ONLINE CONDITIONAL ROBUST OPTIMIZATION WITH CONFORMAL ROBUSTNESS CONTROL

Anonymous authors
Paper under double-blind review

ABSTRACT

We study online conditional robust optimization, where decisions are made by minimizing worst-case loss over prediction sets before the true labels are observed. Existing online conformal prediction methods can ensure long-run decision robustness under arbitrary distribution shifts, but may yield overly conservative decisions because predictive coverage is sufficient but not necessary for robustness. To address this issue, we propose Online Conformal Robustness Control (OCRC), which formulates this task as an online constrained learning problem, minimizing average worst-case loss on past observations subject to an explicit robustness constraint. OCRC dynamically updates prediction-set thresholds directly from feedback on past robustness violations. We establish long-run robustness control for arbitrary data sequences and asymptotically exact marginal robustness under stochastic data settings. Experiments on synthetic and real-world data show that OCRC maintains the target robustness level while achieving substantially lower decision risks than coverage-based methods.

1 INTRODUCTION

Robust decision-making is central to risk-sensitive applications such as portfolio management (Goldfarb & Iyengar, 2003), energy storage (Yan et al., 2020), and radiotherapy treatment planning (Fredriksson et al., 2011). In these settings, decisions are made under uncertainty, and unfavorable outcomes can lead to substantial losses. Minimizing the expected loss alone may therefore be insufficient: decision-makers also need explicit control over the probability of incurring large losses. The central challenge is to provide such robustness guarantees while avoiding excessive conservatism that compromises decision efficiency (Bertsimas & Sim, 2004).

In the offline setting, conditional robust optimization (CRO) (Chenreddy et al., 2022) uses historical data to construct a prediction set $\mathcal{U}(X) \subseteq \mathcal{Y}$ for the unknown label $Y \in \mathcal{Y}$ given the covariate $X \in \mathcal{X}$. For a task-specific loss function $\phi(y, z)$, CRO selects a decision from a feasible set $\mathcal{Z}$ that minimizes the worst-case loss over $\mathcal{U}(X)$ and takes the resulting optimal value as its risk certificate:

$$z(x) = \arg\min_{z \in \mathcal{Z}} \max_{y \in \mathcal{U}(x)} \phi(y, z), \quad r(x) = \max_{y \in \mathcal{U}(x)} \phi(y, z(x)).$$

For a prescribed tolerance $\alpha \in (0, 1)$, the decision and certificate satisfy the robustness requirement if $\mathbb{P}\{\phi(Y, z(X)) \leq r(X)\} \geq 1 - \alpha$. Since the loss is bounded by the certificate whenever $Y \in \mathcal{U}(X)$, marginal predictive coverage $\mathbb{P}\{Y \in \mathcal{U}(X)\} \geq 1 - \alpha$ is sufficient for robustness. Conformal prediction (Vovk et al., 2005; Lei et al., 2018) provides such coverage guarantees in finite samples under exchangeability, motivating its use in CRO (Johnstone & Cox, 2021; Sun et al., 2023; Patel et al., 2024). Recent work further studies decision-aware calibration and direct robustness control (Kiyani et al., 2025; Hu et al., 2026), emphasizing that prediction-set design should account for downstream decision efficiency.

This paper focuses on the online CRO problem, where observations $\{(X_t, Y_t)\}_{t \geq 1}$ arrive sequentially and may exhibit temporal dependence or distribution shift. At time t, the decision-maker observes only the covariate X_t, then constructs a prediction set $\mathcal{U}_t(X_t)$ using the information currently available. After that, the CRO decision is made by $z_t(X_t) = \arg\min_{z \in \mathcal{Z}} \max_{y \in \mathcal{U}_t(X_t)} \phi(y, z)$, with the corresponding certificate $r_t(X_t) = \max_{y \in \mathcal{U}_t(X_t)} \phi(y, z_t(X_t))$ before observing Y_t. Because exchangeability need not hold in this setting, we seek guarantees that control the long-run frequency

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

of robustness violations. Adaptive conformal inference (ACI) (Gibbs & Candes, 2021) updates its calibration level using past coverage errors and achieves long-run coverage control under arbitrary distribution shifts. The same coverage-to-robustness implication then yields long-run robustness for the induced CRO decisions. However, coverage is sufficient but not necessary for robustness: an outcome outside the prediction set may still incur a loss below the associated certificate. Updating prediction sets in response to such noncoverage events can therefore lead to unnecessarily conservative decisions and certificates.

This motivates us to replace coverage with robustness as the quantity to be controlled online. Rather than requiring the outcome to fall inside a prediction set, we directly require the realized loss to stay below the certificate, so that outcomes that are harmless to the decision no longer inflate the certificate. Concretely, let the prediction set be parameterized by $\theta \in \Theta$, written as $\mathcal{U}_\theta(X_t)$, which in turn determines the decision $z_\theta(X_t)$ and the certificate $r_\theta(X_t)$. At each step we choose θ_t and aim to solve the online constrained problem

$$\min_{\theta_1,\ldots,\theta_T\in\Theta} \sum_{t=1}^{T} r_{\theta_t}(X_t) \quad \text{s.t.} \quad \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{\phi(Y_t, z_{\theta_t}(X_t)) > r_{\theta_t}(X_t)\} \le \alpha,$$

which keeps the long-run frequency of robustness violations below a prescribed level α while making the certificates as small as possible.

However, directly replacing coverage with robustness violation is not straightforward. Unlike the scalar threshold in ACI, the parameter θ can simultaneously affect the prediction set, the robust decision, and the resulting risk certificate. More importantly, the robustness condition set induced by θ, namely $\{y : \phi(y, z_\theta(X_t)) \le r_\theta(X_t)\}$, generally does not form a nested family as θ varies, regardless of whether the original prediction sets $\mathcal{U}_\theta(X_t)$ are nested. Consequently, there is no scalar ordering on θ to which the ACI-style coverage control mechanism can be directly applied.

We address this difficulty by introducing an external scalar threshold $q_t \ge 0$ into the robustness certificate. The resulting certificate becomes $r_{\theta_t}(X_t) + q_t$, while q_t serves as a dedicated scalar variable for controlling the realized robustness violation. It can therefore be updated in an ACI-style manner using the robustness feedback, while θ_t is simultaneously optimized to reduce the certificate. This decoupling leads to our *Online Conformal Robustness Control* (OCRC) framework, which jointly updates the prediction-set parameter and the external threshold through an online primal-dual algorithm.

We summarize our contributions as follows.

1. We propose Online Conformal Robustness Control (OCRC), which introduces an external scalar threshold q_t that decouples robustness control from the size and geometry of the prediction set, and jointly updates θ_t and q_t within a primal-dual online algorithm, thereby avoiding the instability of directly feeding robustness feedback into a multidimensional prediction-set parameter.
2. We provide theoretical guarantees that control the long-run frequency of robustness violations under arbitrary data, together with convergence analysis of the proposed algorithm in stationary stochastic environments.
3. We design experimental protocols and baselines to validate whether OCRC is less conservative than online coverage-based methods, covering both stationary and distribution-shift simulations as well as a real-world electricity load dataset.

2 Related Works

Uncertainty quantification of decision-making. A central question in decision-making under uncertainty is how to quantify the uncertainty of predictions so that the resulting decisions remain reliable. Conformal risk control addresses this by generalizing conformal prediction from coverage to the expected value of a general loss, calibrating a parameter so that the expected loss is bounded by a prescribed level with finite-sample, distribution-free guarantees (Angelopoulos et al., 2024; Bates et al., 2021). Learn-then-Test extends this to non-monotone losses through multiple hypothesis testing (Angelopoulos et al., 2021), and conformal decision theory applies conformal calibration

to decisions parameterized by a single scalar (Lekeufack et al., 2024), with applications to safe planning (Lindemann et al., 2023). From a decision-theoretic perspective, Kiyani et al. (2025) prove that prediction sets are optimal for risk-averse decision makers who optimize value at risk, and that the optimal sets need not be nested, which complements our observation that the robustness condition set induced by θ is generally non-nested. In the online setting, Rolling RC (Feldman et al., 2023) extends risk control to arbitrary data streams by replacing exchangeability with a long-run risk guarantee, but it still controls risks defined on the prediction set itself. Our work instead controls the downstream robustness violation, a different object that cannot be handled by the scalar mechanism used in these methods.

Robust optimization. Classical robust optimization optimizes against a deterministic uncertainty set, providing worst-case guarantees without distributional assumptions (Ben-Tal et al., 2009), but can be overly conservative when contextual information is available. Conditional Robust Optimization (CRO) addresses this by allowing the uncertainty set to vary with observed covariates (Chenreddy et al., 2022). In parallel, conformal prediction provides finite-sample, distribution-free validity for constructing prediction sets, and several works bridge the two in the batch setting: Johnstone & Cox (2021) translate conformal prediction regions into ellipsoidal uncertainty sets, Sun et al. (2023) propose a predict-then-calibrate paradigm that decouples prediction model training from uncertainty quantification, and Patel et al. (2024) construct nonconvex prediction regions via conditional generative models. More recently, Hu et al. (2026) observe that coverage is sufficient but not necessary for robustness, and propose Conformal Robustness Control (CRC), which directly constrains downstream robustness instead of coverage. These methods all operate under exchangeability and produce a fixed decision rule after calibration, whereas our setting requires decisions to be made sequentially under non-stationary data.

Online conformal prediction. For sequential data, the exchangeability assumption is often violated. Gibbs & Candes (2021) propose Adaptive Conformal Inference (ACI), which updates a scalar threshold via online gradient descent and achieves long-run coverage under arbitrary distribution shift. Subsequent work addresses ACI's step-size sensitivity through different mechanisms: Gibbs et al. (2024) use exponential reweighting, Zaffran et al. (2022) employ Bernstein Online Aggregation, and Bhatnagar et al. (2023) adopt strongly adaptive online learning. Podkopaev et al. (2024) extend the coin-betting framework to avoid step-size tuning entirely. Xu & Xie (2021) and Xu & Xie (2023) predict future residual quantiles but rely on quantile estimation assumptions. Angelopoulos et al. (2023) recast the problem through control theory as Conformal PID, while Weinstein & Ramdas (2020) and Bao et al. (2025) study selective conformal inference. Despite this diversity, most methods adapt a scalar threshold indexing a nested family, whereas decision policies in our setting are parameterized by multiple dimensions.

3 Online Conformal Robustness Control

3.1 Problem Setup

We consider a sequential data stream

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_T, Y_T) \in \mathcal{X} \times \mathcal{Y},$$

which may be non-i.i.d., dependent, non-stationary, or even adversarial. At time t, the decision maker observes X_t, chooses $\theta_t \in \Theta$ based only on past information, makes a robust decision, and then observes Y_t. Let $\mathcal{F}_{t-1}$ denote the information available before observing Y_t, and assume that θ_t is $\mathcal{F}_{t-1}$-measurable.

Under the CRO framework, let $\mathcal{U}_\theta(x) \subseteq \mathcal{Y}$ be a parameterized prediction set. The corresponding robust decision and robustness certificate are defined as

$$z_\theta(x) = \arg\min_{z \in \mathcal{Z}} \max_{y \in \mathcal{U}_\theta(x)} \phi(y, z), \qquad r_\theta(x) = \max_{y \in \mathcal{U}_\theta(x)} \phi(y, z_\theta(x)). \tag{1}$$

After observing Y_t, define the *robustness margin* $V_t(\theta) = \phi(Y_t, z_\theta(X_t)) - r_\theta(X_t)$. The decision at time t is robust if $V_t(\theta_t) \leq 0$. Our goal is to control the long-run frequency of robustness violations,

$$\frac{1}{T} \sum_{t=1}^{T} \mathbf{1}\{V_t(\theta_t) > 0\} \leq \alpha + o(1),$$

while keeping the cumulative certificate $\sum_{t=1}^{T} r_{\theta_t}(X_t)$ as small as possible.

A natural attempt is to mimic ACI and replace the coverage indicator in the pinball loss with the robustness violation indicator, giving

$$\ell_t(\theta) = V_t(\theta)\big(\mathbf{1}\{V_t(\theta) > 0\} - \alpha\big),$$

its gradient equilibrium condition (Angelopoulos et al., 2025) is

$$\sum_{t=1}^{T} \big(\mathbf{1}\{V_t(\theta) > 0\} - \alpha\big)\nabla_\theta V_t(\theta) = 0.$$

Ignoring the factor $\nabla_\theta V_t(\theta)$, this is exactly the robustness control we need. However, this factor couples the robustness feedback with the geometry of the parameterized decision problem: unlike the scalar threshold in ACI, θ simultaneously controls the prediction set, the robust decision, and the certificate, so the induced robustness condition set $\{y : V_t(\theta) \le 0\}$ is generally multidimensional and non-nested in θ. There is thus no scalar ordering on θ to which the ACI-style mechanism can be directly applied, and the realized feedback need not drive the system toward equilibrium. This motivates the external threshold introduced in the next subsection, which separates robustness control from the prediction-set parameterization.

3.2 Adding an external threshold

We instead opt to place robustness control as our primary concern. By adding an external threshold in the robustness margin, it can be controlled as a direct result of ACI. We reformulate the following notation

$$V_t(\theta, q) = \phi(Y_t, z_\theta(X_t)) - r_\theta(X_t) - q, \quad b_t(\theta, q) = r_\theta(X_t) + q \tag{2}$$

where $b_t(\theta, q)$ acts as the new robustness certificate, $q \in \mathbb{R}_{\ge 0}$ is an external threshold, updated by

$$q_{t+1} = \Pi_{[0,+\infty)}\{q_t + \gamma(H_\sigma(V_t(\theta)) - \alpha)\},$$

which is designed to control the frequency of robustness violations. We derive the algorithm from an optimization perspective. Denote

$$f_t(\theta) = r_\theta(X_t), \quad g_t(\theta, q) = H_\sigma(\phi(Y_t, z_\theta(X_t)) - r_\theta(X_t) - q) - \alpha.$$

By constructing a loss function associated with θ and q

$$\mathcal{L}_t(\theta, q) = f_t(\theta) + \kappa \int_0^q g_t(\theta, q')\mathrm{d}q', \tag{3}$$

the update (2) can be seen as a projected gradient ascent step on L_t with respect to q. We therefore intentionally formulate an online primal-dual algorithm based on $\mathcal{L}_t$.

Algorithm 1: Primal-dual OCRC with external threshold

Input: Robustness level $1-\alpha$, parameter θ_t and threshold q_t from last step, learning rates η_θ, η_q, smoothing parameter σ, coefficient κ.

1 Observe covariate X_t.
2 Compute $z_t = z_{\theta_t}(X_t)$ and $r_t = r_{\theta_t}(X_t)$ as in (1).
3 Deploy decision z_t and certificate $r_t + q_t$. Observe outcome Y_t.
4 Compute $V_t = V_t(\theta_t, q_t)$ as in (2). Compute the smooth indicator $H_t = H_\sigma(V_t(\theta_t, q_t))$.
5 $q_{t+1} \leftarrow \Pi_{[0,+\infty)}\{q_t + \gamma(H_t - \alpha)\}$.
6 $\theta_{t+1} \leftarrow \Pi_\Theta\{\theta_t - \eta_\theta \nabla_\theta \mathcal{L}_t(\theta_t, q_t)\}$ with $\mathcal{L}_t$ defined in (3).

Output: q_{t+1}, θ_{t+1}.

Remark 1. *The derivative of $\mathcal{L}_t$ with respect to q should be $\kappa(H_t - \alpha)$, but we still retain the ACI-style update for q, treating $\eta_q = \gamma/\kappa$ as the true step-size in the following primal-dual analysis. At the computational level, we do not need to compute the integral term in $\mathcal{L}_t$. Let G_σ be an antiderivative of H_σ. Then*

$$\begin{aligned}
\nabla_\theta \mathcal{L}_t(\theta, q) &= \nabla_\theta\{f_t(\theta) + \kappa[G_\sigma(V_t(\theta, 0)) - G_\sigma(V_t(\theta, q)) - \alpha q]\} \\
&= \nabla_\theta f_t(\theta) + \kappa[H_\sigma(V_t(\theta, 0))\nabla_\theta V_t(\theta, 0) - H_\sigma(V_t(\theta, q))\nabla_\theta V_t(\theta, q)] \\
&= \nabla_\theta f_t(\theta) + \kappa q \nabla_\theta\{H_\sigma(V_t(\theta, q)) - \alpha\} + o(\kappa q).
\end{aligned}$$

The second line is used for practical computation and theoretical proof. The third line can illustrate the practical role of q in the algorithm: both to control robustness and to act as a Lagrange multiplier of a constrained optimization problem.

Next, we show that the introduction of q_t does not destroy the structure of the original problem. Note that Algorithm 1 is primarily aimed at tracking the saddle point of $\mathcal{L}_t$. Consider the hindsight best fixed decision of the original problem at time T,

$$\min_{\theta\in\Theta} \sum_{t=1}^{T} f_t(\theta) \quad \text{s.t.} \quad \sum_{t=1}^{T} g_t(\theta, 0) \le 0. \tag{4}$$

We summarize the relationship between the solutions of the two problems in the following theorem.

Theorem 1. *Assume that Θ is compact, $f : \Theta \to \mathbb{R}$ and $g : \Theta \times [0, \infty) \to \mathbb{R}$ are continuous, and that the original feasible region $D = \{\theta \in \Theta : g(\theta, 0) \le 0\} \neq \emptyset$.*

Let θ^* and f^* denote the optimal solution and the optimal function value of the problem

$$\min_{\theta\in\Theta} f(\theta) \quad \textit{s.t.} \quad g(\theta, 0) \le 0.$$

Let $(\theta^*_\kappa, q^*_\kappa)$ and f^*_κ be the solution and the value of

$$\min_{\theta\in\Theta} \max_{q\in\mathbb{R}_{\ge 0}} \left\{ f(\theta) + \kappa \int_0^q g(\theta, q')\mathrm{d}q' \right\}.$$

If $g(\theta, q)$ is monotonically decreasing with respect to q, then

$$f^*_\kappa \to f^*,\ \theta^*_\kappa \to \theta^*,\ q^*_\kappa \to 0, \quad \textit{as} \quad \kappa \to \infty.$$

Replace f, g by $\sum_{t=1}^T f_t, \sum_{t=1}^T g_t$, to find that the problem (4) is equivalent to the mini-max problem of $\sum_{t=1}^T \mathcal{L}_t$ in the sense of $\kappa \to \infty$, while Algorithm 1 is solving the latter.

Remark 2. *Theorem 1 compares the* hindsight *optima of the two problems: it states that the minimax solution $(\theta^*_\kappa, q^*_\kappa)$ of the penalized problem converges to the optimal solution θ^* of the original constrained problem as $\kappa \to \infty$. It does not by itself relate the cumulative losses incurred by the online iterates of Algorithm 1, which are affected by the dynamics of the updates and by the data stream. This comparison is nevertheless the relevant one, because in online learning the natural benchmark is exactly the hindsight fixed decision: the regret of an online algorithm is defined against the best fixed θ chosen with full knowledge of the sequence, i.e., the hindsight optimum of the constrained problem. Establishing that the minimax problem solved by Algorithm 1 has the same hindsight optimum as the original constrained problem is therefore a prerequisite for any meaningful regret guarantee. Moreover, under the stationary stochastic setting considered in Section 4, the hindsight optimum coincides with the population-level target θ^* that the iterates are shown to approach, so the equivalence established here directly supports the convergence result in Theorem 4.*

4 Theoretical Results

We present theoretical results for Algorithm 1 in this section. We start with several mild assumptions.

Assumption 1. *$\mathcal{X}$ and $\mathcal{Y}$ are close and bounded. Θ is compact and bounded. ϕ is bounded by B_ϕ.*

Assumption 2. *$\phi(y, z)$ is L_ϕ-Lipschitz and s_ϕ-smooth in z for any $y \in \mathcal{Y}$. $z_\theta(x)$ is L_z-Lipschitz and s_z-smooth in θ for any x. $r_\theta(x)$ is L_r-Lipschitz and s_r-smooth in θ for any x.*

Assumption 3. *$H_\sigma(u) = H(u/\sigma)$. $H : \mathbb{R} \to [0, 1]$ is a smooth approximation of the indicator. It should be increasing, L_H-Lipschitz and s_H-smooth.*

$$\lim_{u\to-\infty} H(u) = 0, \quad \lim_{u\to+\infty} H(u) = 1.$$

Therefore, its inverse function $H^{-1}_\sigma(\alpha) = \sigma H^{-1}(\alpha)$ exists.

Under Assumptions 2 and 3, f_t is L_r- Lipschitz and s_r-smooth; g_t is $L_H(L_\phi L_z + L_r)/\sigma$-Lipschitz and smooth with parameter $s_H(L_\phi L_z + L_r)^2/\sigma^2 + L_H(L_\phi L_z + L_r)(s_\phi L_z + L_\phi s_z + s_r)/\sigma$.

The next theorem provides robustness control under arbitrary data. It is a direct generalization of the ACI (Gibbs & Candes, 2021) coverage result.

Theorem 2. *Let $\{\theta_t\}_{t=1}^T$ and $\{q_t\}_{t=1}^T$ be the sequences generated by Algorithm 1. Set $q_1 = 0$. Under Assumptions 1-3, q_t are bounded by $2B_\phi - H_\sigma^{-1}(\alpha) + \gamma \max\{\alpha, 1-\alpha\}$, thus*

$$\sum_{t=1}^{T} g_t(\theta_t, q_t) \leq \frac{2B_\phi - H_\sigma^{-1}(\alpha) + \gamma \max\{\alpha, 1-\alpha\}}{\gamma}.$$

Note that $H_\sigma^{-1}(\alpha) = \sigma H^{-1}(\alpha) = O(\sigma)$. Setting $\sigma = o(1)$, consequently,

$$\left[\sum_{t=1}^{T} g_t(\theta_t, q_t)\right]_+ = O(\gamma^{-1}).$$

Next, we discuss the efficiency of the algorithm, namely the convergence of θ. Although this is impossible under completely arbitrary data, we have considered efficiency under convexity and stationarity assumptions. Our subsequent theorems will be based on the following setting by default.

$$\sigma = \Theta(T^{-b}),\ \kappa = \Theta(T^a),\ \eta_\theta = \Theta(T^{-\frac{5a+8b+1}{2}}),\ \eta_q = \Theta(T^{-\frac{1-a}{4}})$$

where $a > 0, b > 0$ with $9a + 8b < 1$.

Theorem 3. *Assume that $\mathcal{L}_t(\theta, q)$ is convex in θ, and that Assumptions 1-3 holds. Under the default setting, the saddle regret be bounded as follows:*

$$R_T^{\text{sad}} = \sup_{\theta,q} \sum_{t=1}^{T} \{\mathcal{L}_t(\theta_t, q) - \mathcal{L}_t(\theta, q_t)\} = O(T^{\frac{9a+8b+3}{4}}).$$

Remark 3. *Sublinear saddle regret is the standard criterion for successful online learning in convex-concave games. It certifies that the ergodic average of the primal-dual iterates forms an asymptotic Nash equilibrium of the averaged game, with the duality gap vanishing at rate $O(R_{\text{sad}}/T)$. In this sense, it provides a principled guarantee that the online dynamics have effectively solved the min-max problem in the long-run average—even though the environment may vary over time.*

Although the convexity assumption relies on the structure of the prediction set U_θ and is nontrivial to construct and verify, we still adopt this theorem, which is a classical result in OCO, to demonstrate the performance of our algorithm in the online regime. To discuss the efficiency of the algorithm for nonconvex loss, we also consider the following stationary stochastic case.

Assumption 4. *$(X_t, Y_t) \sim P_t$, such that expectation of the loss function is time-invariant.*

$$\mathbb{E}[\mathcal{L}_t(\theta, q)|\mathcal{F}_{t-1}] = \ell(\theta, q).$$

Define

$$\xi_t^\theta(\theta, q) = \nabla_\theta \mathcal{L}_t(\theta, q) - \mathbb{E}[\nabla_\theta \mathcal{L}_t(\theta, q)|\mathcal{F}_{t-1}],$$
$$\xi_t^q(\theta, q) = \nabla_q \mathcal{L}_t(\theta, q) - \mathbb{E}[\nabla_q \mathcal{L}_t(\theta, q)|\mathcal{F}_{t-1}].$$

Assume bounded variance

$$\mathbb{E}\left[\|\xi_t^\theta\|^2 + |\xi_t^q|^2\right] \leq \sigma_\xi^2.$$

This assumption is mainly aimed at i.i.d. data. Although our algorithm is designed for distribution drift settings, the following theorem characterizes its convergence rate in stationary environments.

Theorem 4. *Assume that Assumption 1-4 holds and that $E[g_t(\theta_1, 0)] \leq 0$. Define $\Phi(\theta) = \max_q \ell(\theta, q)$. Under the default setting,*

$$\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\|\nabla \Phi(\theta_t)\|^2\right] = O(T^{\frac{5a+8b-1}{8}}).$$

In a stationary stochastic setting, Algorithm 1 becomes a kind of Stochastic Gradient Descent Ascent (SGDA). This result follows from Lin et al. (2020), a notable work in SGDA.

Theorem 2 controls the smoothed robustness violation. We next translate it to hard event $\mathbf{1}\{V_t(\theta_t, q_t) > 0\}$.

Assumption 5. *There exists $\rho > 0$ such that for all $\epsilon > 0$,*

$$\sup_{(\theta_t, q_t)_{t\geq 1}} \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{|V_t(\theta_t, q_t)| \leq \epsilon\} \leq \rho\epsilon.$$

Theorem 5. *Suppose Assumption 5 holds. Assume that H_σ satisfies, for constants C_H, c_H and approximation error ϵ_σ,*

$$\mathbf{1}\{u > 0\} \leq H_\sigma(u) + C_H \mathbf{1}\{|u| \leq c_H\sigma\} + \epsilon_\sigma \quad \textit{for all } u \in \mathbb{R}.$$

Then

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{V_t(\theta_t, q_t) > 0\} \leq \alpha + \frac{1}{T}\sum_{t=1}^{T} g_t(\theta_t, q_t) + O(\sigma) + O(\epsilon_\sigma).$$

5 Experiments

In this section, we provide some experimental results of OCRC, compared with coverage-based methods. The baseline method s maintains prediction sets with long-run coverage $1-\alpha$, following the principle described in Section 1. The source code is included in the supplementary material.

Regarding the specific settings, we adopt the commonly used ellipsoidal prediction set for $\mathcal{X} \subseteq \mathbb{R}^m, \mathcal{Y} \subseteq \mathbb{R}^n$ (Sun et al., 2023)

$$U_\theta(x) = \{y \in \mathbb{R}^n : (y - \mu_\theta(x))^\top \Sigma_\theta^{-1}(x)(y - \mu_\theta(x)) \leq 1\}.$$

In the OCRC method, the parameterization of the μ_θ and Σ_θ can take various forms, such as regression, neural networks, etc. For simplicity, however, we adopt the following relatively simple form in our simulation experiments.

$$\theta = (B, A_0, A_1, \ldots, A_m), \quad B \in \mathbb{R}^{n\times m}, A_d \in \mathbb{R}^{n\times n}, d = 0, 1, \ldots, m;$$
$$\mu_\theta(x) = Bx, \quad L_\theta(x) = A_0 + \sum_{d=1}^{m} x_d A_d, \quad \Sigma_\theta(x) = L_\theta(x) L_\theta^\top(x).$$

For coverage-based methods, however, controlling the coverage rate via online conformal methods typically requires the prediction sets to be nested with respect to a scalar. We therefore keep $\mu_0(x)$ and $\Sigma_0(x)$ the same as the oracle exact conditional mean and covariance for Y_t given $X_t = x$. Let

$$\theta = \tau \in \mathbb{R}, \quad \mu_\theta(x) = \mu_0(x), \quad \Sigma_\theta(x) = \tau\Sigma_0(x).$$

The $1-\alpha$ coverage of the prediction set is maintained by tuning τ online. In particular, we compared ACI (Gibbs & Candes, 2021) and Conformal PID (Angelopoulos et al., 2023).

5.1 Stationary data simulation

We first conduct simulation experiments under stationary environments, where Theorem 4 is applicable, to test the basic performance of the algorithm. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$ and $\mathcal{Z} = \{z \in [0,1]^2 : \|z\|_1 = 1\}$. Define the loss function $\phi(y, z) = -y^\top z$. The data $\{(X_t, Y_t)\}_{t=1}^T$ arrive sequentially but are i.i.d. from P_{XY}, where

$$X \sim N((1,1)^\top, 2.25 \cdot I_2), \quad Y = \begin{pmatrix} -5 & -2 \\ -3 & -1 \end{pmatrix} X + e, \quad e \sim N((0,0)^\top, I_2).$$

Although the data seem unbounded by the setting, in practice the vast majority of observations do not deviate significantly from the center. Similar settings can be found in Hu et al. (2026).

Result 1. The robustness and risk certificate of both algorithms stabilize after a certain period of time. It can be seen that Online CRO, based on a coverage rate of $1-\alpha$, tends to be overly conservative, whereas OCRC is able to maintain the robustness violation rate closer to $1-\alpha$, thereby achieving a lower risk certificate. Note that the coverage of OCRC does not stabilize but instead decreases slowly, which is because OCRC's mechanism for controlling robustness is decoupled from the size of prediction set. It suffices that the prediction set includes the predictions under appropriate pessimistic scenarios, such that the corresponding decision loss exceeds that under the $1-\alpha$ cases.

Simulations without stationarity can be found in Appendix B.1.

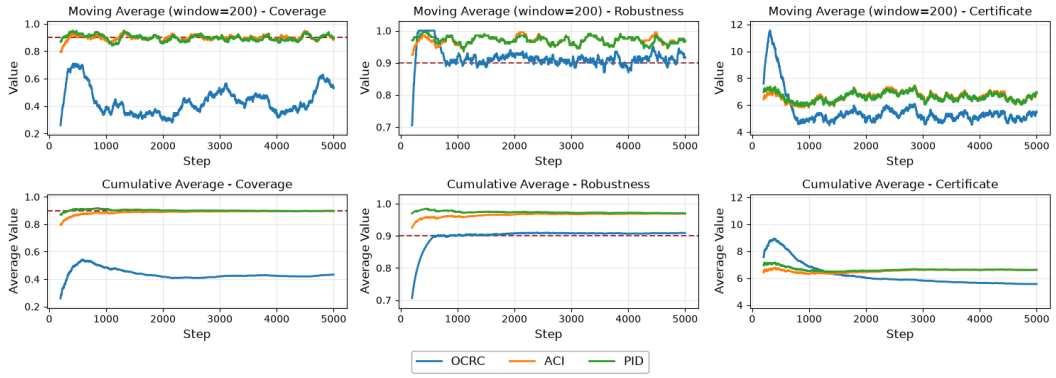

Figure 1: Coverage, robustness, and risk certificate under $T = 5000$ and $\alpha = 0.1$. Top: moving averages (window 200); bottom: cumulative averages. The dashed line denotes $1 - \alpha$.

5.2 Real-data: Load-Climate-Open

We conducted experiments on the real-world dataset to validate the performance of the algorithm. **Load-Climate-Open** is an open-access dataset designed for day-ahead electricity load forecasting. It covers the period from 2015 to 2019 at an hourly resolution and provides regional load data alongside climate and renewable generation features. It can be downloaded from `https://zenodo.org/records/20733833`.

Denote the hourly regional electricity load as the label $\mathbf{Y}$, and the 13 features provided in the dataset as $\mathbf{X}$. We define a decision problem similar to that in the preceding simulation. Suppose that the electricity cost is functional depended on the electricity load, and that a factory must meet a fixed total electricity consumption over a period of m hours. Let $z \in \mathbb{R}^m_{\geq 0}$, $\|z\|_1 = 1$ denote the proportion of electricity consumed by the factory in each hour over one period. We organize the data from m consecutive hours into the covariates $X_t \in \mathbb{R}^{m\times 13}$ and outcome $Y_t \in \mathbb{R}^m$ of one period. The loss is defined as $\phi(y, z) = h(y)^\top z$. In the real-data experiments, we adopt the following parameterization of the prediction set.

$$\theta = (\vartheta, A_0), \quad \mu_\theta(x) = \mathrm{MLP}_\vartheta(x), \quad \Sigma_\theta = A_0 A_0^\top,$$

where MLP_ϑ is a multilayer neural network with hidden layers of dimensions 128 and 64; A_0 is a lower triangular matrix.

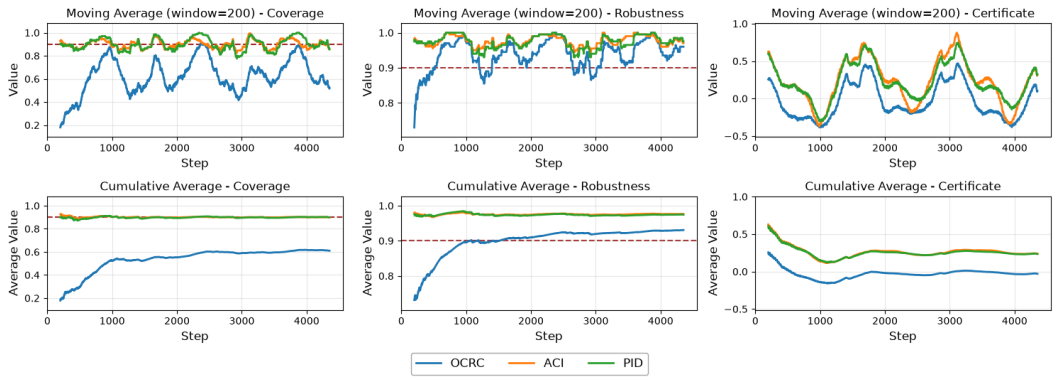

Figure 2: The same settings as Figure 1, except for $T = 4353$.

Result 2. We set $m = 6$ and $h(y) = y$, with 40% of samples used for standardization and MLP pretraining. OCRC achieves lower certificates while maintaining the target robustness level. Its advantage becomes more pronounced in the higher-dimensional setting, where coverage-based methods must control all coordinates although violations in a few dimensions may have little effect on the decision loss.

5.3 REAL-DATA: DEPI DATASET

Experiments are conducted on the **DEPI dataset**, a cleaned and merged version of the M5 Forecasting dataset released by the DEPI Data Science Project at MCIT. The original M5 data, sourced from Walmart, spans daily sales from 2011 to 2016 across 3,049 products in 10 stores located in California, Texas, and Wisconsin. The DEPI version consolidates the original `sales`, `calendar`, and `price` files into a single table with redundant columns removed. It can be downloaded in `https://www.kaggle.com/datasets/davidhalim2004/depi-dataset`.

We select the top $p = 3$ states by total sales and aggregate transactions to a *state–day* granularity. For each time step t, we construct: **Covariates** $X_t \in \mathbb{R}^{p \times F}$: per-state features covering (i) historical sales lags and rolling statistics (lag1/7/14/28, 7/14/28-day rolling mean and std), (ii) calendar and event variables (weekday, month, weekend flag, day-of-year, event flag and impact), and (iii) price signals (selling price and log price ratio); **Labels** $Y_t \in \mathbb{R}^p$: the next-day total sales for the p states at time $t + 1$, matching the information available to a retailer when placing a replenishment order.

The decision-maker chooses a stocking quantity z before observing the realized sales y, and incurs an overage cost a per unit of leftover stock and an underage cost b per unit of unmet demand. The resulting loss is $\phi(y, z) = a^\top (z - y)_+ + b^\top (y - z)_+$, where $(\cdot)_+ = \max(\cdot, 0)$. The prediction set is parameterized in the same form as in the previous experiment.

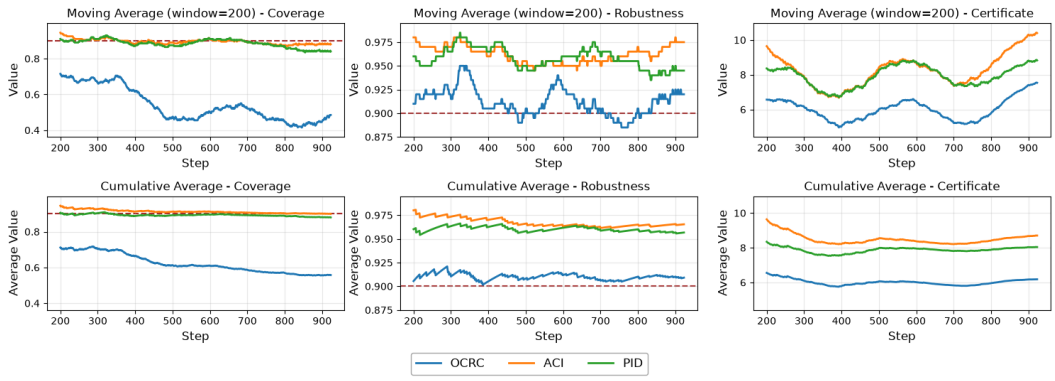

Figure 3: The same settings as Figure 1, except for $T = 924$.

6 CONCLUSION

We proposed Online Conformal Robustness Control (OCRC) for online conditional robust optimization under non-stationary and adversarial data. By replacing prediction-set coverage with direct robustness-violation feedback, OCRC introduces an external scalar threshold that decouples robustness control from prediction-set optimization. We show that the resulting minimax formulation recovers the hindsight constrained optimum as the penalty grows, while providing robustness control under arbitrary data and convergence guarantees under convexity and stationary stochastic settings. Experiments on synthetic and real-world electricity-load and retail-sales datasets demonstrate that OCRC maintains robustness closer to the target level while producing smaller risk certificates than coverage-based online CRO.

Future work will investigate tighter convergence guarantees under more general dependent and non-stationary data, as well as extensions to richer decision structures and adaptive robustness thresholds. It would also be interesting to replace the single external threshold with multiple thresholds that adapt to different components of the robustness violation, and to study whether the resulting algorithm can further reduce conservatism in high-dimensional decision problems.

7 AI USE STATEMENT

The authors used ChatGPT and DeepSeek to polish the language, correct grammar, assist with code debugging, and generate baseline scripts. All AI-generated code was reviewed, tested, and validated by the authors. All scientific content, ideas, experiments, and analyses were conceived and conducted by the authors. The authors reviewed and edited all AI-assisted text and take full responsibility for the content of this paper.

REFERENCES

Anastasios N. Angelopoulos, Stephen Bates, Emmanuel J. Candès, Michael I. Jordan, and Lihua Lei. Learn then test: Calibrating predictive algorithms to achieve risk control. *CoRR*, abs/2110.01052, 2021. URL https://arxiv.org/abs/2110.01052.

Anastasios N. Angelopoulos, Emmanuel J. Candes, and Ryan J. Tibshirani. Conformal PID control for time series prediction, 2023. URL https://arxiv.org/abs/2307.16895.

Anastasios N. Angelopoulos, Michael I. Jordan, and Ryan J. Tibshirani. Gradient equilibrium in online learning: Theory and applications. *Journal of Machine Learning Research*, 26(305):1–68, 2025. URL http://jmlr.org/papers/v26/25-0356.html.

Anastasios Nikolas Angelopoulos, Stephen Bates, Adam Fisch, Lihua Lei, and Tal Schuster. Conformal risk control. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=33XGfHLtZg.

Yajie Bao, Yuyang Huo, Haojie Ren, and Changliang Zou. CAP: A general algorithm for online selective conformal prediction with FCR control. *Journal of Machine Learning Research*, 26 (287):1–74, 2025. URL http://jmlr.org/papers/v26/24-0452.html.

Stephen Bates, Anastasios Angelopoulos, Lihua Lei, Jitendra Malik, and Michael Jordan. Distribution-free, risk-controlling prediction sets. *J. ACM*, 68(6), September 2021. ISSN 0004-5411. doi: 10.1145/3478535. URL https://doi.org/10.1145/3478535.

Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust Optimization*. Princeton University Press, Princeton, NJ, 2009. ISBN 9780691143682.

Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004. doi: 10.1287/opre.1030.0065.

Aadyot Bhatnagar, Huan Wang, Caiming Xiong, and Yu Bai. Improved online conformal prediction via strongly adaptive online learning. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 2337–2363. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/bhatnagar23a.html.

Abhilash Reddy Chenreddy, Nymisha Bandi, and Erick Delage. Data-driven conditional robust optimization. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 9525–9537. Curran Associates, Inc., 2022. doi: 10.52202/068431-0692. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/3df874367ce2c43891aab1ab23ae6959-Paper-Conference.pdf.

Shai Feldman, Liran Ringel, Stephen Bates, and Yaniv Romano. Risk control for online learning models, 2023. URL https://openreview.net/forum?id=uqLDy0HGPR7.

Albin Fredriksson, Anders Forsgren, and Björn Hårdemark. Minimax optimization for handling range and setup uncertainties in proton therapy. *Medical Physics*, 38(3):1672–1684, 2011. doi: https://doi.org/10.1118/1.3556559. URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.3556559.

Isaac Gibbs and Emmanuel Candes. Adaptive conformal inference under distribution shift. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 1660–1672. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/0d441de75945e5acbc865406fc9a2559-Paper.pdf.

Isaac Gibbs, John J. Cherian, and Emmanuel J. Candès. Conformal prediction with conditional guarantees, 2024. URL https://arxiv.org/abs/2305.12616.

D. Goldfarb and G. Iyengar. Robust portfolio selection problems. 28(1):1–38, February 2003. ISSN 0364-765X. doi: 10.1287/moor.28.1.1.14260. URL https://doi.org/10.1287/moor.28.1.1.14260.

Yang Hu, Jieren Tan, Changliang Zou, Yajie Bao, and Haojie Ren. Conformal robustness control: A new strategy for robust decision. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 102698–102726, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/a80188f9b1246d4b95d396165cf99207-Paper-Conference.pdf.

Chancellor Johnstone and Bruce Cox. Conformal uncertainty sets for robust optimization. In Lars Carlsson, Zhiyuan Luo, Giovanni Cherubin, and Khuong An Nguyen (eds.), *Proceedings of the Tenth Symposium on Conformal and Probabilistic Prediction and Applications*, volume 152 of *Proceedings of Machine Learning Research*, pp. 72–90. PMLR, 08–10 Sep 2021. URL https://proceedings.mlr.press/v152/johnstone21a.html.

Shayan Kiyani, George J. Pappas, Aaron Roth, and Hamed Hassani. Decision theoretic foundations for conformal prediction: Optimal uncertainty quantification for risk-averse agents. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 30943–30965. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/kiyani25a.html.

Jing Lei, Max G'Sell, Alessandro Rinaldo, Ryan J. Tibshirani, and Larry Wasserman. Distribution-free predictive inference for regression. *Journal of the American Statistical Association*, 113 (523):1094–1111, 2018. doi: 10.1080/01621459.2017.1307116. URL https://doi.org/10.1080/01621459.2017.1307116.

Jordan Lekeufack, Anastasios N. Angelopoulos, Andrea Bajcsy, Michael I. Jordan, and Jitendra Malik. Conformal decision theory: Safe autonomous decisions from imperfect predictions, 2024. URL https://arxiv.org/abs/2310.05921.

Tianyi Lin, Chi Jin, and Michael Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6083–6093. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/lin20a.html.

Lars Lindemann, Matthew Cleaveland, Gihyun Shim, and George J. Pappas. Safe planning in dynamic environments using conformal prediction. *IEEE Robotics and Automation Letters*, 8(8): 5116–5123, 2023. doi: 10.1109/LRA.2023.3292071.

Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. Trading regret for efficiency: Online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(81):2503–2528, 2012. URL http://jmlr.org/papers/v13/mahdavi12a.html.

Yash P. Patel, Sahana Rayan, and Ambuj Tewari. Conformal contextual robust optimization. In Sanjoy Dasgupta, Stephan Mandt, and Yingzhen Li (eds.), *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 2485–2493. PMLR, 02–04 May 2024. URL https://proceedings.mlr.press/v238/patel24a.html.

Aleksandr Podkopaev, Dong Xu, and Kuang-Chih Lee. Adaptive conformal inference by betting. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 40886–40907. PMLR, 21–27 Jul 2024. URL https://proceedings.mlr.press/v235/podkopaev24a.html.

Chunlin Sun, Linyu Liu, and Xiaocheng Li. Predict-then-calibrate: A new perspective of robust contextual LP. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in Neural Information Processing Systems*, volume 36, pp. 17713–17741. Curran Associates, Inc., 2023. doi: 10.52202/075280-0778. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/397271e11322fae8ba7f827c50ca8d9b-Paper-Conference.pdf.

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic Learning in A Random World*, volume 29. Springer, 2005.

Asaf Weinstein and Aaditya Ramdas. Online control of the false coverage rate and false sign rate. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 10193–10202. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/weinstein20a.html.

Chen Xu and Yao Xie. Conformal prediction interval for dynamic time-series. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 11559–11569. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/xu21h.html.

Chen Xu and Yao Xie. Sequential predictive conformal inference for time series. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 38707–38727. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/xu23r.html.

Xiaohe Yan, Chenghong Gu, Xin Zhang, and Furong Li. Robust optimization-based energy storage operation for system congestion management. *IEEE Systems Journal*, 14(2):2694–2702, 2020. doi: 10.1109/JSYST.2019.2932897.

Margaux Zaffran, Olivier Feron, Yannig Goude, Julie Josse, and Aymeric Dieuleveut. Adaptive conformal predictions for time series. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 25834–25866. PMLR, 17–23 Jul 2022. URL https://proceedings.mlr.press/v162/zaffran22a.html.

A Proofs

A.1 Proof of Theorem 1

Define

$$\psi(\theta) = \max_{q \geq 0} \int_0^q g(\theta, q')dq',$$

then θ^*_κ, f^*_κ are solution and value of $\min_\theta f(\theta) + \kappa\psi(\theta)$, respectively. It is easy to find that

- ψ is continuous (through Berge's Maximum Theorem),
- $\psi(\theta) \geq 0, \quad$ for all $\theta \in \Theta$,
- $\psi(\theta) = 0$ if and only if $\theta \in D$.

Therefore, for $0 < \kappa_1 \leq \kappa_2$, we have $f^*_{\kappa_1} \leq f^*_{\kappa_2}$. On the other hand,

$$f^*_{\kappa_2} \leq f(\theta^*) + \kappa_2 \psi(\theta^*) = f^*.$$

Consequently, $f^*_{\kappa_1} \leq f^*_{\kappa_2} \leq f^*$. Furthermore, $\limsup_{\kappa\to\infty} f^*_\kappa \leq f^*$.

We next prove $\liminf_{\kappa\to\infty} f^*_\kappa \geq f^*$. Assume, for the sake of contradiction, that there exists a subsequence $\{\kappa_j\}_{j\geq 1}$ and $\epsilon > 0$, such that $\lim_{j\to\infty} f^*_{\kappa_j} = f^* - \epsilon$. According to the compactness of Θ, $\{\theta^*_{\kappa_j}\}_{j\geq 1}$ converges to a certain point $\bar{\theta}$.

If $\bar{\theta} \in D$, $\psi(\bar{\theta}) = 0$, by continuity $\lim_{j\to\infty} \psi(\theta^*_{\kappa_j}) = 0$, thus

$$f(\theta^*_{\kappa_j}) \leq f(\theta^*_{\kappa_j}) + \kappa \psi(\theta^*_{\kappa_j}) \leq f^* - \epsilon.$$

Taking the limit on both sides yields $f(\bar{\theta}) \leq f^* - \epsilon < f^*$, which contradicts the definition that f^* is the optimal value of the original problem.

If $\bar{\theta} \notin D$, $\psi(\bar{\theta}) > 0$, by continuity $\psi(\theta^*_{\kappa_j}) \geq \psi(\bar{\theta})/2 > 0$ for sufficiently large j. Thus

$$f(\theta^*_{\kappa_j}) + \kappa_j \psi(\theta^*_{\kappa_j}) \geq \inf_{\theta\in\Theta} f(\theta) + \kappa_j \psi(\bar{\theta})/2 \to \infty.$$

It contradicts the fact that $f^* - \epsilon$ is finite.

Overall, $f^* \leq \liminf_{\kappa\to\infty} f^*_\kappa \leq \limsup_{\kappa\to\infty} f^*_\kappa \leq f^*$, i.e. $\lim_{\kappa\to\infty} f^*_\kappa = f^*$. Moreover, in the above argument, the contradiction in the case where $\bar{\theta} \notin D$ does not rely on the contradiction assumption, therefore $\text{dist}(\theta^*_\kappa, D) \to 0$. It then immediately follows that $\theta^*_\kappa \to \theta^*, q^*_\kappa \to 0$.

A.2 Proof of Theorem 2

We prove by induction that $q_t \leq 2B_\phi - H_\sigma^{-1}(\alpha) + \gamma \max\{\alpha, 1-\alpha\}$ for all $t = 1, 2, \ldots$. Assume it holds for t.

- If $q_t \leq 2B_\phi - H_\sigma^{-1}(\alpha)$, since $g_t(\theta_t, q_t) \leq \max\{\alpha, 1-\alpha\}$, $q_{t+1} \leq 2B_\phi + \gamma \max\{\alpha, 1-\alpha\}$.
- If $q_t > 2B_\phi - H_\sigma^{-1}(\alpha)$, $\phi(Y_t, z_{\theta_t}(X_t)) - r_\theta(X_t) \leq 2B_\phi < q_t + H_\sigma^{-1}(\alpha)$, which means $g_t(\theta_t, q_t) < 0$ and $q_{t+1} < q_t$.

The induction is complete. Summing the update relation yields

$$\sum_{t=1}^{T} g_t(\theta_t, q_t) \leq \frac{q_{T+1} - q_1}{\gamma}.$$

Set $q_1 = 0$ and substitute to complete the proof.

A.3 Proof of Theorem 3

We use the following lemma in Mahdavi et al. (2012).

Lemma 1. *For all $\theta \in \Theta$ and $q \geq 0$,*

$$\sum_{t=1}^{T} \{\mathcal{L}_t(\theta_t, q) - \mathcal{L}_t(\theta, q_t)\} \leq \frac{\|\theta_1 - \theta\|^2 - \|\theta_{T+1} - \theta\|^2}{2\eta_\theta} + \frac{|q_1 - q|^2 - |q_{T+1} - q|^2}{2\eta_q} + \sum_{t=1}^{T} \left[\frac{\eta_\theta}{2} \|\nabla_\theta \mathcal{L}_t(\theta_t, q_t)\|^2 + \frac{\eta_q}{2} \|\nabla_q \mathcal{L}_t(\theta_t, q_t)\|^2\right].$$

To prove this, expand the squared distance after one update.

$$\|\theta_{t+1} - \theta\|^2 \leq \|\theta_t - \theta\|^2 - 2\eta_\theta \nabla_\theta \mathcal{L}_t(\theta_t, q_t)^\top (\theta_t - \theta) + \eta_\theta^2 \|\nabla_\theta \mathcal{L}_t(\theta_t, q_t)\|^2.$$

By convexity,

$$\begin{aligned} \mathcal{L}_t(\theta_t, q_t) - \mathcal{L}_t(\theta, q_t) &\leq \nabla_\theta \mathcal{L}_t(\theta_t, q_t)^\top (\theta_t - \theta) \\ &\leq \frac{\|\theta_{t+1} - \theta\|^2 - \|\theta_t - \theta\|^2}{2\eta_\theta} + \frac{\eta_\theta}{2} \|\nabla_\theta \mathcal{L}_t(\theta_t, q_t)\|^2. \end{aligned} \tag{5}$$

Similarly by concavity,

$$\begin{aligned}\mathcal{L}_t(\theta_t, q) - \mathcal{L}_t(\theta_t, q_t) &\le \nabla_q \mathcal{L}_t(\theta_t, q_t)^\top (q - q_t) \\ &\le \frac{\|q - q_t\|^2 - \|q - q_{t+1}\|^2}{2\eta_q} + \frac{\eta_q}{2}\|\nabla_q \mathcal{L}_t(\theta_t, q_t)\|^2. \end{aligned} \tag{6}$$

Combine (5) and (6), then sum up for $t = 1, 2, \ldots, T$ to finish the proof of the lemma. Using the boundedness of Θ and Lipschitz continuity (which means $\|\nabla \mathcal{L}_t\| = O(\kappa/\sigma)$),

$$\sum_{t=1}^{T} \{\mathcal{L}_t(\theta_t, q) - \mathcal{L}_t(\theta, q_t)\} = O\left(\frac{\kappa^2 \eta_\theta}{\sigma^2} T + \frac{1}{\eta_\theta} + \frac{\kappa^2 \eta_q}{\sigma^2} T + \frac{1}{\eta_q}\right).$$

Substitute the default setting to get the bound.

A.4 Proof of Theorem 4

Note that $\ell(\theta, q) = \mathbb{E}[\mathcal{L}_t(\theta, q)|\mathcal{F}_{t-1}]$ shares the same Lipschitz and smooth parameters with $\mathcal{L}_t$. According to Assumption 2 and 3, they satisfy $L_\ell = O(\kappa/\sigma)$ and $s_\ell = O(\kappa/\sigma^2)$. Since $g_t(\theta, q)$ is monotonically decreasing in q, $\ell(\theta, q)$ is concave in q. We next follow the conclusion of nonconvex-concave SGDA by Lin et al. (2020) without proof.

Lemma 2 ((Lin et al., 2020) Theorem 4.9)**.** *Given an SGDA process*

$$\begin{cases} x_{t+1} = \Pi_{\mathcal{X}}\{x_t - \eta_x \nabla_x F_t(x_t, y_t)\}, \\ y_{t+1} = \Pi_{\mathcal{Y}}\{y_t + \eta_y \nabla_y F_t(x_t, y_t)\}, \end{cases}$$

assume that $f(x, y) := \mathbb{E}[F_t(x, y)|\mathcal{F}_{t-1}]$ is time-invariant, L-Lipschitz, s-smooth and concave in y. Also, assume that

$$\mathbb{E}[\|\nabla f(x, y) - \nabla F_t(x, y)\|^2] \le \sigma_\xi^2 < \infty, \quad \textit{for all } t.$$

Define $\Phi(x) = \max_{y \in \mathcal{Y}} f(x, y)$. Given $\epsilon > 0$, set the step-sizes $\eta_x = \Theta(\epsilon^4/(s^3 D^2 (L^2 + \sigma_\xi^2)))$, $\eta_y = \Theta(\epsilon^2/(s\sigma^2))$, then total time

$$T = O\left(\left(\frac{s^3 (L^2 + \sigma_\xi^2) D^2 \hat{\Delta}_\Phi}{\epsilon^6} + \frac{s^3 D^2 \hat{\Delta}_0}{\epsilon^4}\right) \max\left\{1, \frac{\sigma^2}{\epsilon^2}\right\}\right),$$

where D is the bound of $\mathcal{Y}$, $\hat{\Delta}_\Phi = \Phi(x_0) - \min_{x \in \mathcal{X}} \Phi(x)$ and $\hat{\Delta}_0 = \Phi(x_0) - f(x, y)$, then $\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[\|\nabla \Phi(x_t)\|^2] \le \epsilon^2$.

In our settings, $\hat{\Delta}_\Phi$ and $\hat{\Delta}_0$ are constants (see analysis of the optimal value when $\kappa \to \infty$ in A.1). Eliminating ϵ using T and substituting into our default setting completes the proof.

A.5 Proof of Theorem 5

For each t,

$$\mathbf{1}\{V_t(\theta_t, q_t) > 0\} \le H_\sigma(V_t(\theta_t, q_t)) + C_H \mathbf{1}\{|V_t(\theta_t, q_t)| \le c_H \sigma\} + \epsilon_\sigma.$$

Averaging over t, the first term is the smoothed robustness violation. The second term is controlled by Assumption 5, setting $\epsilon = c_H \sigma$. The final term is the smoothing approximation error.

B Additional Experiments

B.1 Distribution shift data simulation

We perform additional simulations, which relax the stationarity and generate $\{X_t\}$ as an AR(5) time series

$$X_{t,1} = 1 + \sum_{s=1}^{5} 0.5^s (X_{t-s,1} - 1) + v_{t,1}, \quad X_{t,2} = 1 + \sum_{s=1}^{5} 0.25^s (X_{t-s,1} - 1) + v_{t,2},$$

where $v_{t,1}, v_{t,2}$ are independent $N(0, 0.04)$ noise variables.

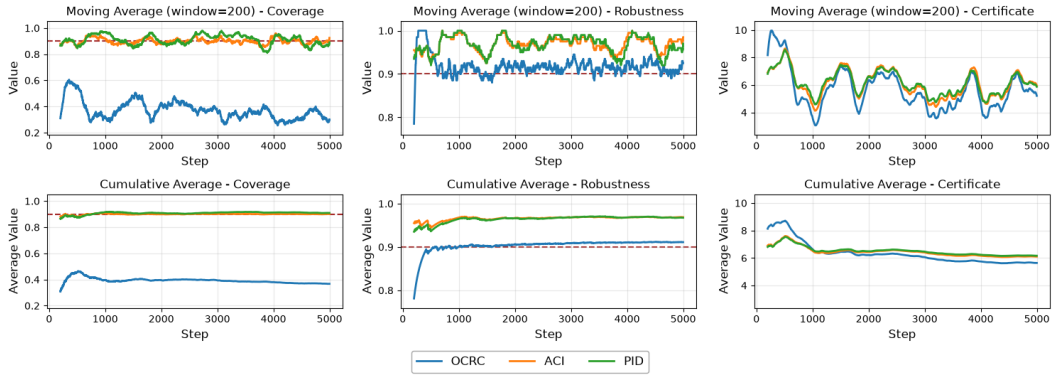

Figure 4: The same settings as Figure 1.



Figure 5: Moving average robustness under different η_q. All other settings remain exactly the same.

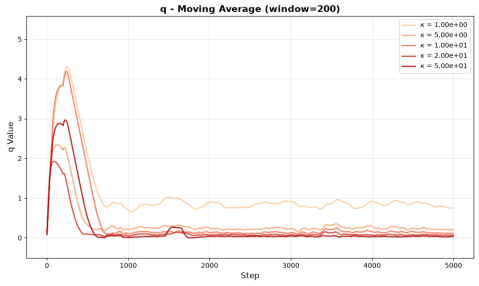

Figure 6: Moving average q value under different κ. All other settings remain exactly the same.

Result 3. The AR model is a stationary time series, but it is not stationary in the sense of loss expectations and thus does not satisfy the assumptions of Theorem 5. The loss distribution actually changes over time. Nevertheless, the experiments show that OCRC is able to track these changes while controlling robustness, and performs better than the coverage-based methods.

We also conducted additional experiments under this setting to illustrate the roles of q and κ.

B.2 Roles of q and κ

We fix the other parameters and vary η_q and κ to observe the convergence behavior of the algorithm. The former is used to illustrate the specific role of q in the algorithm, while the latter serves to corroborate the result of Theorem 1.

Result 4. As can be seen from the left figure, when η_q is small, the variation of robustness is relatively rigid. This is because the poor initialization of the parameter leads to a very low robustness initially, causing the prediction set to expand sharply in the early stage. After the parameter stabilizes, the prediction set has accumulated too much error and takes a long time to shrink back to an appropriate size. In contrast, when η_q is large, q smooths this process by absorbing part of the early-stage error, resulting in a more stable evolution of the prediction set size. This also illustrates the role of q in response to changes in the environment.

Result 5. In the right figure, a larger κ corresponds to a smaller value of q after stabilization, which corroborates the result of Theorem 1. When κ is sufficiently large, q has little influence on the original problem, yet it enables the algorithm to achieve robustness control under arbitrary data.