasing

Corollary 3 predicts that when $f_t = 0$ for all t, the regret is exactly zero, so $\mathcal{E}$ holds unconditionally and the logarithmic CCV bound of Theorem 1(ii) applies without any posterior condition. Figure 4 validates this prediction. Panel (a) shows the empirical CCV of sub-policy B (with error bars over seeds) alongside the theoretical bound $16G^2 \ln(Te)/\alpha + 6GD$ as T ranges from 50 to 20,000; the empirical CCV lies far below the bound, and $\Pr[\mathcal{E}] = 1.0$ for every T, as predicted by Corollary 3. Panel (b) quantifies the tightness by plotting the ratio of empirical CCV to the theoretical bound: the ratio ranges from 0.2% to 4.5%, indicating that the worst-case bound is highly conservative in this benign setting.

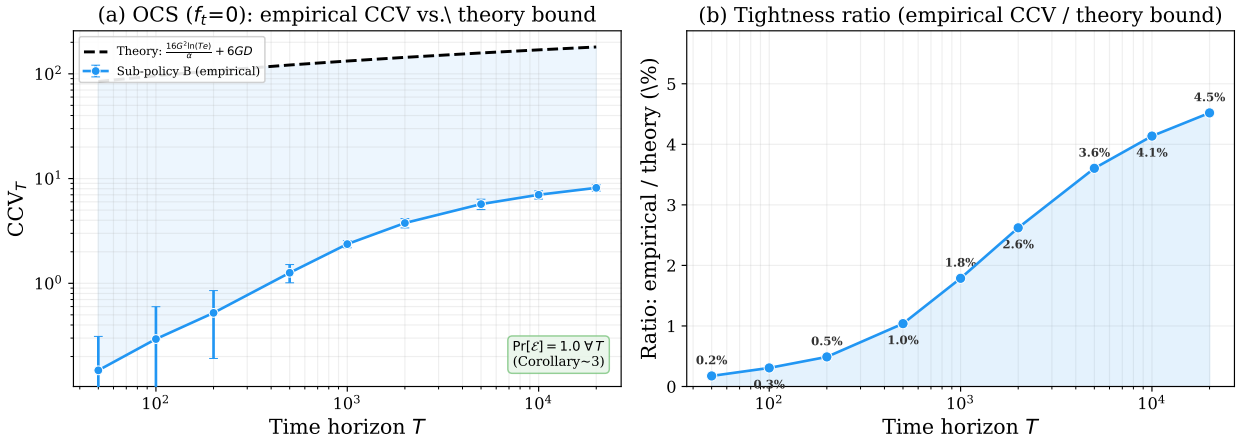

Figure 4: Online convex set chasing ($f_t = 0$, Corollary 3). (a) Empirical CCV (with error bars) vs. theoretical bound on log-log scale; $\Pr[\mathcal{E}] = 1.0$ for all T. (b) Ratio of empirical CCV to theoretical bound (%), showing the bound is conservative ($< 5\%$).