

Omniprediction with Long-Term Constraints

Yahav Bechavod

Jiuyao Lu

Aaron Roth

University of Pennsylvania

YAHAV@SEAS.UPENN.EDU

JIU YAOLU@WHARTON.UPENN.EDU

AAROTH@CIS.UPENN.EDU

Editors: Steve Hanneke and Tor Lattimore

Abstract

We introduce and study the problem of online omniprediction with long-term constraints. At each round, a forecaster is tasked with generating predictions for an underlying (adaptively, adversarially chosen) state that are broadcast to a collection of downstream agents, who must each choose an action. Each of the downstream agents has both a utility function mapping actions and state to utilities, and a vector-valued constraint function mapping actions and states to vector-valued costs. The utility and constraint functions can arbitrarily differ across downstream agents. Their goal is to choose actions that guarantee themselves no regret while simultaneously guaranteeing that they do not cumulatively violate the constraints across time. We show how to make a single set of predictions so that each of the downstream agents can guarantee this by acting as a simple function of the predictions, guaranteeing each of them $\tilde{O}(|\mathcal{A}|\sqrt{T})$ regret and $O(|\mathcal{A}|)$ cumulative constraint violation for a finite action space \mathcal{A} . We also show how to extend our guarantees to arbitrary intersecting contextually defined *subsequences*, guaranteeing each agent both regret and constraint violation bounds not just marginally, but simultaneously on each subsequence, against a benchmark set of actions simultaneously tailored to each subsequence.

Keywords: omniprediction, long-term constraints, online learning, swap regret, decision calibration

1. Introduction

The traditional problem of “online learning from expert advice” studies a sequential optimization problem that is decoupled across time. At each iteration $t \in \{1, \dots, T\}$, an adversary chooses a “state” $y_t \in \mathcal{Y}$, and simultaneously, the learner/decision-maker chooses an action $a_t \in \mathcal{A}$. The learner/decision-maker then experiences utility $u(a_t, y_t) \in [0, 1]$ ¹. The goal of the learner is to maximize $\sum_{t=1}^T u(a_t, y_t)$, and the standard “external regret benchmark” compares the realized utility to the utility that could have been obtained in hindsight by the best -fixed- action:

$$\text{Reg}_{\text{ext}}(1 : T) = \max_{a \in \mathcal{A}} \sum_{t=1}^T (u(a, y_t) - u(a_t, y_t)).$$

In strategic settings, the stronger “swap regret” benchmark is often studied because it offers protection against manipulation (agents with no swap regret guarantees cannot be exploited in the same ways that agents with standard no regret guarantees can be (Braverman et al., 2018; Deng et al., 2019; Rubinstein and Zhao, 2024; Arunachaleswaran et al., 2024)):

$$\text{Reg}_{\text{swap}}(1 : T) = \max_{\phi: \mathcal{A} \rightarrow \mathcal{A}} \sum_{t=1}^T (u(\phi(a_t), y_t) - u(a_t, y_t)).$$

1. The formalism of choosing a “state” $y \in \mathcal{Y}$ is fully general; for instance, in the experts problem we can take $\mathcal{Y} = [0, 1]^{|\mathcal{A}|}$ where y_a represents the utility of action a .

Swap regret compares the realized accumulated utility of the learner to the best utility they could have obtained in hindsight by applying a “swap function” ϕ to their history of actions, consistently mapping each action a to a different action $\phi(a) \in \mathcal{A}$. External regret is a special case of swap regret in which the “swap functions” ϕ are restricted to be constant valued.

These are both “decoupled” optimization problems because they are linearly separable across time — i.e. the ideal solution would always be (if it were possible) to separately optimize every day, and play $a_t \in \arg \max_{a \in \mathcal{A}} u(a, y_t)$. Regret minimization is a classical problem and there are many algorithms for solving it, obtaining both external and swap regret bounds that scale as $O(\sqrt{T})$, which is optimal (Hannan, 1957; Littlestone and Warmuth, 1994; Cesa-Bianchi et al., 1997; Arora et al., 2012; Blum and Mansour, 2007).

We write “learner/decision-maker” above because there are really two problems to solve: *predicting* the value of the next $y_t \in \mathcal{Y}$ and choosing to *act* accordingly by choosing $a_t \in \mathcal{A}$. Traditional algorithmic approaches combine these two problems and directly learn distributions over actions. But a recent line of work called “omniprediction”, beginning with Gopalan et al. (2022) (initially in a batch/distributional learning setting) proposes to decouple these two problems by producing *predictions* p_t that can be simultaneously consumed by many downstream decision makers who may have different utility functions u , while providing all of them with corresponding guarantees. This line of work has been extended to the online adversarial learning setting (Noarov et al., 2023; Garg et al., 2024; Roth and Shi, 2024; Dwork et al., 2025; Okoroafor et al., 2025; Lu et al., 2025), in which many downstream decision makers are explicitly promised regret bounds — sometimes external regret bounds (Kleinberg et al., 2023; Garg et al., 2024; Dwork et al., 2025; Okoroafor et al., 2025) and sometimes stronger swap regret bounds (Noarov et al., 2023; Roth and Shi, 2024; Hu and Wu, 2024; Lu et al., 2025). These works explicitly take advantage of the separable nature of the problem: They give algorithms for making predictions p_t (subject to various kinds of “calibration” style constraints) such that downstream decision makers can be guaranteed strong regret bounds by individually “best responding” to the predictions, each playing an action $a_t = \arg \max_{a \in \mathcal{A}} u(a, \hat{p}_t)$ — i.e. locally optimizing for their own utility function on a day-by-day basis as if the predictions were correct, ignoring the history of outcomes and actions played.

But many online optimization problems are *not* linearly separable in this way. Consider an energy management system where predictions correspond to electricity prices, and actions include: purchasing electricity for immediate use, buying excess power to charge batteries, discharging stored energy, or temporarily shutting down operations. Different downstream decision-makers face distinct constraints: residential users might defer appliance usage to low-cost periods, manufacturers require minimum uptime to meet production quotas, and grant-funded operations must stay within fixed budgets. These constraints couple decisions across time, making prior omniprediction methods inapplicable. Our goal is to give omniprediction style bounds for decision makers like this who are bound to satisfy constraints across time. To crystallize the tension, observe that the goal of omniprediction is to *decouple* prediction from decision making — but decision makers facing long term constraints must necessarily *couple* their decisions across time, and this coupling behaves differently for differently constrained decision makers.

The literature on “learning with long-term constraints” *has* studied different variants of this problem for a single, combined learner/decision-maker who directly learns in action space (Mannor et al., 2009; Sun et al., 2017; Yu and Neely, 2020; Guo et al., 2022; Castiglioni et al., 2022; Qiu et al., 2023; Sinha and Vaze, 2024). In the problem of learning with long-term constraints, not only does a decision maker have a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$, but they have a set of *constraints*

$\{c_j\}_{j \in [J]}$, each of the form $c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]$.² The goal now is to maximize $\sum_{t=1}^T u(a_t, y_t)$, while satisfying each constraint globally: for each $j \in [J]$, satisfying: $\sum_{t=1}^T c_j(a_t, y_t) \leq 0$. It has been known since [Mannor et al. \(2009\)](#) that in adversarial settings it is not possible to compete with the best fixed action in hindsight that satisfies the constraints marginally. Instead, what has become the standard benchmark is the set of actions that in hindsight satisfy the realized constraint *every round*:

$$\mathcal{A}_{1:T}^c = \{a \in \mathcal{A} : c_j(a, y_t) \leq 0 \text{ for every } t \in [T] \text{ and } j \in [J]\}.$$

The goal in this literature is to minimize external regret with respect to this benchmark:

$$\text{Reg}_{\text{ext}}(1 : T) = \max_{a \in \mathcal{A}_{1:T}^c} \sum_{t=1}^T (u(a, y_t) - u(a_t, y_t)).$$

while simultaneously minimizing the maximum cumulative constraint violation:

$$\text{CCV}(1 : T) = \max_j \sum_{t=1}^T c_j(a_t, y_t).$$

We introduce the omniprediction variant of this problem, in which a centralized provider of “predictions as a service” publishes forecasts p_t of the outcome y_t . Many different decision makers, who differ both in their utility functions and constraint functions then consume these predictions to choose actions. Our main result is an algorithm that can simultaneously guarantee each such decision maker regret bounded as $\text{Reg}(1 : T) \in \tilde{O}(|\mathcal{A}|\sqrt{T})$ (matching the optimal dependence on T , even absent constraints, and defined with respect to the benchmark determined by their own personal constraints) and cumulative constraint violation $\text{CCV}(1 : T) \in O(|\mathcal{A}|)$ — i.e. independent of T . In fact, our techniques yield stronger swap regret bounds at the same rate. In the constrained setting, we define swap regret with respect to the benchmark as:

$$\text{Reg}_{\text{swap}}(1 : T) = \max_{\phi: \mathcal{A} \rightarrow \mathcal{A}_{1:T}^c} \sum_{t=1}^T (u(\phi(a_t), y_t) - u(a_t, y_t)).$$

We also show how these results can be extended to arbitrary collections of *subsequences* \mathcal{S} , where each $S \in \mathcal{S}$ is a subset of the timesteps $S \subseteq \{1, \dots, T\}$ that can be determined dynamically by external context. For example, a subsequence S might consist of “hot days” that might be correlated with higher electricity prices; the subsequences may intersect arbitrarily. The set of benchmark actions will be different for each subsequence, as some actions may satisfy the constraints at every round in a subsequence S even if they do not satisfy the constraint on every round in $\{1, \dots, T\}$:

$$\mathcal{A}_S^c = \{a \in \mathcal{A} : c_j(a, y_t) \leq 0 \text{ for every } t \in S \text{ and } j \in [J]\}.$$

We can similarly define regret on a subsequence (to the benchmark \mathcal{A}_S^c particular to that subsequence) and cumulative constraint violation on a subsequence:

$$\begin{aligned} \text{Reg}_{\text{swap}}(S) &= \max_{\phi: \mathcal{A} \rightarrow \mathcal{A}_S^c} \sum_{t \in S} (u(\phi(a_t), y_t) - u(a_t, y_t)), \\ \text{CCV}(S) &= \max_j \sum_{t \in S} c_j(a_t, y_t). \end{aligned}$$

2. In the literature on learning with long-term constraints, incoming constraints are usually defined only as a function of the action set. We instead define a fixed set of constraints for each decision-maker, but allow dependence on the label set as well. Hence we can equivalently handle adversarial constraints by encoding adaptivity in the incoming labels.

Given an arbitrary collection of subsequences \mathcal{S} , we show how to obtain simultaneously for each downstream agent, and simultaneously for each subsequence $S \in \mathcal{S}$ regret and cumulative constraint violation bounds: $\text{Reg}_{\text{swap}}(S) \in \tilde{O}(|\mathcal{A}|\sqrt{|S|})$ where $|S|$ is the length of subsequence S and $\text{CCV}(S) \in O(|\mathcal{A}|)$ — where the constant in the cumulative constraint violation bound now depends linearly on $|\mathcal{S}|$ (the number of subsequences of interest), but the constant in the regret bound depends only logarithmically on $|\mathcal{S}|$. Crucially, the benchmark with which regret is measured is tailored to the subsequence — we are able to compete with a richer set of actions on subsequences that turn out, ex-post to be “easier” than others. A full discussion of related work appears in Appendix A.

1.1. Our Results in More Detail

The outcome space $\mathcal{Y} = [0, 1]^d$, and so outcomes y are d -dimensional vectors. At every round t our algorithm produces predictions $p_t \in \mathcal{Y}$ that are also d -dimensional vectors. Our goal is for downstream decision makers to be able to treat these predictions p_t “as if” they were actual outcomes y_t . Each decision maker has an action space \mathcal{A} and is parameterized by a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ that we assume is linear in \mathcal{Y} — i.e. such that for every a , $u(a, \cdot)$ is a linear function. For a single downstream decision maker, this models the experts learning problem without loss of generality by taking $d = |\mathcal{A}|$ and letting $u(a, y) = y_a$, but is more general, since we do not restrict either the form of the utility function or d in this way. Each downstream decision maker is also parameterized by a vector valued constraint function $\mathbf{c} : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]^J$. We assume nothing about its functional form — e.g. \mathbf{c} need not be linear or even convex in \mathcal{Y} .

We work in the standard finite-action setting. There is a natural continuous-action reduction in an important special case. Suppose actions are mixed strategies over k base actions and utilities/constraints are evaluated in expectation. Then a constrained best response under J constraints can be taken to be an extreme point supported on at most $J + 1$ actions. Discretizing this sparse class gives an ε -net of size roughly $k^{J+1}\varepsilon^{-J}$, reducing the problem to the finite-action setting up to the usual discretization error. This is practical when J is 1 or a small constant, as in much prior work on long-term constraints, but becomes impractical as J grows.

Handling Constraints via Elimination In work studying online omniprediction for unconstrained online learning problems (Kleinberg et al., 2023; Noarov et al., 2023; Roth and Shi, 2024; Garg et al., 2024; Dwork et al., 2025; Okoroafor et al., 2025), downstream decision makers with utility function u “best respond” to forecasts, playing action $a_t \in \text{BR}^u(p_t) = \arg \max_{a \in \mathcal{A}} u(a, p_t)$. In our problem of learning with long-term constraints, we similarly study decision makers who play *constrained* best responses. Given a subset of the action space $\mathcal{C} \subseteq \mathcal{A}$, the constrained best response to a prediction p_t is $\text{CBR}_{\mathcal{C}}^u(p_t) = \arg \max_{a \in \mathcal{C}} u(a, p_t)$ — i.e. the action in the constrained set \mathcal{C} that would be utility maximizing were the prediction correct. The question is how decision makers choose the constraint set \mathcal{C} . We study two simple methods, corresponding to two different kinds of benchmarks we can compete with. Most straightforwardly, we study decision makers who start with $\mathcal{C} = \mathcal{A}$, and then eliminate actions that have been observed to violate any constraints on any previous round — i.e. at round t ,

$$\mathcal{C} = \widehat{\mathcal{A}}_t^{\mathbf{c}} = \{a \in \mathcal{A} : c_j(a, y_{t'}) \leq 0 \text{ for all } j \in [J], t' < t\}.$$

This is formalized in Algorithm 1. We show in Theorems 8 and 14 that (when paired with predictions produced by our algorithm), this results in decision makers who cumulatively violate any con-

straint at most $|\mathcal{A}|$ times over the course of the interaction, and have diminishing external and swap regret to the standard benchmark $\mathcal{A}_{1:T}^c$. We also consider decision makers who do not eliminate actions the first time they are observed to violate the constraint, but are more conservative, eliminating them only after the action has been observed to have a cumulative constraint violation crossing a well-chosen threshold. This is formalized in Algorithm 2. We show in Theorems 12 and 20 that this lets us compete with a richer benchmark class, $\mathcal{A}_{1:T}^{\mathbb{E}[c]} = \{a : \mathbb{E}_{y_t \sim Y_t}[c_j(a, y_t)] \leq 0 \text{ for every } t \in [T] \text{ and } j \in [J]\}$, where here the expectation in the definition of the benchmark is taken over the distribution over outcomes chosen by the adversary at each round — again both with external and swap regret bounds. In other words, the benchmark allows us to compare to actions that sometimes violate constraints in realization, so long as they do not in expectation over the adversary’s randomness. Competing with this richer benchmark class results in a cumulative constraint violation that now grows as $\tilde{O}(|\mathcal{A}|\sqrt{T})$.

Making Predictions Our algorithm for making predictions is parameterized by the collection of utility/constraint functions (u, c) , one for each downstream decision maker. It makes predictions that are unbiased *conditionally on the decisions of each downstream decision maker*, in the sense that for each (u, c) :

$$\left\| \sum_{t=1}^T \mathbb{1}[\text{CBR}_{\hat{\mathcal{A}}_t^c}^u(p_t) = a](p_t - y_t) \right\|_{\infty} \leq O(\sqrt{T}).$$

This can be accomplished using an appropriate parameterization of the algorithm of Noarov et al. (2023), with bias bounds that depend only logarithmically on the number of downstream decision makers (u, c) of interest, and running time that depends linearly on the number of downstream decision makers. Similar “decision calibration” conditions have been used to give regret bounds in the unconstrained case (Noarov et al., 2023; Roth and Shi, 2024; Lu et al., 2025) — We show that they continue to give regret bounds in the constrained case (Theorems 14, 20) because the elimination method used to eliminate actions by the downstream decision makers maintains the invariant that all of the actions that eventually end up in the benchmark set $\mathcal{A}_{1:T}^c$ are contained in the feasible set $\hat{\mathcal{A}}_t^c$ used by the downstream agents at every intermediate step $t \leq T$.

Extension to Subsequences Finally, in Section 5 and Appendix D we show how all of our results can be extended to give regret and constraint violation bounds $\text{Reg}(S), \text{CCV}(S)$ for arbitrary collections of context-sensitive *subsequences* $S \in \mathcal{S}$. For the first and simpler benchmark, the idea is simple and requires modifying both the elimination strategy used by the downstream decision makers and the prediction strategy. The downstream decision makers must now maintain sets of feasible actions $\hat{\mathcal{A}}_{t,S}^c$ separately, one for each subsequence $S \in \mathcal{S}$, as before, eliminating actions whenever they have been observed to violate constraints on previous rounds *that are contained in the subsequence* S . Given a new round t , they then solve the constrained best response problem over the set $\mathcal{C} = \cup_{S \in \mathcal{S}: t \in S} \hat{\mathcal{A}}_{t,S}^c$ — i.e. by taking the union of all actions that have not been eliminated on *any* of the subsequences that contain the current round. The predictions that we make must now also satisfy a richer set of conditional bias conditions: informally speaking, they must be unbiased conditionally both on the action chosen by each of the downstream decision makers, and now also membership of the current round in each of the subsequences $S \in \mathcal{S}$. Once again this can be accomplished via an instantiation of the algorithm of Noarov et al. (2023) at a cost in bias that is only logarithmic in $|\mathcal{S}|$ and in running time that is linear in $|\mathcal{S}|$, the number of subsequences of interest. However, for the second, stronger benchmark, the simple approach described above does

not suffice informally because we now don't just count rounds at which constraints were violated, but we sum quantitative constraint violations which can be negative (i.e. on some rounds we can have slack). Consequently, an action can be feasible on a subsequence (and hence never eliminated for that subsequence) while leading to large cumulative constraint violation on intersections of that subsequence with other subsequences. This is problematic for the learner who must give guarantees simultaneously for many intersecting subsequences. In Appendix D we present a more sophisticated elimination algorithm that is able to bound the cumulative constraint violation on all subsequences, despite their intersections.

2. Model and Preliminaries

Let \mathcal{X} denote the feature space and \mathcal{Y} denote the outcome/label space. Throughout, we consider $\mathcal{Y} = [0, 1]^d$. We consider a set of agents \mathcal{N} with an arbitrary action space \mathcal{A} . Each agent is equipped with a tuple (u, c_1, \dots, c_J) , which includes a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$ ³. We also sometimes write the constraint functions as a single vector valued function $\mathbf{c} = (c_1, \dots, c_J)$. Since agents are uniquely defined by their corresponding tuples, we treat agents and their tuples interchangeably. We assume that the utility functions are linear and Lipschitz-continuous in y .

Assumption 1 Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$. We assume that for every action $a \in \mathcal{A}$, $u(a, y)$ is linear in y , i.e. $u(a, k_1 y_1 + k_2 y_2) = k_1 u(a, y_1) + k_2 u(a, y_2)$ for all $k_1, k_2 \in \mathbb{R}$, $y_1, y_2 \in \mathcal{Y}$. Moreover, we assume that $u(a, y)$ is L -Lipschitz in y in the ℓ_∞ norm: for any $y_1, y_2 \in \mathcal{Y}$, $|u(a, y_1) - u(a, y_2)| \leq L \|y_1 - y_2\|_\infty$.

Remark 1 For simplicity we assume that the utility functions are linear in y , but we can equally well handle the case in which the utility functions are affine in y , as we can augment the label space with an additional coordinate that takes constant value 1. This preserves the convexity of the label space and allows for arbitrary constant offsets in the utility of each action.

Assuming linear/affine utility functions is only more general than the standard assumption that decision makers are risk neutral in the sense that in the face of randomness, decision makers act to maximize their expected utility. If \mathcal{Y} represents the set of probability distributions over outcomes, any risk neutral decision maker has a linear utility function by linearity of expectation.

Remark 2 Linearity and Lipschitzness enter the analysis when converting decision-calibration conditions into utility guarantees, such as in Lemmas 15 and 16.

A natural extension would be to replace y by a feature map $\phi(y)$ and require utilities to be linear in $\phi(y)$. This would change the prediction problem and the resulting bounds, and is therefore outside the scope of this paper.

We take the role of an online/sequential forecaster producing predictions that will be consumed by agents. We consider the following repeated interaction between a forecaster, agents, and an adversary. In every round $t \in [T]$:

- (1) The adversary selects a feature vector $x_t \in \mathcal{X}$ and a distribution over outcomes $Y_t \in \Delta \mathcal{Y}$;

3. Sometimes online adversarial learning problems are described by an adversary choosing a different utility and/or constraint function at each step. This is equivalent to having a fixed state-dependent utility function/constraint functions, and having an adversary choose state.

- (2) The forecaster observes the feature x_t , produces a distribution over predictions $\pi_t \in \Delta\mathcal{Y}$, from which a prediction $p_t \in \mathcal{Y}$ is sampled;
- (3) Each agent chooses an action a_t as a function of the prediction p_t and the history;
- (4) The adversary reveals an outcome $y_t \sim Y_t$, and the agent obtains utility $u(a_t, y_t)$ and the constraint loss vector $\{c_j(a_t, y_t)\}_{j \in [J]}$.

Agents aim to maximize their cumulative utilities over the T rounds: $\sum_{t=1}^T u(a_t, y_t)$ while minimizing their cumulative constraint violation (CCV) over T rounds:

$$\text{CCV}(1 : T) := \max_{j \in [J]} \sum_{t=1}^T c_j(a_t, y_t) \leq o(T).$$

We treat utility maximization as an objective and cumulative constraint violation as a requirement: CCV must be sublinear in T for every constraint.

We compare the utility of actions chosen based on our predictions against the utility of actions from different benchmark classes. A standard benchmark class from the literature on learning with long term constraints is the set of actions that satisfy the *realized* constraints at every round, denoted:

$$\mathcal{A}_{1:T}^c := \{a : c_j(a, y_t) \leq 0 \text{ for every } t \in [T] \text{ and } j \in [J]\}.$$

A somewhat more permissive benchmark class is the set of actions that satisfy the constraints *in expectation* every round, denoted:

$$\mathcal{A}_{1:T}^{\mathbb{E}[c]} := \{a : \mathbb{E}_{y_t \sim Y_t} [c_j(a, y_t)] \leq 0 \text{ for every } t \in [T] \text{ and } j \in [J]\}.$$

We note that these benchmark classes are agent-specific as they depend on agents' constraint functions. For a given benchmark class, we will compete with the best fixed action in hindsight from the class. We measure our performance using a notion of regret, which we call *constrained external regret*.

Definition 3 (\mathcal{B} -Constrained External Regret) Fix an agent with action space \mathcal{A} and utility function u . Let $\mathcal{B} \subseteq \mathcal{A}$ be a benchmark class of actions. For a sequence of actions a_1, \dots, a_T and outcomes y_1, \dots, y_T , the agent's \mathcal{B} -constrained external regret against the benchmark class \mathcal{B} is:

$$\text{Reg}_{\text{ext}}(u, \mathcal{B}, 1 : T) := \max_{a \in \mathcal{B}} \sum_{t=1}^T (u(a, y_t) - u(a_t, y_t)).$$

We will also compete with a more demanding benchmark based on action modification rules. For a given benchmark class \mathcal{B} , an action modification rule is any function $\phi : \mathcal{A} \rightarrow \mathcal{B}$ that consistently maps an agent's actions to alternatives within the benchmark class. Competing with the best such rule in hindsight leads to the notion of *constrained swap regret*. This is a stronger notion than constrained external regret, as constrained external regret can be viewed as a special case where the action modification rule is restricted to being a constant function.

Definition 4 (\mathcal{B} -Constrained Swap Regret) Fix an agent with action space \mathcal{A} and utility function u . Let $\mathcal{B} \subseteq \mathcal{A}$ be a benchmark class of actions. For a sequence of actions a_1, \dots, a_T and outcomes y_1, \dots, y_T , the agent's \mathcal{B} -constrained swap regret against the benchmark class \mathcal{B} is:

$$\text{Reg}_{\text{swap}}(u, \mathcal{B}, 1 : T) := \max_{\phi : \mathcal{A} \rightarrow \mathcal{B}} \sum_{t=1}^T (u(\phi(a_t), y_t) - u(a_t, y_t)).$$

All guarantees we provide in this paper, such as regret bounds, will be stated under the assumption that the corresponding benchmark class is non-empty.

3. Elimination-Based Algorithms and Cumulative Constraint Violation Bounds

In the literature on unconstrained omniprediction, the typical decision rule is the best response to the predictions, also known as the optimal post-processing of the predictions (Gopalan et al., 2022, 2023a,b; Globus-Harris et al., 2023; Garg et al., 2024; Gopalan et al., 2024; Okoroafor et al., 2025; Hu et al., 2025; Lu et al., 2025). This decision rule treats the predictions as accurate and chooses the action to maximize the agent’s utility.

Definition 5 (Best Response) Fix a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and a prediction $p \in \mathcal{Y}$. The best response to p according to u is the action $\text{BR}^u(p) = \arg \max_{a \in \mathcal{A}} u(a, p)$.

In the constrained setting, one would hope to avoid actions that (significantly) violate the constraints, and choose only from the feasible actions. If the agent knew the benchmark class $\mathcal{A}_{1:T}^c$, they could play a constrained version of the best response, $\arg \max_{a \in \mathcal{A}_{1:T}^c} u(a, p)$, i.e., the action that maximizes the predicted utility among $\mathcal{A}_{1:T}^c$. But the agents do not know what this benchmark class is during the interaction (it isn’t defined until the end of the interaction), and so we let each agent maintain a candidate action set $\widehat{\mathcal{A}}_t^c$, which is updated on every round $t \in [T]$ and serves as a surrogate for the benchmark class $\mathcal{A}_{1:T}^c$. The agent will play $\arg \max_{a \in \widehat{\mathcal{A}}_t^c} u(a, p)$, the action that maximizes the predicted utility among their maintained candidate action set $\widehat{\mathcal{A}}_t^c$. Similarly, when competing with the other benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$, agents will maintain a surrogate $\widehat{\mathcal{A}}_t^{\mathbb{E}[c]}$, and maximize the predicted utility over actions in $\widehat{\mathcal{A}}_t^{\mathbb{E}[c]}$. In general, we say that agents *constrained best respond* to our predictions.

Definition 6 (C-Constrained Best Response) Let $\mathcal{C} \subseteq \mathcal{A}$ be an action set. Fix a utility $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and a prediction $p \in \mathcal{Y}$. The \mathcal{C} -constrained best response to p according to u is the action $\text{CBR}_{\mathcal{C}}^u(p) = \arg \max_{a \in \mathcal{C}} u(a, p)$.

It is straightforward for an agent to identify non-members of the benchmark class $\mathcal{A}_{1:T}^c$ and maintain the candidate action set $\widehat{\mathcal{A}}_t^c$. At the beginning, there is no information about the benchmark class $\mathcal{A}_{1:T}^c$ and the agent can set $\widehat{\mathcal{A}}_1^c = \mathcal{A}$. In every round $t \in [T]$, if any action $a \in \mathcal{A}$ violates any of the J constraints, a is known not to be in the benchmark class $\mathcal{A}_{1:T}^c$, and hence can be permanently eliminated from the candidate action set $\widehat{\mathcal{A}}_t^c$. This process is formally stated in Algorithm 1.

Algorithm 1 Elimination-Based Algorithm for Benchmark Class $\mathcal{A}_{1:T}^c$

```

Initialize  $\widehat{\mathcal{A}}_1^c = \mathcal{A}$ 
for  $t = 1$  to  $T$  do
    Receive the prediction  $p_t$ 
    Play the  $\widehat{\mathcal{A}}_t^c$ -constrained best response to  $p_t$ ,  $a_t = \text{CBR}_{\widehat{\mathcal{A}}_t^c}^u(p_t)$ 
    Observe the outcome  $y_t$ , receive utility  $u(a_t, y_t)$  and constraint realizations  $(c_j(a_t, y_t))_{j \in [J]}$ 
    Let  $\widehat{\mathcal{A}}_{t+1}^c = \widehat{\mathcal{A}}_t^c \setminus \{a \in \widehat{\mathcal{A}}_t^c : \exists j \in [J], c_j(a, y_t) > 0\}$ 
end

```

By design, the elimination operation in Algorithm 1 is only triggered by actions that violate a constraint. Since actions in the benchmark class $\mathcal{A}_{1:T}^c$ satisfy the constraints every round, they will never be eliminated from the candidate action set $\widehat{\mathcal{A}}_t^c$ by Algorithm 1.

Proposition 7 Fix an agent in \mathcal{N} with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Suppose the agent runs Algorithm 1 to compete with the benchmark class $\mathcal{A}_{1:T}^c$, then $\mathcal{A}_{1:T}^c \subseteq \widehat{\mathcal{A}}_t^c$ for any $t \in [T]$.

Because an action is permanently eliminated upon violating the constraints, each constraint violation of Algorithm 1 must be triggered by a distinct action. Consequently, the cumulative constraint violation of Algorithm 1 can be upper bounded by the number of actions.

Theorem 8 Fix an agent in \mathcal{N} with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Suppose the agent runs Algorithm 1 to compete with the benchmark class $\mathcal{A}_{1:T}^c$, then the cumulative constraint violation of the agent is bounded by:

$$\text{CCV}(1 : T) \leq |\mathcal{A}|.$$

Remark 9 In fact, our proof of Theorem 8 leads to a stronger result, $\text{CCV}^+(1 : T) := \max_{j \in [J]} \sum_{t=1}^T c_j^+(a_t, y_t) \leq |\mathcal{A}| - |\mathcal{A}_{1:T}^c|$, where $c_j^+(a_t, y_t) := \max\{0, c_j(a_t, y_t)\}$. The right-hand side can be strengthened because actions in the benchmark class $\mathcal{A}_{1:T}^c$ never trigger any constraint violation. The left-hand side can be strengthened because our proof bounds any non-positive term $c_j(a, y_t)$ by 0. CCV^+ is a common definition of cumulative constraint violation in the literature on learning with long-term constraints (see e.g. [Sinha and Vaze \(2024\)](#); [Lekeuffack and Jordan \(2024\)](#)).

Next, we introduce our elimination algorithm for the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$. The identification of non-members is straightforward for the benchmark class $\mathcal{A}_{1:T}^c$ but not for $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$. Since $c_j(a_t, y_t)$ and $\mathbb{E}_{y_t \sim Y_t}[c_j(a_t, y_t)]$ can have different signs, it is impossible to decide whether an action is in $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$ based on its performance in a single round. Therefore, we will let agents eliminate actions by assessing their performances over multiple rounds. To be specific, for each action $a \in \mathcal{A}$ and each constraint c_j , we let $v_{1:t}^{a,j}$ denote the total violation of c_j accumulated by the agent from playing action a up to round t , i.e.,

$$v_{1:t}^{a,j} := \sum_{s=1}^t \mathbb{1}[a_s = a] c_j(a, y_s).$$

As we will show in the proof of Theorem 10, $v_{1:t}^{a,j}$ will concentrate around its ‘expectation version’, $\sum_{s=1}^t \mathbb{1}[a_s = a] \mathbb{E}_{y_s \sim Y_s}[c_j(a, y_s)]$, which is non-positive for any benchmark action $a \in \mathcal{A}_{1:T}^{\mathbb{E}[c]}$ and is potentially positive for any non-benchmark action $a \notin \mathcal{A}_{1:T}^{\mathbb{E}[c]}$. So, for each action $a \in \mathcal{A}$, we will let agents monitor $(v_{1:t}^{a,j})_{j \in [J]}$ in every round and eliminate action a once any of $(v_{1:t}^{a,j})_{j \in [J]}$ exceeds a prefixed threshold τ . This process is formally stated in Algorithm 2. We will give guarantees for Algorithm 2 that hold with high probability, where the probability is determined by τ .

Algorithm 2 never eliminates any benchmark action in $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$ from the candidate action set $\mathcal{A}_t^{\mathbb{E}[c]}$ across the T rounds with high probability over the randomness of the outcomes y_1, \dots, y_T .

Proposition 10 Fix an agent in \mathcal{N} with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Suppose the agent runs Algorithm 2 with threshold parameter τ to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$, then with probability at least $1 - 2|\mathcal{A}|JT e^{-\frac{\tau^2}{8T}}$:

$$\mathcal{A}_{1:T}^{\mathbb{E}[c]} \subseteq \widehat{\mathcal{A}}_t^{\mathbb{E}[c]} \text{ for every } t \in [T].$$

Algorithm 2 Elimination-Based Algorithm for Benchmark Class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$

Input: A threshold parameter τ

Initialize $\widehat{\mathcal{A}}_1^{\mathbb{E}[c]} = \mathcal{A}$

For every $a \in \mathcal{A}$ and $j \in [J]$, set $v_0^{a,j} = 0$

for $t = 1$ **to** T **do**

 Receive the prediction p_t

 Play the $\widehat{\mathcal{A}}_t^{\mathbb{E}[c]}$ -constrained best response to p_t , $a_t = \text{CBR}_{\widehat{\mathcal{A}}_t^{\mathbb{E}[c]}}^u(p_t)$

 Observe the outcome y_t , obtain utility $u(a_t, y_t)$ and constraint increments $c_j(a_t, y_t)$, $j \in [J]$

 Update $v_{1:t}^{a_t,j} = v_{1:t-1}^{a_t,j} + c_j(a_t, y_t)$ for every $j \in [J]$

 Let $v_{1:t}^{a,j} = v_{1:t-1}^{a,j}$ for every $a \in \mathcal{A} \setminus \{a_t\}$ and $j \in [J]$

if $\exists j \in [J], v_t^{a_t,j} > \tau$ **then**

 | $\widehat{\mathcal{A}}_{t+1}^{\mathbb{E}[c]} = \widehat{\mathcal{A}}_t^{\mathbb{E}[c]} \setminus \{a_t\}$

end

else

 | $\widehat{\mathcal{A}}_{t+1}^{\mathbb{E}[c]} = \widehat{\mathcal{A}}_t^{\mathbb{E}[c]}$

end

end

Proposition 10 provides a guarantee for a single agent. To extend this to all agents in \mathcal{N} , we apply a union bound. By setting the threshold $\tau = 4\sqrt{T \ln(|\mathcal{A}||\mathcal{N}|JT/\delta)}$ for a desired failure probability $\delta \in (0, 1)$, we arrive at the following corollary, which holds with probability at least $1 - \delta$.

Corollary 11 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent runs Algorithm 2 with threshold parameter $4\sqrt{T \ln(|\mathcal{A}||\mathcal{N}|JT/\delta)}$ to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$, then with probability at least $1 - \delta$, for every agent:*

$$\mathcal{A}_{1:T}^{\mathbb{E}[c]} \subseteq \widehat{\mathcal{A}}_t^{\mathbb{E}[c]} \text{ for every } t \in [T].$$

Suppose an action is eliminated by Algorithm 2 in round t . Its contribution to the cumulative constraint violation of the agent must be smaller than τ before round t , otherwise it would be eliminated before round t . In round t , it violates the constraints by at most 1. Therefore, the agent accumulates at most $\tau + 1$ constraint violation when playing a specific action. This fact leads to a bound on the cumulative constraint violation of Algorithm 2.

Theorem 12 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent runs Algorithm 2 with threshold parameter $4\sqrt{T \ln(|\mathcal{A}||\mathcal{N}|JT/\delta)}$ to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$, then with probability at least $1 - \delta$, the cumulative constraint violation of every agent is bounded by:*

$$\text{CCV}(1 : T) \leq 4|\mathcal{A}|\sqrt{T \ln(|\mathcal{A}||\mathcal{N}|JT/\delta)} + |\mathcal{A}|.$$

We remark that both Algorithms 1 and 2 are run by agents. While they take the prediction p_t as input in every round, their cumulative constraint violation bounds in Theorems 8 and 12 are proven

without any assumptions on the predictions. These bounds are consequences of the elimination rules and hold true for any arbitrary forecasters.

4. Conditionally Unbiased Predictions and Regret Bounds

In this section, we show how to produce predictions that lead to $\tilde{O}(|\mathcal{A}|\sqrt{T})$ regret against the benchmark class $\mathcal{A}_{1:T}^c$. Guarantees for the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$ with the same dependence on $|\mathcal{A}|$ and T can be derived similarly and are relegated to Appendix C.

While the cumulative constraint violation bounds of Algorithms 1 and 2 do not rely on the quality of the predictions, to obtain the regret bounds, we need the predictions to be unbiased — conditionally on various subsequences. Notably, we define the following notion of conditional unbiasedness.

Definition 13 ((\mathcal{N}, α_1) -Decision Calibration) *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent runs Algorithm 1 to maintain a candidate action set $\hat{\mathcal{A}}_t^c$ and $\hat{\mathcal{A}}_t^c$ -constrained best respond to the prediction p_t in every round $t \in [T]$. Let $\alpha_1 : \mathbb{R} \rightarrow \mathbb{R}$. We say that a sequence of predictions p_1, \dots, p_T is (\mathcal{N}, α_1) -decision calibrated with respect to a sequence of outcomes y_1, \dots, y_T if for every $a \in \mathcal{A}$ and $(u, \mathbf{c}) \in \mathcal{N}$:*

$$\left\| \sum_{t=1}^T \mathbb{1}[\text{CBR}_{\hat{\mathcal{A}}_t^c}^u(p_t) = a](p_t - y_t) \right\|_\infty \leq \alpha_1(T^{u, \mathbf{c}}(a)).$$

where $T^{u, \mathbf{c}}(a) = \sum_{t=1}^T \mathbb{1}[\text{CBR}_{\hat{\mathcal{A}}_t^c}^u(p_t) = a]$.

Assumption 2 *We assume that $\alpha_1 : \mathbb{R} \rightarrow \mathbb{R}$ is a concave function. This will be the case in all the bounds we give; in general, this condition holds for any sublinear error bound T^r for $r < 1$.*

We now show that decision calibration implies no constrained swap regret, and hence no constrained external regret. At a high level, decision calibrated predictions allow agents to accurately assess (on average) the utilities of their chosen actions and the counterfactual actions produced by any action modification rule. Because agents choose optimally based on these accurate utility estimates, their decisions are inherently competitive against any alternative benchmark.

Theorem 14 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent runs Algorithm 1 to compete with the benchmark class $\mathcal{A}_{1:T}^c$. If the sequence of predictions p_1, \dots, p_T is (\mathcal{N}, α_1) -decision calibrated, then the $\mathcal{A}_{1:T}^c$ -constrained swap regret of every agent is bounded by:*

$$\text{Reg}_{\text{swap}}(u, \mathcal{A}_{1:T}^c, 1 : T) \leq 2L|\mathcal{A}|\alpha_1(T/|\mathcal{A}|).$$

To prove the theorem, first note that for any action modification rule $\phi : \mathcal{A} \rightarrow \mathcal{A}_{1:T}^c$, we can decompose the regret against ϕ into three parts as:

$$\begin{aligned} & \sum_{t=1}^T (u(\phi(a_t), y_t) - u(a_t, y_t)) \\ &= \sum_{t=1}^T (u(\phi(a_t), y_t) - u(\phi(a_t), p_t)) + \sum_{t=1}^T (u(\phi(a_t), p_t) - u(a_t, p_t)) + \sum_{t=1}^T (u(a_t, p_t) - u(a_t, y_t)) \end{aligned}$$

By Proposition 7 and the fact that $\phi(a_t) \in \mathcal{A}_{1:T}^c$, we have that $\phi(a_t) \in \widehat{\mathcal{A}}_t^c$ for every $t \in [T]$. Since agents $\widehat{\mathcal{A}}_t^c$ -constrained best respond to the prediction p_t in round t , $a_t = \arg \max_{b \in \widehat{\mathcal{A}}_t^c} u(b, p_t)$. This implies that:

$$u(\phi(a_t), p_t) \leq u(a_t, p_t).$$

So, the second part in the decomposition is upper bounded by 0.

Thus, it suffices to bound the other two parts, i.e., the difference in utility under our predictions p_t and the outcomes y_t for both the chosen actions and the swapped-in actions. We show this in the next two lemmas using decision calibration.

Lemma 15 *If the sequence of predictions p_1, \dots, p_T is (\mathcal{N}, α_1) -decision calibrated, then for any $(u, \mathbf{c}) \in \mathcal{N}$:*

$$\left| \sum_{t=1}^T (u(a_t, p_t) - u(a_t, y_t)) \right| \leq L|\mathcal{A}|\alpha_1(T/|\mathcal{A}|).$$

Lemma 16 *If the sequence of predictions p_1, \dots, p_T is (\mathcal{N}, α_1) -decision calibrated, then for any $(u, \mathbf{c}) \in \mathcal{N}$ and $\phi : \mathcal{A} \rightarrow \mathcal{A}_{1:T}^c$:*

$$\left| \sum_{t=1}^T (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| \leq L|\mathcal{A}|\alpha_1(T/|\mathcal{A}|).$$

Applying Lemmas 15 and 16 to the decomposition of the regret against ϕ finishes the proof.

4.1. Algorithm for Calibration and Decision Calibration

We now address the algorithmic challenge of producing decision calibrated predictions. Our approach builds upon the UNBIASED-PREDICTION algorithm from Noarov et al. (2023), which makes conditionally unbiased predictions in the online setting. The algorithm and its guarantees are presented in Appendix G; we refer interested readers to the original work for further details.

We will instantiate UNBIASED-PREDICTION to make decision calibrated predictions; we will refer to this instantiation as DECISION-CALIBRATION-REALIZATION. Our guarantees will directly inherit from the guarantees of UNBIASED-PREDICTION.

Theorem 17 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent will run Algorithm 1 to compete with the benchmark class $\mathcal{A}_{1:T}^c$. There is an instantiation of UNBIASED-PREDICTION (Noarov et al., 2023) —which we call DECISION-CALIBRATION-REALIZATION—producing predictions $p_1, \dots, p_T \in \mathcal{Y}$ satisfying that for any sequence of outcomes $y_1, \dots, y_T \in \mathcal{Y}$, the following holds with probability at least $1 - \delta$ for any $(u, \mathbf{c}) \in \mathcal{N}$ and $a \in \mathcal{A}$:*

$$\left\| \sum_{t=1}^T \mathbb{1}[\text{CBR}_{\widehat{\mathcal{A}}_t^c}^u(p_t) = a](p_t - y_t) \right\|_\infty \leq O\left(\ln(d|\mathcal{A}||\mathcal{N}|T) + \sqrt{\ln(d|\mathcal{A}||\mathcal{N}|T) \cdot T^{u, \mathbf{c}}(a)} + \sqrt{\ln(d|\mathcal{A}||\mathcal{N}|/\delta) \cdot T}\right).$$

Substituting the above bounds into Theorem 14 and combining with Theorem 8, we arrive at the following corollary bounding the cumulative constraint violation and the $\mathcal{A}_{1:T}^c$ -constrained swap regret.

Corollary 18 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent runs Algorithm 1 to compete with the benchmark class $\mathcal{A}_{1:T}^c$. The sequence of predictions p_1, \dots, p_T outputted by DECISION-CALIBRATION-REALIZATION ensures that with probability at least $1 - \delta$, the cumulative constraint violation and the $\mathcal{A}_{1:T}^c$ -constrained swap regret of every agent are bounded by:*

$$\begin{aligned} \text{CCV}(1 : T) &\leq |\mathcal{A}|, \\ \text{Reg}_{\text{swap}}(u, \mathcal{A}_{1:T}^c, 1 : T) &\leq O(L|\mathcal{A}|\sqrt{T \ln(d|\mathcal{A}||\mathcal{N}|T/\delta)}). \end{aligned}$$

Since constrained swap regret is stronger than constrained external regret, the same regret bound also holds for $\text{Reg}_{\text{ext}}(u, \mathcal{A}_{1:T}^c, 1 : T)$.

5. Simultaneous Guarantees for Multiple Subsequences

We now summarize how our framework extends to arbitrary, possibly intersecting, contextually defined subsequences. The full details are deferred to Appendix D.

Let \mathcal{S} be a collection of subsequences, where each subsequence $S \in \mathcal{S}$ is a subset of $[T]$. We denote the active subsequences in round $t \in [T]$ as $\mathcal{S}_t = \{S \in \mathcal{S} : t \in S\}$. We assume that $\cup_{S \in \mathcal{S}} S = [T]$, so that \mathcal{S}_t is nonempty for every round t . These subsequences need not be fixed in advance but can be defined dynamically based on contextual information.

For each subsequence $S \in \mathcal{S}$, the relevant benchmark classes are subsequence-specific:

$$\begin{aligned} \mathcal{A}_S^c &= \{a : c_j(a, y_t) \leq 0 \text{ for every } t \in S \text{ and } j \in [J]\}, \\ \mathcal{A}_S^{\mathbb{E}[c]} &= \left\{ a : \mathbb{E}_{y_t \sim Y_t} [c_j(a, y_t)] \leq 0 \text{ for every } t \in S \text{ and } j \in [J] \right\}. \end{aligned}$$

Regret over a subsequence $S \in \mathcal{S}$ with respect to a benchmark class \mathcal{B} is defined as:

$$\begin{aligned} \text{Reg}_{\text{ext}}(u, \mathcal{B}, S) &= \max_{a \in \mathcal{B}} \sum_{t \in S} (u(a, y_t) - u(a_t, y_t)), \\ \text{Reg}_{\text{swap}}(u, \mathcal{B}, S) &= \max_{\phi: \mathcal{A} \rightarrow \mathcal{B}} \sum_{t \in S} (u(\phi(a_t), y_t) - u(a_t, y_t)). \end{aligned}$$

The cumulative constraint violation over S is:

$$\text{CCV}(S) := \max_{j \in [J]} \sum_{t \in S} c_j(a_t, y_t).$$

The goal is to obtain regret and cumulative-constraint-violation guarantees simultaneously for every agent and every $S \in \mathcal{S}$, with the benchmark tailored to the particular subsequence.

For both benchmark classes, the forecaster enforces decision calibration not only conditional on the agent's chosen action, but also conditional on membership of the current round in each subsequence $S \in \mathcal{S}$.

Realization-based benchmark classes $\{\mathcal{A}_S^c\}_{S \in \mathcal{S}}$. For the benchmark classes $\{\mathcal{A}_S^c\}_{S \in \mathcal{S}}$, the extension is direct. Each agent runs separate instantiations of Algorithm 1 to maintain a candidate action set $\widehat{\mathcal{A}}_{t,S}^c$ for each subsequence S . At round t , the agent aggregates the active candidate sets by taking the union

$$U_t = \cup_{S \in \mathcal{S}_t} \widehat{\mathcal{A}}_{t,S}^c,$$

and plays the U_t -constrained best response to the prediction.

This yields, with probability at least $1 - \delta$, simultaneously for every agent and every $S \in \mathcal{S}$,

$$\begin{aligned} \text{CCV}(S) &\leq |\mathcal{A}||\mathcal{S}|, \\ \text{Reg}_{\text{swap}}(u, \mathcal{A}_S^c, S) &\leq O\left(L|\mathcal{A}|\sqrt{|S| \ln \frac{d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T}{\delta}}\right), \end{aligned}$$

with the same regret bound for constrained external regret.

Expectation-based benchmark classes $\{\mathcal{A}_S^{\mathbb{E}[c]}\}_{S \in \mathcal{S}}$. The benchmark classes $\{\mathcal{A}_S^{\mathbb{E}[c]}\}_{S \in \mathcal{S}}$ require more care. Naively running independent copies of Algorithm 2 for each subsequence is not enough. Because constraint violations are signed, slack on $S \setminus S'$ can mask large positive violation on $S \cap S'$. Thus an action may remain feasible from the perspective of one subsequence while causing large cumulative violation on another intersecting subsequence.

Our solution has the same high-level structure as the solution for the realization-based benchmark classes $\{\mathcal{A}_S^c\}_{S \in \mathcal{S}}$. We maintain a candidate action set $\widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]}$ for each subsequence and gradually eliminate actions from it; the agent selects an action from the union set, $U_t = \cup_{S \in \mathcal{S}_t} \widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]}$, in each round. The key difference lies in a more sophisticated elimination rule.

To avoid the masking effect, we attribute each played action to a single responsible subsequence. Index $\mathcal{S} = \{S_1, \dots, S_I\}$. When action a_t is played at round t , we credit the subsequence S_{r_t} with the smallest index that is active and for which a_t is a candidate action:

$$r_t = \min \left\{ i \in [I] : t \in S_i, a_t \in \widehat{\mathcal{A}}_{t,S_i}^{\mathbb{E}[c]} \right\}.$$

For every action a , constraint j , and pair of subsequences $S, S' \in \mathcal{S}$, we track the attributed violation accumulated on S during rounds for which S' is responsible,

$$w_{[1,t] \cap S}^{a,j,S'} = \sum_{s=1}^t \mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] c_j(a, y_s).$$

An action a_t is eliminated from its responsible subsequence S_{r_t} once any of these attributed terms exceeds a subsequence-specific threshold. This rule prevents double-counting while ensuring that violations on intersections cannot be masked by slack elsewhere.

This yields, with probability at least $1 - \delta$, simultaneously for every agent and every $S \in \mathcal{S}$,

$$\begin{aligned} \text{CCV}(S) &\leq \sum_{S' \in \mathcal{S}} 4|\mathcal{A}|\sqrt{|S'| \ln \frac{2|\mathcal{A}||\mathcal{N}||\mathcal{S}|^2 J|S'|}{\delta}} + |\mathcal{A}||\mathcal{S}|, \\ \text{Reg}_{\text{swap}}(u, \mathcal{A}_S^{\mathbb{E}[c]}, S) &\leq O\left(L|\mathcal{A}|\sqrt{|S| \ln \frac{d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T}{\delta}}\right), \end{aligned}$$

again with the same regret bound for constrained external regret.

References

- Daron Anderson, George Iosifidis, and Douglas J Leith. Lazy lagrangians with predictions for online learning. *arXiv preprint arXiv:2201.02890*, 2022.
- Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.
- Eshwar Ram Arunachaleswaran, Natalie Collina, and Jon Schneider. Pareto-optimal algorithms for learning in games. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 490–510, 2024.
- Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8(6), 2007.
- Mark Braverman, Jieming Mao, Jon Schneider, and Matt Weinberg. Selling to a no-regret buyer. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 523–538, 2018.
- Matteo Castiglioni, Andrea Celli, Alberto Marchesi, Giulia Romano, and Nicola Gatti. A unifying framework for online optimization with long-term constraints. In Sanmi Koyejo, S. Mohamed, A. Agarwal, Danielle Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, 2022.
- Nicolò Cesa-Bianchi, Yoav Freund, David Haussler, David P. Helmbold, Robert E. Schapire, and Manfred K. Warmuth. How to use expert advice. *J. ACM*, 44(3):427–485, May 1997.
- Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, Noah Golowich, Robert Kleinberg, and Princewill Okoroafor. Improved bounds for calibration via stronger sign preservation games. *arXiv preprint arXiv:2406.13668*, 2024.
- Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Strategizing against no-regret learners. *Advances in neural information processing systems*, 32, 2019.
- Cynthia Dwork, Chris Hays, Nicole Immorlica, Juan C. Perdomo, and Pranay Tankala. From fairness to infinity: Outcome-indistinguishable (omni)prediction in evolving graphs. In Nika Haghtalab and Ankur Moitra, editors, *The Thirty Eighth Annual Conference on Learning Theory, 30-4 July 2025, Lyon, France*, volume 291 of *Proceedings of Machine Learning Research*, pages 1564–1637. PMLR, 2025.
- Sumegha Garg, Christopher Jung, Omer Reingold, and Aaron Roth. Oracle efficient online multi-calibration and omniprediction. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2725–2792. SIAM, 2024.
- Ira Globus-Harris, Varun Gupta, Christopher Jung, Michael Kearns, Jamie Morgenstern, and Aaron Roth. Multicalibrated regression for downstream fairness. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pages 259–286, 2023.

- Ira Globus-Harris, Declan Harrison, Michael Kearns, Aaron Roth, and Jessica Sorrell. Multicalibration as boosting for regression. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, volume 202 of *Proceedings of Machine Learning Research*, pages 11459–11492. PMLR, 2023.
- Parikshit Gopalan, Adam Tauman Kalai, Omer Reingold, Vatsal Sharan, and Udi Wieder. Omnipredictors. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPICs*, pages 79:1–79:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.
- Parikshit Gopalan, Lunjia Hu, Michael P. Kim, Omer Reingold, and Udi Wieder. Loss minimization through the lens of outcome indistinguishability. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPICs*, pages 60:1–60:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023a.
- Parikshit Gopalan, Michael P. Kim, and Omer Reingold. Swap agnostic learning, or characterizing omniprediction via multicalibration. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine, editors, *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023b.
- Parikshit Gopalan, Princewill Okoroafor, Prasad Raghavendra, Abhishek Sherry, and Mihir Singhal. Omnipredictors for regression and the approximate rank of convex functions. In Shipra Agrawal and Aaron Roth, editors, *The Thirty Seventh Annual Conference on Learning Theory, June 30 - July 3, 2023, Edmonton, Canada*, volume 247 of *Proceedings of Machine Learning Research*, pages 2027–2070. PMLR, 2024.
- Hengquan Guo, Xin Liu, Honghao Wei, and Lei Ying. Online convex optimization with hard constraints: Towards the best of two worlds and beyond. In Sanmi Koyejo, S. Mohamed, A. Agarwal, Danielle Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, 2022.
- James Hannan. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games*, 3(2):97–139, 1957.
- Úrsula Hébert-Johnson, Michael P. Kim, Omer Reingold, and Guy N. Rothblum. Multicalibration: Calibration for the (computationally-identifiable) masses. In Jennifer G. Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 1944–1953. PMLR, 2018.
- Lunjia Hu and Yifan Wu. Predict to minimize swap regret for all payoff-bounded tasks. In *65th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2024, Chicago, IL, USA, October 27-30, 2024*, pages 244–263. IEEE, 2024.

- Lunjia Hu, Inbal Rachel Livni Navon, Omer Reingold, and Chutong Yang. Omnipredictors for constrained optimization. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, volume 202 of *Proceedings of Machine Learning Research*, pages 13497–13527. PMLR, 2023.
- Lunjia Hu, Kevin Tian, and Chutong Yang. Omnipredicting single-index models with multi-index models. In Michal Koucký and Nikhil Bansal, editors, *Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC 2025, Prague, Czechia, June 23-27, 2025*, pages 1762–1773. ACM, 2025.
- Bobby Kleinberg, Renato Paes Leme, Jon Schneider, and Yifeng Teng. U-calibration: Forecasting for an unknown agent. In Gergely Neu and Lorenzo Rosasco, editors, *The Thirty Sixth Annual Conference on Learning Theory, COLT 2023, 12-15 July 2023, Bangalore, India*, volume 195 of *Proceedings of Machine Learning Research*, pages 5143–5145. PMLR, 2023.
- Jordan Lekeufack and Michael I Jordan. An optimistic algorithm for online convex optimization with adversarial constraints. *arXiv preprint arXiv:2412.08060*, 2024.
- Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Inf. Comput.*, 108(2): 212–261, 1994.
- Jiuyao Lu, Aaron Roth, and Mirah Shi. Sample efficient omniprediction and downstream swap regret for non-linear losses. In Nika Haghtalab and Ankur Moitra, editors, *The Thirty Eighth Annual Conference on Learning Theory, 30-4 July 2025, Lyon, France*, volume 291 of *Proceedings of Machine Learning Research*, pages 3829–3878. PMLR, 2025.
- Shie Mannor, John N. Tsitsiklis, and Jia Yuan Yu. Online learning with sample path constraints. *J. Mach. Learn. Res.*, 10:569–590, 2009.
- Michael J Neely and Hao Yu. Online convex optimization with time-varying constraints. *arXiv preprint arXiv:1702.04783*, 2017.
- Georgy Noarov, Ramya Ramalingam, Aaron Roth, and Stephan Xie. High-dimensional prediction for sequential decision making. In *International Conference on Machine Learning, ICML 2025, 13-19 July 2025, Vancouver, Canada*, 2023.
- Princewill Okoroafor, Robert Kleinberg, and Michael P Kim. Near-optimal algorithms for omniprediction. *arXiv preprint arXiv:2501.17205*, 2025.
- Mingda Qiao and Gregory Valiant. Stronger calibration lower bounds via sidestepping. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 456–466. ACM, 2021.
- Shuang Qiu, Xiaohan Wei, and Mladen Kolar. Gradient-variation bound for online convex optimization with constraints. In Brian Williams, Yiling Chen, and Jennifer Neville, editors, *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational*

- Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, pages 9534–9542. AAAI Press, 2023.
- Aaron Roth and Mirah Shi. Forecasting for swap regret for all downstream agents. In Dirk Berge-
mann, Robert Kleinberg, and Daniela Sabán, editors, *Proceedings of the 25th ACM Conference
on Economics and Computation, EC 2024, New Haven, CT, USA, July 8-11, 2024*, pages 466–
488. ACM, 2024.
- Aviad Rubinstein and Junyao Zhao. Strategizing against no-regret learners in first-price auctions.
In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 894–921,
2024.
- Abhishek Sinha and Rahul Vaze. Optimal algorithms for online convex optimization with adver-
sarial constraints. In Amir Globersons, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich
Paquet, Jakub M. Tomczak, and Cheng Zhang, editors, *Advances in Neural Information Process-
ing Systems 38: Annual Conference on Neural Information Processing Systems 2024, NeurIPS
2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024.
- Wen Sun, Debadepta Dey, and Ashish Kapoor. Safety-aware algorithms for adversarial contextual
bandit. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Con-
ference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages
3280–3288. PMLR, 06–11 Aug 2017.
- Xiaohan Wei, Hao Yu, and Michael J. Neely. Online primal-dual mirror descent under stochastic
constraints. *Proc. ACM Meas. Anal. Comput. Syst.*, 4(2), jun 2020.
- Xinlei Yi, Xiuxian Li, Tao Yang, Lihua Xie, Yiguang Hong, Tianyou Chai, and Karl H Johans-
son. Distributed online convex optimization with adversarial constraints: Reduced cumulative
constraint violation bounds under slater’s condition. *arXiv preprint arXiv:2306.00149*, 2023.
- Hao Yu and Michael J. Neely. A low complexity algorithm with $o(\sqrt{T})$ regret and $o(1)$ constraint vi-
olations for online convex optimization with long term constraints. *Journal of Machine Learning
Research*, 21(1):1–24, 2020.
- Shengjia Zhao, Michael P. Kim, Roshni Sahoo, Tengyu Ma, and Stefano Ermon. Calibrating pre-
dictions to decisions: A novel approach to multi-class calibration. In Marc’Aurelio Ranzato,
Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan, editors, *Ad-
vances in Neural Information Processing Systems 34: Annual Conference on Neural Information
Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pages 22313–22324,
2021.

Appendix A. Related Work

Our work extends both past work on *omniprediction* and *prediction for downstream regret* as well as past work on *online learning with long term constraints*. We discuss both strands of related work below.

Omniprediction and Prediction for Downstream Regret Gopalan et al. (2022) introduced the problem of *omniprediction* in the batch/distributional setting. In the traditional “omniprediction” setting the label space is binary $\mathcal{Y} = \{0, 1\}$, the “action space” is considered to be the set of real valued predictions in $[0, 1]$ and “decision makers” are modeled as optimizing a loss function $\ell : [0, 1] \times \{0, 1\} \rightarrow \mathbb{R}$, such as squared or absolute loss. Gopalan et al. (2022) show that multicalibration (c.f. Hébert-Johnson et al. (2018)) implies omniprediction. Subsequent work (Gopalan et al., 2023a,b) studied both weaker and stronger conditions than multicalibration and the corresponding variants of omniprediction that they imply. Garg et al. (2024) introduced the problem of omniprediction in the online adversarial setting and gave oracle efficient algorithms for it; subsequent work derived optimal regret bounds for the problem (Okoroafor et al., 2025; Dwork et al., 2025). All of the papers discussed so far study binary outcomes ($\mathcal{Y} = \{0, 1\}$), but a notable exception is Gopalan et al. (2024) who study vector valued outcomes and decision makers with loss functions that are convex in the outcome. Two papers (Globus-Harris et al., 2023; Hu et al., 2023) study the problem of omniprediction with constraints. Like most of the rest of the literature, they focus on binary outcomes $\mathcal{Y} = \{0, 1\}$ and real valued action spaces $\mathcal{A} = [0, 1]$ in the batch setting. As their focus is on machine learning, they both focus on specific kinds of constraints that arise as “group fairness” constraints in binary classification.

A parallel literature Kleinberg et al. (2023); Noarov et al. (2023); Roth and Shi (2024); Hu and Wu (2024) has focused on making sequential predictions of adversarially chosen outcomes from some space \mathcal{Y} so that downstream agents, who may have arbitrary discrete action spaces, can optimize their own loss. Several papers from this literature Noarov et al. (2023); Roth and Shi (2024) have gone beyond the binary label setting and have focused (as we do) on arbitrary loss functions that for each action are *linear* in the state $y \in [0, 1]^d$ to be predicted. The goal is to guarantee that all downstream decision makers have diminishing regret in the worst case. Motivated by decision makers interacting in competitive environments, Noarov et al. (2023); Roth and Shi (2024); Hu and Wu (2024) give results to guarantee downstream decision makers diminishing *swap* regret at near optimal rates. Because calibration is known to be unobtainable in online adversarial settings at $O(\sqrt{T})$ rates (Qiao and Valiant, 2021; Dagan et al., 2024), this literature has from the beginning used techniques (like “U-calibration” (Kleinberg et al., 2023)) and extensions of decision calibration Zhao et al. (2021); Noarov et al. (2023); Roth and Shi (2024)) that sidestep these lower bounds while being strong enough to guarantee downstream decision makers no (swap) regret. Lu et al. (2025) give a unifying view on both literatures. Our model and techniques derive mainly from the “prediction for downstream regret literature” (arbitrary action space, high dimensional outcome space, swap regret guarantees in an online adversarial setting) — one way to view our work is as extending omniprediction with constraints (Hu et al., 2023; Globus-Harris et al., 2023) to this setting.

Online Learning with Long-Term Constraints The problem of learning with long-term constraints first appeared in Mannor et al. (2009), who showed that even for simple utility functions and a single adaptive constraint, it is impossible for the learner to get sublinear external regret with

respect to the benchmark of actions that satisfy the constraint *marginally*. Instead, follow up work (e.g. Sun et al. (2017); Guo et al. (2022); Anderson et al. (2022); Yi et al. (2023); Sinha and Vaze (2024); Lekeufack and Jordan (2024)) adopted a stricter external regret benchmark, of actions that satisfy the constraints *on every round*. Generally speaking, these works operate in the setting of *on-line convex optimization* rather than considering a finite set of actions. Among these, Sinha and Vaze (2024) obtained the fastest known convergence rates without additional assumptions, scaling with $\tilde{O}(\sqrt{T})$ for each of the regret and the cumulative constraint violation. Lekeufack and Jordan (2024) provided an algorithm for the experts setting with rates scaling each with $\tilde{O}(\sqrt{T \ln(d)})$ where d is the number of experts, and regret is measured against the set of all probability vectors over the experts that (in expectation) satisfy the constraints. This literature also explores benchmarks where constraints are not violated *in expectation* (Neely and Yu, 2017; Wei et al., 2020), in cases where there assumed to exist an underlying distribution incoming constraint functions. This can be viewed analogously to $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$ in our formulation, for the case where the distribution over states also encodes a distribution over constraints that is fixed across the rounds of interaction. Our formulation, however, is more permissive as it allows the distribution over constraints to vary across rounds.

Appendix B. Omitted Results from Section 3

Proof of Theorem 8. Fix any $j \in [J]$. We will bound the cumulative violation against the j -th constraint, i.e., $\sum_{t=1}^T c_j(a_t, y_t)$. The final result follows by taking the maximum over all $j \in [J]$.

First, we partition the sum based on the action played. For each action $a \in \mathcal{A}$, let $R_a = \{t \in [T] : a_t = a\}$ be the set of rounds when a is played. This gives:

$$\sum_{t=1}^T c_j(a_t, y_t) = \sum_{a \in \mathcal{A}: R_a \neq \emptyset} \sum_{t \in R_a} c_j(a, y_t).$$

$\max R_a$ corresponds to the last round when a is played. For any earlier round $t \in R_a \setminus \{\max R_a\}$, action a is played and is not eliminated, which implies that the constraint is satisfied, i.e., $c_j(a, y_t) \leq 0$. Therefore, only the final term, $c_j(a, y_{\max R_a})$, can contribute positively to the sum for each action. The sum is thus bounded by:

$$\begin{aligned} & \sum_{a \in \mathcal{A}: R_a \neq \emptyset} \left(\sum_{t \in R_a \setminus \{\max R_a\}} c_j(a, y_t) + c_j(a, y_{\max R_a}) \right) \\ & \leq \sum_{a \in \mathcal{A}: R_a \neq \emptyset} c_j(a, y_{\max R_a}) \\ & \leq \sum_{a \in \mathcal{A}: R_a \neq \emptyset} 1 \\ & \leq |\mathcal{A}|. \end{aligned}$$

■

Proof of Proposition 10. Fix any $a \in \mathcal{A}_{1:T}^{\mathbb{E}[c]}$ and $j \in [J]$. By the definition of the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$, we have that for every $t \in [T]$:

$$\begin{aligned} \Pr[v_{1:t}^{a,j} > \tau] &= \Pr \left[\sum_{s=1}^t \mathbb{1}[a_s = a] c_j(a, y_s) > \tau \right] \\ &\leq \Pr \left[\sum_{s=1}^t \mathbb{1}[a_s = a] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)]) > \tau \right]. \end{aligned}$$

We will adopt a martingale concentration inequality to bound the above probability. For this purpose, we define the filtration $\mathcal{F}_t = \sigma(\{(x_s, p_s, y_s)\}_{s=1}^t, x_{t+1}, p_{t+1})$. Intuitively, \mathcal{F}_t represents the complete history of information available to the agent before making their decision in round $t + 1$, which is why x_{t+1} and p_{t+1} are included. We first show that $\{\mathbb{1}[a_s = a](c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)])\}_{s=1}^T$ is a sequence of martingale differences with respect to this filtration.

We have that $a_t = \text{CBR}_{\widehat{\mathcal{A}}_t^{\mathbb{E}[c]}}^u(p_t)$. $\widehat{\mathcal{A}}_t^{\mathbb{E}[c]}$ is the candidate action set maintained by the agent based on the record of the actions up to round $t - 1$, so $\widehat{\mathcal{A}}_t^{\mathbb{E}[c]} \in \mathcal{F}_{t-1}$. $p_t \in \mathcal{F}_{t-1}$ by the definition of the filtration. Therefore, $a_t \in \mathcal{F}_{t-1}$.

Consequently, for any $t \in [T]$:

$$\begin{aligned} &\mathbb{E} \left[\mathbb{1}[a_t = a] (c_j(a, y_t) - \mathbb{E}_{y_t \sim Y_t} [c_j(a, y_t)]) \mid \mathcal{F}_{t-1} \right] \\ &= \mathbb{1}[a_t = a] \mathbb{E} \left[c_j(a, y_t) - \mathbb{E}_{y_t \sim Y_t} [c_j(a, y_t)] \mid \mathcal{F}_{t-1} \right] \\ &= 0. \end{aligned}$$

This implies that $\{\mathbb{1}[a_s = a](c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)])\}_{s=1}^t$ is a sequence of martingale differences. By Azuma-Hoeffding inequality (Lemma 47), we derive that for any $a \in \mathcal{A}_{1:T}^{\mathbb{E}[c]}$, $j \in [J]$, $t \in [T]$:

$$\begin{aligned} \Pr[v_{1:t}^{a,j} > \tau] &\leq \Pr \left[\sum_{s=1}^t \mathbb{1}[a_s = a] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)]) > \tau \right] \\ &\leq 2e^{-\frac{\tau^2}{8t}} \\ &\leq 2e^{-\frac{\tau^2}{8T}}. \end{aligned}$$

Using the union bound over every $a \in \mathcal{A}_{1:T}^{\mathbb{E}[c]}$, $j \in [J]$, $t \in [T]$, we derive that:

$$\begin{aligned} \Pr \left[\exists a \in \mathcal{A}_{1:T}^{\mathbb{E}[c]}, j \in [J], t \in [T], \text{ such that } v_{1:t}^{a,j} > \tau \right] &\leq \sum_{a \in \mathcal{A}_{1:T}^{\mathbb{E}[c]}} \sum_{j \in [J]} \sum_{t \in [T]} 2e^{-\frac{\tau^2}{8T}} \\ &\leq 2|\mathcal{A}|JT e^{-\frac{\tau^2}{8T}}. \end{aligned}$$

By the elimination rule of Algorithm 2, this implies that the agent never eliminates any action in $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$ across the T rounds with probability at least $1 - 2|\mathcal{A}|JT e^{-\frac{\tau^2}{8T}}$. \blacksquare

Proof of Theorem 12. By Corollary 11, no agent in \mathcal{N} eliminates any benchmark action in $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$ across the T rounds with probability at least $1 - \delta$. We will condition on this event, which ensures that the Algorithm 2 is well-behaved and avoids the risk of eliminating all actions in \mathcal{A} .

Fix any $j \in [J]$. We will bound the cumulative violation against the j -th constraint, i.e., $\sum_{t=1}^T c_j(a_t, y_t)$. The final result follows by taking the maximum over all $j \in [J]$.

Similarly to the proof of Theorem 8, we let R_a denote the set of rounds when a is played, i.e.,

$$R_a = \{t : a_t = a\}$$

Then we have that:

$$\begin{aligned} \sum_{t=1}^T c_j(a_t, y_t) &= \sum_{a \in \mathcal{A}: R_a \neq \emptyset} \sum_{t \in R_a} c_j(a, y_t) \\ &= \sum_{a \in \mathcal{A}: R_a \neq \emptyset} \sum_{t=1}^{\max R_a} \mathbb{1}[t \in R_a] c_j(a, y_t) \\ &= \sum_{a \in \mathcal{A}: R_a \neq \emptyset} \sum_{t=1}^{\max R_a} \mathbb{1}[a_t = a] c_j(a, y_t) \\ &= \sum_{a \in \mathcal{A}: R_a \neq \emptyset} \left(\sum_{t=1}^{\max R_a - 1} \mathbb{1}[a_t = a] c_j(a, y_t) + c_j(a, y_{\max R_a}) \right) \\ &= \sum_{a \in \mathcal{A}: R_a \neq \emptyset} (v_{\max R_a - 1}^{a,j} + c_j(a, y_{\max R_a})) \\ &\leq \sum_{a \in \mathcal{A}: R_a \neq \emptyset} (\tau + 1) \\ &\leq 4|\mathcal{A}| \sqrt{T \ln \frac{|\mathcal{A}| |\mathcal{N}| J T}{\delta}} + |\mathcal{A}|. \end{aligned}$$

■

Proof of Lemma 15. Using the linearity of u , we can write:

$$\begin{aligned} \left| \sum_{t=1}^T (u(a_t, p_t) - u(a_t, y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_t = a] (u(a, p_t) - u(a, y_t)) \right| \\ &= \left| \sum_{a \in \mathcal{A}} \left(u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} \left| u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[a_t = a] (p_t - y_t) \right\|_{\infty} \\ &\leq L \sum_{a \in \mathcal{A}} \alpha_1(T^{u, \mathbf{c}}(a)). \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , and the third inequality follows from (\mathcal{N}, α_1) -decision calibration. By concavity of α_1 and the fact that $\sum_{a \in \mathcal{A}} T^{u, \mathbf{c}}(a) = \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_t = a] = T$, this expression is at most:

$$L|\mathcal{A}|\alpha_1(T/|\mathcal{A}|).$$

■

Proof of Lemma 16. The proof is similar to the proof of Lemma 15. Using the linearity of u , we can write:

$$\begin{aligned} \left| \sum_{t=1}^T (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_t = a] (u(\phi(a), p_t) - u(\phi(a), y_t)) \right| \\ &= \left| \sum_{a \in \mathcal{A}} \left(u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} \left| u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[a_t = a] (p_t - y_t) \right\|_{\infty} \\ &\leq L \sum_{a \in \mathcal{A}} \alpha_1(T^{u, \mathbf{c}}(a)) \\ &\leq L|\mathcal{A}|\alpha_1(T/|\mathcal{A}|). \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , the third inequality follows from (\mathcal{N}, α_1) -decision calibration, and the fourth inequality follows from concavity of α_1 . ■

Appendix C. Omitted Results from Section 4

In this section, we show how to produce predictions that lead to $\tilde{O}(|\mathcal{A}|\sqrt{T})$ regret against the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$. By Corollary 11, with probability at least $1 - \delta$ over the randomness of the outcomes y_1, \dots, y_T , $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]} \subseteq \hat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}$ for every agent and every $t \in [T]$. We will condition on this event throughout this section. Therefore, all the guarantees we give in this section hold with probability at least $1 - \delta$.

Similarly to Definition 13, we will need the following notion of conditional unbiasedness.

Definition 19 ((\mathcal{N}, β_1) -Decision Calibration) *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent runs Algorithm 2 with threshold parameter $4\sqrt{T \ln \frac{|\mathcal{A}||\mathcal{N}|JT}{\delta}}$ to maintain a candidate action set $\hat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}$ and $\hat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}$ -constrained best respond to the prediction p_t in every round $t \in [T]$. Let $\beta_1 : \mathbb{R} \rightarrow \mathbb{R}$. We say that a sequence of predictions p_1, \dots, p_T is (\mathcal{N}, β_1) -decision*

calibrated with respect to a sequence of outcomes y_1, \dots, y_T if for every $a \in \mathcal{A}$ and $(u, \mathbf{c}) \in \mathcal{N}$:

$$\left\| \sum_{t=1}^T \mathbb{1}[\text{CBR}_{\widehat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}}^u(p_t) = a](p_t - y_t) \right\|_{\infty} \leq \beta_1(T^{u, \mathbb{E}[\mathbf{c}]}(a))$$

where $T^{u, \mathbb{E}[\mathbf{c}]}(a) = \sum_{t=1}^T \mathbb{1}[\text{CBR}_{\widehat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}}^u(p_t) = a]$.

Both Definitions 13 and 19 require the predictions to be unbiased conditional on the actions chosen by agents. They differ only in agents' decision rules. In Definition 13, agents $\widehat{\mathcal{A}}_t^{\mathbf{c}}$ -constrained best respond to the predictions to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbf{c}}$, whereas in Definition 19, agents $\widehat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}$ -constrained best respond to the predictions to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$. For simplicity, we refer to both conditions as decision calibration.

Similarly to Theorem 14, decision calibration implies no $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$ -constrained swap regret.

Theorem 20 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent runs Algorithm 2 with threshold parameter $4\sqrt{T \ln \frac{|\mathcal{A}||\mathcal{N}|JT}{\delta}}$ to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$. If the sequence of predictions p_1, \dots, p_T is (\mathcal{N}, β_1) -decision calibrated, then with probability at least $1 - \delta$, the $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$ -constrained swap regret of every agent is bounded by:*

$$\text{Reg}_{\text{swap}}(u, \mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}, 1 : T) \leq 2L|\mathcal{A}|\beta_1(T/|\mathcal{A}|)$$

The proof is similar to the proof of Theorem 14. For any action modification rule $\phi : \mathcal{A} \rightarrow \mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$, we can write:

$$\begin{aligned} & \sum_{t=1}^T (u(\phi(a_t), y_t) - u(a_t, y_t)) \\ &= \sum_{t=1}^T (u(\phi(a_t), y_t) - u(\phi(a_t), p_t)) + \sum_{t=1}^T (u(\phi(a_t), p_t) - u(a_t, p_t)) + \sum_{t=1}^T (u(a_t, p_t) - u(a_t, y_t)) \end{aligned}$$

Conditioning on the event in Corollary 11 and using the fact that $\phi(a_t) \in \mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$, we have that $\phi(a_t) \in \widehat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}$ for every $t \in [T]$. Since agents $\widehat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}$ -constrained best response to the prediction p_t in round t , $a_t = \arg \max_{b \in \widehat{\mathcal{A}}_t^{\mathbb{E}[\mathbf{c}]}} u(b, p_t)$. This implies that:

$$u(\phi(a_t), p_t) \leq u(a_t, p_t).$$

So, the second part in the decomposition is bounded by 0.

Thus, it suffices to bound the other two parts using decision calibration.

Lemma 21 *If the sequence of predictions p_1, \dots, p_T is (\mathcal{N}, β_1) -decision calibrated, then for any $(u, \mathbf{c}) \in \mathcal{N}$:*

$$\left| \sum_{t=1}^T (u(a_t, p_t) - u(a_t, y_t)) \right| \leq L|\mathcal{A}|\beta_1(T/|\mathcal{A}|)$$

Proof The proof is similar to the proof of Lemma 15. Using the linearity of u , we can write:

$$\begin{aligned}
 \left| \sum_{t=1}^T (u(a_t, p_t) - u(a_t, y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_t = a] (u(a, p_t) - u(a, y_t)) \right| \\
 &= \left| \sum_{a \in \mathcal{A}} \left(u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} \left| u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[a_t = a] (p_t - y_t) \right\|_{\infty} \\
 &\leq L \sum_{a \in \mathcal{A}} \beta_1(T^{u, \mathbb{E}[\mathbf{c}]}(a))
 \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , and the third inequality follows from (\mathcal{N}, β_1) -decision calibration. By concavity of β_1 and the fact that $\sum_{a \in \mathcal{A}} T^{u, \mathbb{E}[\mathbf{c}]}(a) = \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_t = a] = T$, this expression is at most:

$$L|\mathcal{A}|\beta_1(T/|\mathcal{A}|)$$

■

Lemma 22 *If the sequence of predictions p_1, \dots, p_T is (\mathcal{N}, β_1) -decision calibrated, then for any $(u, \mathbf{c}) \in \mathcal{N}$ and $\phi : \mathcal{A} \rightarrow \mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$:*

$$\left| \sum_{t=1}^T (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| \leq L|\mathcal{A}|\beta_1(T/|\mathcal{A}|).$$

Proof Using the linearity of u , we can write:

$$\begin{aligned}
 \left| \sum_{t=1}^T (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[a_t = a] (u(\phi(a), p_t) - u(\phi(a), y_t)) \right| \\
 &= \left| \sum_{a \in \mathcal{A}} \left(u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} \left| u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[a_t = a] y_t \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[a_t = a] (p_t - y_t) \right\|_{\infty} \\
 &\leq L \sum_{a \in \mathcal{A}} \beta_1(T^{u, \mathbb{E}[\mathbf{c}]}(a)) \\
 &\leq L|\mathcal{A}|\beta_1(T/|\mathcal{A}|)
 \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , the third inequality follows from (\mathcal{N}, β_1) -decision calibration, and the fourth inequality follows from concavity of β_1 . \blacksquare

We can now complete the proof of Theorem 20.

Proof Fix any $(u, \mathbf{c}) \in \mathcal{N}$ and $\phi : \mathcal{A} \rightarrow \mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$. Applying Lemmas 21 and 22 to the decomposition of the regret against ϕ , we have that:

$$\begin{aligned} & \sum_{t=1}^T (u(\phi(a_t), y_t) - u(a_t, y_t)) \\ &= \sum_{t=1}^T (u(\phi(a_t), y_t) - u(\phi(a_t), p_t)) + \sum_{t=1}^T (u(\phi(a_t), p_t) - u(a_t, p_t)) + \sum_{t=1}^T (u(a_t, p_t) - u(a_t, y_t)) \\ &\leq 2L|\mathcal{A}|\beta_1(T/|\mathcal{A}|) \end{aligned}$$

\blacksquare

We will instantiate UNBIASED-PREDICTION to make decision calibrated predictions as defined in Definition 19. We will refer to this instantiation as DECISION-CALIBRATION-EXPECTATION. It has the following guarantees.

Theorem 23 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent will run Algorithm 2 with threshold parameter $4\sqrt{T \ln \frac{|\mathcal{A}||\mathcal{N}|JT}{\delta}}$ to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$. There is an instantiation of UNBIASED-PREDICTION (Noarov et al., 2023) —which we call DECISION-CALIBRATION-EXPECTATION— producing predictions $p_1, \dots, p_T \in \Delta\mathcal{Y}$ satisfying, with probability at least $1 - \delta$, for any $(u, \mathbf{c}) \in \mathcal{N}$ and $a \in \mathcal{A}$:*

$$\left\| \sum_{t=1}^T \mathbb{1}[\text{CBR}_{\mathcal{A}_t^{\mathbb{E}[\mathbf{c}]}}^u(p_t) = a](p_t - y_t) \right\|_{\infty} \leq O \left(\ln(d|\mathcal{A}||\mathcal{N}|T) + \sqrt{\ln(d|\mathcal{A}||\mathcal{N}|T) \cdot T^{u, \mathbb{E}[\mathbf{c}]}(a)} + \sqrt{\ln \frac{d|\mathcal{A}||\mathcal{N}|}{\delta} \cdot T} \right)$$

Substituting the above bounds into Theorem 20 and combining with Theorem 12, we arrive at the following corollary bounding the cumulative constraint violation and the $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$ -constrained swap regret. To obtain an overall failure probability of at most δ , we instantiate the regret guarantee and the cumulative-constraint-violation guarantee each with failure probability $\delta/2$, and then apply a union bound.

Corollary 24 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent runs Algorithm 2 with threshold parameter $4\sqrt{T \ln \frac{2|\mathcal{A}||\mathcal{N}|JT}{\delta}}$ to compete with the benchmark class $\mathcal{A}_{1:T}^{\mathbb{E}[\mathbf{c}]}$. The sequence of predictions p_1, \dots, p_T outputted by DECISION-CALIBRATION-EXPECTATION ensures that with probability at*

least $1 - \delta$, the cumulative constraint violation and the $\mathcal{A}_{1:T}^{\mathbb{E}[c]}$ -constrained swap regret of every agent are bounded by:

$$\begin{aligned} \text{CCV}(1 : T) &\leq 4|\mathcal{A}|\sqrt{T \ln(2|\mathcal{A}||\mathcal{N}|JT/\delta)} + |\mathcal{A}|, \\ \text{Reg}_{\text{swap}}(u, \mathcal{A}_{1:T}^{\mathbb{E}[c]}, 1 : T) &\leq O\left(L|\mathcal{A}|\sqrt{T \ln \frac{d|\mathcal{A}||\mathcal{N}|T}{\delta}}\right). \end{aligned}$$

Since constrained swap regret is stronger than constrained external regret, the same regret bound also holds for $\text{Reg}_{\text{ext}}(u, \mathcal{A}_{1:T}^{\mathbb{E}[c]}, 1 : T)$.

Appendix D. Simultaneous Guarantees for Multiple Subsequences: Details

Our analysis so far has focused on achieving low regret against the best single benchmark action in hindsight, evaluated over all T rounds. We now extend this framework to a more demanding and general setting by considering performance over multiple, arbitrary subsequences of rounds. We will compete with different benchmarks over different subsequences, and provide more granular guarantees that hold simultaneously across all subsequences.

Let \mathcal{S} be a collection of subsequences, where each subsequence $S \in \mathcal{S}$ is a subset of $[T]$. We assume that the union of these subsequences covers the entire time horizon, i.e., $\cup_{S \in \mathcal{S}} S = [T]$. This ensures that every round belongs to at least one subsequence, and thus has a benchmark for us to compete with. We denote the active subsequences in round $t \in [T]$ as $\mathcal{S}_t = \{S \in \mathcal{S} : t \in S\}$, which is non-empty by our assumption. These subsequences need not be fixed in advance but can be defined dynamically based on contextual information. A subsequence $S \in \mathcal{S}$ is generally characterized by an indicator function $h_S : [T] \times \mathcal{X} \rightarrow \{0, 1\}$. For any round $t \in [T]$ with an observed feature x_t , the round is part of the subsequence S if and only if $h_S(t, x_t) = 1$. This flexible definition allows subsequences to be based on the round index t , the feature x_t , or both. We could also generalize the definition of subsequence to let it depend arbitrarily on the history.

We now extend our benchmark definitions to the multi-subsequence setting. Fix a subsequence $S \in \mathcal{S}$, one benchmark class of interest is the set of actions that satisfy the realized constraints over the entire subsequence, denoted as:

$$\mathcal{A}_S^c = \{a : c_j(a, y_t) \leq 0 \text{ for every } t \in S \text{ and } j \in [J]\}.$$

Another benchmark class is the set of actions that satisfy the constraints in expectation over the entire subsequence, denoted as:

$$\mathcal{A}_S^{\mathbb{E}[c]} = \left\{ a : \mathbb{E}_{y_t \sim Y_t} [c_j(a, y_t)] \leq 0 \text{ for every } t \in S \text{ and } j \in [J] \right\}.$$

Previously, our notions of constrained external regret and constrained swap regret were defined over the entire time horizon of T rounds. We now extend these definitions to handle arbitrary subsequences. Note that the “best action in hindsight” can be different on each subsequence, so regret bounds that hold simultaneously on many subsequences are comparing to a substantially richer benchmark class than regret bounds that hold marginally over a single subsequence.

Definition 25 (\mathcal{B} -Constrained External Regret over Subsequence S) Fix an agent with action space \mathcal{A} and utility function u . Let \mathcal{B} be a benchmark class of actions. Fix a subsequence $S \subseteq [T]$. For a

sequence of actions a_1, \dots, a_T and outcomes y_1, \dots, y_T , the agent's \mathcal{B} -constrained external regret against the benchmark class \mathcal{B} over the subsequence S is:

$$\text{Reg}_{\text{ext}}(u, \mathcal{B}, S) = \max_{a \in \mathcal{B}} \sum_{t \in S} (u(a, y_t) - u(a_t, y_t)).$$

Definition 26 (\mathcal{B} -Constrained Swap Regret over Subsequence S) Fix an agent with action space \mathcal{A} and utility function u . Let \mathcal{B} be a benchmark class of actions. Fix a subsequence $S \subseteq [T]$. For a sequence of actions a_1, \dots, a_T and outcomes y_1, \dots, y_T , the agent's \mathcal{B} -constrained swap regret against the benchmark class \mathcal{B} over the subsequence S is:

$$\text{Reg}_{\text{swap}}(u, \mathcal{B}, S) = \max_{\phi: \mathcal{A} \rightarrow \mathcal{B}} \sum_{t \in S} (u(\phi(a_t), y_t) - u(a_t, y_t)).$$

Similarly, we define the cumulative constraint violation over a subsequence:

$$\text{CCV}(S) := \max_{j \in [J]} \sum_{t \in S} c_j(a_t, y_t).$$

Agents aim to achieve low regret and satisfy the long-term constraints simultaneously over every subsequence in \mathcal{S} . The latter requires that $\text{CCV}(S) \leq o(|S|)$ for any $S \in \mathcal{S}$, where $|S|$ is the length of subsequence S .

To compete with the benchmark classes $\{\mathcal{A}_S^c\}_{S \in \mathcal{S}}$, we use Algorithm 1 as the fundamental building block for our multi-subsequence setting. We will let each agent instantiate a copy of Algorithm 1 for every subsequence. In each round $t \in [T]$, these parallel instantiations provide the agent with a candidate action set $\hat{\mathcal{A}}_{t,S}^c$ for every active subsequence $S \in \mathcal{S}_t$. The agent then aggregates these candidate action sets by taking their union, $U_t = \cup_{S \in \mathcal{S}_t} \hat{\mathcal{A}}_{t,S}^c$, and U_t -constrained best respond to our prediction p_t . This procedure is formally stated in Algorithm 3.

Algorithm 3 Elimination-Based Algorithm for Benchmark Classes $\{\mathcal{A}_S^c\}_{S \in \mathcal{S}}$

Initialize $\hat{\mathcal{A}}_{1,S}^c = \mathcal{A}$ for every subsequence $S \in \mathcal{S}$

for $t = 1$ **to** T **do**

 Receive the prediction p_t

 Aggregate candidate action sets from active subsequences, $U_t = \cup_{S \in \mathcal{S}_t} \hat{\mathcal{A}}_{t,S}^c$

 Play the U_t -constrained best response to p_t , $a_t = \text{CBR}_{U_t}^u(p_t)$

 Observe the outcome y_t , obtain utility $u(a_t, y_t)$ and constraint increments $c_j(a_t, y_t)$, $j \in [J]$

 Obtain $\hat{\mathcal{A}}_{t+1,S}^c$ by executing one step of Algorithm 1 for every active subsequence $S \in \mathcal{S}_t$

 Let $\hat{\mathcal{A}}_{t+1,S}^c = \hat{\mathcal{A}}_{t,S}^c$ for every non-active subsequence $S \in \mathcal{S} \setminus \mathcal{S}_t$

end

Since each instantiation of Algorithm 1 is guaranteed not to eliminate the benchmark actions corresponding to its subsequence (Proposition 7), the union set U_t at each round $t \in [T]$ will therefore contain the benchmark actions for every active subsequence $S \in \mathcal{S}_t$.

Proposition 27 Let \mathcal{S} be a collection of subsequences. Fix an agent in \mathcal{N} with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Suppose the agent runs Algorithm 3 to compete with the benchmark class \mathcal{A}_S^c over every subsequence $S \in \mathcal{S}$, then $\mathcal{A}_S^c \subseteq \hat{\mathcal{A}}_{t,S}^c \subseteq U_t$ for any $t \in [T]$ and $S \in \mathcal{S}_t$.

Fix any subsequence $S \in \mathcal{S}$. Since the agent chooses actions from the union sets, an action's selection can be sourced from the candidate action set of any subsequence in \mathcal{S} . As a result, each action can contribute at most $|\mathcal{S}|$ times to the cumulative constraint violation over S before it is fully eliminated. Consequently, the cumulative constraint violation of Algorithm 3 over any subsequence can be upper bounded by $|\mathcal{A}||\mathcal{S}|$.

Theorem 28 *Let \mathcal{S} be a collection of subsequences. Fix an agent in \mathcal{N} with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Suppose the agent runs Algorithm 3 to compete with the benchmark class \mathcal{A}_S^c over every subsequence $S \in \mathcal{S}$, then the cumulative constraint violation of the agent over any subsequence $S \in \mathcal{S}$ is bounded by:*

$$\text{CCV}(S) \leq |\mathcal{A}||\mathcal{S}|.$$

Proof Fix any $j \in [J]$ and $S \in \mathcal{S}$. We will bound the cumulative violation against the j -th constraint over S , i.e., $\sum_{t \in S} c_j(a_t, y_t)$. The final result follows by taking the maximum over all $j \in [J]$.

Similarly to the proof of Theorem 8, we first partition the sum based on the action played. For each action $a \in \mathcal{A}$, let $R_{S,a} = \{t : t \in S, a_t = a\}$. This gives:

$$\sum_{t \in S} c_j(a_t, y_t) = \sum_{a \in \mathcal{A}: R_{S,a} \neq \emptyset} \sum_{t \in R_{S,a}} c_j(a, y_t).$$

For an action a to be played in round t , it must be a candidate action for at least one active subsequence S' . As a result, the above equation can be upper bounded by:

$$\sum_{a \in \mathcal{A}: R_{S,a} \neq \emptyset} \sum_{t \in R_{S,a}} \sum_{S' \in \mathcal{S}} \mathbb{1}[t \in S', a \in \hat{\mathcal{A}}_{t,S'}^c] c_j(a, y_t) = \sum_{a \in \mathcal{A}: R_{S,a} \neq \emptyset} \sum_{S' \in \mathcal{S}} \sum_{t \in R_{S,a} \cap S': a \in \hat{\mathcal{A}}_{t,S'}^c} c_j(a, y_t).$$

For notational convenience, we define $R_{S,a,S'} = \{t \in R_{S,a} \cap S' : a \in \hat{\mathcal{A}}_{t,S'}^c\}$. For any $t \in R_{S,a,S'} \setminus \{\max R_{S,a,S'}\}$, action a is played and is not eliminated from the candidate action set corresponding to S' , which implies that the constraint is satisfied, i.e., $c_j(a, y_t) \leq 0$. Therefore, only the final term, $c_j(a, y_{\max R_{S,a,S'}})$, can contribute positively to the sum for each action. The sum is thus bounded by:

$$\begin{aligned} & \sum_{a \in \mathcal{A}: R_{S,a} \neq \emptyset} \sum_{S' \in \mathcal{S}} \left(\sum_{t \in R_{S,a,S'} \setminus \{\max R_{S,a,S'}\}} c_j(a, y_t) + c_j(a, y_{\max R_{S,a,S'}}) \right) \\ & \leq \sum_{a \in \mathcal{A}: R_{S,a} \neq \emptyset} \sum_{S' \in \mathcal{S}} c_j(a, y_{\max R_{S,a,S'}}) \\ & \leq \sum_{a \in \mathcal{A}: R_{S,a} \neq \emptyset} \sum_{S' \in \mathcal{S}} 1 \\ & \leq |\mathcal{A}||\mathcal{S}|. \end{aligned}$$

■

Handling the expectation-based benchmarks $\{\mathcal{A}_S^{\mathbb{E}[c]}\}_{S \in \mathcal{S}}$, in contrast, requires a more sophisticated approach than simply running parallel copies of Algorithm 2. The challenge arises because

each instantiation of the algorithm has only a local view of an action’s performance. An action could appear feasible from the perspective of a subsequence S , but within a smaller, overlapping portion of another subsequence S' , that same action could be highly problematic. Low (i.e. very negative) constraint violation over the difference $S \setminus S'$ can mask the high constraint violation on the intersection $S \cap S'$, preventing S from eliminating the action. This action would then remain in the union set U_t and continue to harm the performance of subsequence S' , even if S' has already eliminated it.

Our solution has the same high-level structure as Algorithm 3. We maintain a candidate action set $\widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]}$ for each subsequence and gradually eliminate actions from it; the agent selects an action from the union set, $U_t = \cup_{S \in \mathcal{S}_t} \widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]}$, in each round. The key difference lies in a more sophisticated elimination rule.

We will monitor each action’s constraint violations more granularly by tracking their performance over the intersections of pairs of subsequence. However, since these intersections overlap, simply summing constraint violations across them would lead to double-counting. To resolve this, we will attribute the action a_t played in each round $t \in [T]$ to a single, uniquely responsible subsequence. This responsible subsequence can be selected in any arbitrary way. To be concrete, we let the collection of subsequences be indexed as $\mathcal{S} = \{S_1, S_2, \dots, S_I\}$ where $I = |\mathcal{S}|$. We credit the subsequence S_{r_t} with the smallest index that is active and for which a_t is a candidate action:

$$r_t = \min \left\{ i \in [I] : t \in S_i, a_t \in \widehat{\mathcal{A}}_{t,S_i}^{\mathbb{E}[c]} \right\}.$$

For each action $a \in \mathcal{A}$, constraint $j \in [J]$, and pair of subsequences $(S, S') \in \mathcal{S}^2$, we define the attributed constraint violation $w_{[1,t] \cap S}^{a,j,S'}$ that tracks the cumulative violation of action a against the j -th constraint during the rounds of S up to round t for which subsequence S' is responsible:

$$w_{[1,t] \cap S}^{a,j,S'} := \sum_{s=1}^t \mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] c_j(a, y_s).$$

Our new elimination rule is based on monitoring these terms against subsequence-specific thresholds, $\{\tau_{S'}\}_{S' \in \mathcal{S}}$. An action a_t is eliminated from the candidate action set of its responsible subsequence, S_{r_t} , if its attributed constraint violations exceed a subsequence-specific threshold $\tau_{S_{r_t}}$. Specifically, we eliminate a_t from $\widehat{\mathcal{A}}_{t,S_{r_t}}^{\mathbb{E}[c]}$ if for any constraint $j \in [J]$ and any subsequence $S \in \mathcal{S}$, the term $w_{[1,t] \cap S}^{a,j,S_{r_t}}$ surpasses $\tau_{S_{r_t}}$. This entire procedure is formally detailed in Algorithm 4.

Algorithm 4 never eliminates any benchmark action in $\mathcal{A}_S^{\mathbb{E}[c]}$ from the candidate action set $\widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]}$ across the T rounds for any subsequence $S \in \mathcal{S}$ with high probability over the randomness of the outcomes y_1, \dots, y_T .

Proposition 29 *Fix an agent in \mathcal{N} with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Suppose the agent runs Algorithm 4 with threshold parameters $\{\tau_S\}_{S \in \mathcal{S}}$ to compete with the benchmark class $\mathcal{A}_S^{\mathbb{E}[c]}$ over every subsequence $S \in \mathcal{S}$, then with probability at least $1 - 2|\mathcal{A}||\mathcal{S}|J \sum_{S \in \mathcal{S}} |S| e^{-\frac{\tau_S^2}{8|S|}}$:*

$$\mathcal{A}_S^{\mathbb{E}[c]} \subseteq \widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]} \subseteq U_t \text{ for every } t \in [T] \text{ and } S \in \mathcal{S}_t.$$

Algorithm 4 Elimination-Based Algorithm for Benchmark Classes $\{\mathcal{A}_S^{\mathbb{E}[c]}\}_{S \in \mathcal{S}}$

Input: Threshold parameters $\{\tau_S\}_{S \in \mathcal{S}}$

 Initialize $\widehat{\mathcal{A}}_{1,S}^{\mathbb{E}[c]} = \mathcal{A}$ for each $S \in \mathcal{S}$

 For every $a \in \mathcal{A}$, $j \in [J]$, and $S \in \mathcal{S}$, set $w_0^{a,j,S} = 0$
for $t = 1$ **to** T **do**

 Receive the prediction p_t

 Aggregate candidate action sets from active subsequences, $U_t = \cup_{S \in \mathcal{S}_t} \widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]}$

 Play the U_t -constrained best response to p_t , $a_t = \text{CBR}_{U_t}^u(p_t)$

 Determine the responsible subsequence with index $r_t = \min\{i \in [I] : t \in S_i, a_t \in \widehat{\mathcal{A}}_{t,S_i}^{\mathbb{E}[c]}\}$

 Observe the outcome y_t , obtain utility $u(a_t, y_t)$ and constraint increments $c_j(a_t, y_t)$, $j \in [J]$

 Update $w_{[1,t] \cap S}^{a,j,S'} = w_{[1,t-1] \cap S}^{a,j,S'} + \mathbb{1}[a_t = a, t \in S, S_{r_t} = S'] c_j(a, y_t)$ for every $a \in \mathcal{A}$, $j \in [J]$, and $S, S' \in \mathcal{S}$
if $\exists j \in [J], S \in \mathcal{S}, w_{[1,t] \cap S}^{a_t,j,S_{r_t}} > \tau_S$ **then**

 | $\widehat{\mathcal{A}}_{t+1,S_{r_t}}^{\mathbb{E}[c]} = \widehat{\mathcal{A}}_{t,S_{r_t}}^{\mathbb{E}[c]} \setminus \{a_t\}$
end
else

 | $\widehat{\mathcal{A}}_{t+1,S_{r_t}}^{\mathbb{E}[c]} = \widehat{\mathcal{A}}_{t,S_{r_t}}^{\mathbb{E}[c]}$
end

 Let $\widehat{\mathcal{A}}_{t+1,S}^{\mathbb{E}[c]} = \widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]}$ for any non-responsible subsequence $S \in \mathcal{S} \setminus \{S_{r_t}\}$
end

Proof Fix any $j \in [J]$, $(S, S') \in \mathcal{S}^2$, and $a \in \mathcal{A}_S^{\mathbb{E}[c]}$, we have that for every $t \in [T]$:

$$\begin{aligned} \Pr \left[w_{[1,t] \cap S}^{a,j,S'} > \tau_S \right] &= \Pr \left[\sum_{s=1}^t \mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] c_j(a, y_s) > \tau_S \right] \\ &\leq \Pr \left[\sum_{s=1}^t \mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)]) > \tau_S \right]. \end{aligned}$$

Similarly to the proof of Proposition 10, we define the filtration $\mathcal{F}_t = \sigma(\{(x_s, p_s, y_s)\}_{s=1}^t, x_{t+1}, p_{t+1})$ and first show that $\{\mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)])\}_{s=1}^T$ is a sequence of martingale differences.

We have that $a_t = \text{CBR}_{U_t}^u(p_t)$. U_t is the union of some candidate action set maintained by the agent based on the record of the actions up to round $t-1$, so $U_t \in \mathcal{F}_{t-1}$. $p_t \in \mathcal{F}_{t-1}$ by definition of the filtration. So $a_t \in \mathcal{F}_{t-1}$ and hence $\mathbb{1}[a_t = a] \in \mathcal{F}_{t-1}$.

We have that $\mathbb{1}[t \in S] = \mathbb{1}[h_{S_i}(t, x_t) = 1]$, where $x_t \in \mathcal{F}_{t-1}$. As a result, $\mathbb{1}[t \in S] \in \mathcal{F}_{t-1}$.

The responsible subsequence at round t has index $r_t = \min\{i \in [I] : t \in S_i, a_t \in \widehat{\mathcal{A}}_{t,S_i}^{\mathbb{E}[c]}\} = \min\{i \in [I] : h_{S_i}(t, x_t) = 1, a_t \in \widehat{\mathcal{A}}_{t,S_i}^{\mathbb{E}[c]}\}$ where $x_t \in \mathcal{F}_{t-1}$, $a_t \in \mathcal{F}_{t-1}$, and $\widehat{\mathcal{A}}_{t,S_i}^{\mathbb{E}[c]} \in \mathcal{F}_{t-1}$. Therefore, $\mathbb{1}[S_{r_t} = S'] \in \mathcal{F}_{t-1}$.

Consequently, for any $t \in [T]$:

$$\begin{aligned} & \mathbb{E} \left[\mathbb{1}[a_t = a, t \in S, S_{r_t} = S'] (c_j(a, y_t) - \mathbb{E}_{y_t \sim Y_t} [c_j(a, y_t)]) \mid \mathcal{F}_{t-1} \right] \\ &= \mathbb{1}[a_t = a, t \in S, S_{r_t} = S'] \mathbb{E} \left[c_j(a, y_t) - \mathbb{E}_{y_t \sim Y_t} [c_j(a, y_t)] \mid \mathcal{F}_{t-1} \right] \\ &= 0. \end{aligned}$$

This implies that $\{\mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)])\}_{s=1}^t$ is a sequence of martingale differences. The subsequence of these terms corresponding to rounds $s \in S$, i.e., $\{\mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)])\}_{s \in S: s \leq t}$, is also a martingale difference sequence, because the selection rule is predictable with respect to the filtration \mathcal{F}_{s-1} . By Azuma-Hoeffding inequality (Lemma 47), we derive that for any $j \in [J]$, $(S, S') \in \mathcal{S}^2$, $a \in \mathcal{A}_S^{\mathbb{E}[c]}$, and $t \in S$:

$$\begin{aligned} \Pr \left[w_{[1,t] \cap S}^{a,j,S'} > \tau_S \right] &\leq \Pr \left[\sum_{s=1}^t \mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)]) > \tau_S \right] \\ &= \Pr \left[\sum_{s \in S: s \leq t} \mathbb{1}[a_s = a, s \in S, S_{r_s} = S'] (c_j(a, y_s) - \mathbb{E}_{y_s \sim Y_s} [c_j(a, y_s)]) > \tau_S \right] \\ &\leq 2e^{-\frac{\tau_S^2}{8|S|}}. \end{aligned}$$

Using the union bound over every $j \in [J]$, $(S, S') \in \mathcal{S}^2$, $a \in \mathcal{A}_{1:T}^{\mathbb{E}[c]}$, and $t \in S$, we derive that:

$$\begin{aligned} \Pr \left[\exists j \in [J], (S, S') \in \mathcal{S}^2, a \in \mathcal{A}_S^{\mathbb{E}[c]}, t \in S, \text{ such that } w_{[1,t] \cap S}^{a,j,S'} > \tau_S \right] &\leq \sum_{j \in [J]} \sum_{(S, S') \in \mathcal{S}^2} \sum_{a \in \mathcal{A}_S^{\mathbb{E}[c]}} \sum_{t \in S} 2e^{-\frac{\tau_S^2}{8|S|}} \\ &\leq 2|\mathcal{A}||\mathcal{S}|J \sum_{S \in \mathcal{S}} |S| e^{-\frac{\tau_S^2}{8|S|}}. \end{aligned}$$

By the elimination rule of Algorithm 4, this implies that the agent never eliminates any action in $\mathcal{A}_S^{\mathbb{E}[c]}$ across the T rounds for any subsequence $S \in \mathcal{S}$ with probability at least $1 - 2|\mathcal{A}||\mathcal{S}|J \sum_{S \in \mathcal{S}} |S| e^{-\frac{\tau_S^2}{8|S|}}$. \blacksquare

Proposition 29 provides a guarantee for a single agent. To extend this to all agents in \mathcal{N} , we apply a union bound. By setting the thresholds $\tau_S = 4\sqrt{|S| \ln \frac{|\mathcal{A}||\mathcal{N}||\mathcal{S}|^2 J |S|}{\delta}}$ for a desired failure probability $\delta \in (0, 1)$, we arrive at the following corollary, which holds with probability at least $1 - \delta$.

Corollary 30 *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent runs Algorithm 4 to compete with the benchmark class $\mathcal{A}_S^{\mathbb{E}[c]}$ over every subsequence $S \in \mathcal{S}$, where the threshold parameter for each subsequence S is set to $4\sqrt{|S| \ln \frac{|\mathcal{A}||\mathcal{N}||\mathcal{S}|^2 J |S|}{\delta}}$. Then with probability at least $1 - \delta$, for every agent:*

$$\mathcal{A}_S^{\mathbb{E}[c]} \subseteq \widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[c]} \subseteq U_t \text{ for every } t \in [T] \text{ and } S \in \mathcal{S}_t.$$

We now bound the cumulative constraint violation of Algorithm 4. The reasoning is a granular extension of the argument for Theorem 12. In Algorithm 4, an action a is eliminated from the candidate action set of the responsible subsequence S' only when one of its attributed constraint violations, $w_{[1,t] \cap S'}^{a,j,S'}$, surpasses the threshold $\tau_{S'}$. This implies that each attributed constraint violation term, $w_{[1,t] \cap S'}^{a,j,S'}$, is bounded by at most $\tau_{S'} + 1$. Summing this bound over all possible responsible subsequences $S' \in \mathcal{S}$ and all actions $a \in \mathcal{A}$ gives the final bound on $\text{CCV}(S)$.

Theorem 31 *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent runs Algorithm 4 to compete with the benchmark class $\mathcal{A}_S^{\mathbb{E}[c]}$ over every subsequence $S \in \mathcal{S}$, where the threshold parameter for each subsequence S is set to $4\sqrt{|S| \ln \frac{|\mathcal{A}||\mathcal{N}||S|^2 J |S|}{\delta}}$. Then with probability at least $1 - \delta$, the cumulative constraint violation of any agent over any subsequence $S \in \mathcal{S}$ is bounded by:*

$$\text{CCV}(S) \leq \sum_{S' \in \mathcal{S}} 4|\mathcal{A}| \sqrt{|S'| \ln \frac{|\mathcal{A}||\mathcal{N}||S|^2 J |S'|}{\delta}} + |\mathcal{A}||S|.$$

Proof By Corollary 30, no agent in \mathcal{N} eliminates any benchmark action in $\mathcal{A}_S^{\mathbb{E}[c]}$ across the T rounds for any subsequence $S \in \mathcal{S}$ with probability at least $1 - \delta$. We will condition on this event, which ensures that the Algorithm 4 is well-behaved and avoids the risk of eliminating all actions in \mathcal{A} .

Fix any $j \in [J]$ and $S \in \mathcal{S}$. We will bound the cumulative violation against the j -th constraint over S , i.e., $\sum_{t \in S} c_j(a_t, y_t)$. The final result follows by taking the maximum over all $j \in [J]$.

We will partition the sum based on the action played and the responsible subsequence. For each action $a \in \mathcal{A}$ and subsequence $S' \in \mathcal{S}$, we define $R_{S,a,S'} = \{t \in [T] : a_t = a, t \in S, S_{r_t} = S'\}$. Then we have that:

$$\begin{aligned} \sum_{t \in S} c_j(a_t, y_t) &= \sum_{t=1}^T \mathbb{1}[t \in S] c_j(a_t, y_t) \\ &= \sum_{a \in \mathcal{A}} \sum_{S' \in \mathcal{S}} \sum_{t=1}^T \mathbb{1}[a_t = a, t \in S, S_{r_t} = S'] c_j(a, y_t) \\ &= \sum_{a \in \mathcal{A}} \sum_{S' \in \mathcal{S}} \sum_{t=1}^{\max R_{S,a,S'}} \mathbb{1}[a_t = a, t \in S, S_{r_t} = S'] c_j(a, y_t) \\ &= \sum_{a \in \mathcal{A}} \sum_{S' \in \mathcal{S}} \left(\sum_{t=1}^{\max R_{S,a,S'} - 1} \mathbb{1}[a_t = a, t \in S, S_{r_t} = S'] c_j(a, y_t) + c_j(a, y_{\max R_{S,a,S'}}) \right) \\ &= \sum_{a \in \mathcal{A}} \sum_{S' \in \mathcal{S}} (w_{[1, \max R_{S,a,S'} - 1] \cap S'}^{a,j,S'} + c_j(a, y_{\max R_{S,a,S'}})) \\ &\leq \sum_{a \in \mathcal{A}} \sum_{S' \in \mathcal{S}} (\tau_{S'} + 1) \\ &= \sum_{S' \in \mathcal{S}} 4|\mathcal{A}| \sqrt{|S'| \ln \frac{|\mathcal{A}||\mathcal{N}||S|^2 J |S'|}{\delta}} + |\mathcal{A}||S|. \end{aligned}$$

■

Next, we focus on the regret against the benchmark class \mathcal{A}_S^c over every subsequence $S \in \mathcal{S}$. Similarly to Section 4, we will obtain the same regret bounds for the benchmark class $\mathcal{A}_S^{\mathbb{E}[c]}$ and relegate the results to Appendix F.

To obtain the regret bounds, we will need the predictions to be decision calibrated, conditional on every subsequence in \mathcal{S} .

Definition 32 ($(\mathcal{N}, \mathcal{S}, \alpha_2)$ -Decision Calibration) *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent will run Algorithm 3 to compete with the benchmark class \mathcal{A}_S^c over every subsequence $S \in \mathcal{S}$. Each agent will maintain a candidate action set $\{U_t\}_{t \in [T]}$. Let $\alpha_2 : \mathbb{R} \rightarrow \mathbb{R}$. We say that a sequence of predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \alpha_2)$ -decision calibrated with respect to a sequence of outcomes y_1, \dots, y_T if for every $S \in \mathcal{S}$, $a \in \mathcal{A}$, and $(u, \mathbf{c}) \in \mathcal{N}$:*

$$\left\| \sum_{t=1}^T \mathbb{1}[t \in S, \text{CBR}_{U_t}^u(p_t) = a](p_t - y_t) \right\|_{\infty} \leq \alpha_2(T^{u, \mathbf{c}, S}(a)).$$

where $T^{u, \mathbf{c}, S}(a) = \sum_{t=1}^T \mathbb{1}[t \in S, \text{CBR}_{U_t}^u(p_t) = a]$.

Similarly to Theorem 14 in the single-sequence case, decision calibration leads to the constrained swap regret bound.

Theorem 33 *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent runs Algorithm 3 to compete with the benchmark class \mathcal{A}_S^c over every subsequence $S \in \mathcal{S}$. If the sequence of predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \alpha_2)$ -decision calibrated, then the \mathcal{A}_S^c -constrained swap regret of any agent over any subsequence $S \in \mathcal{S}$ is bounded by:*

$$\text{Reg}_{\text{swap}}(u, \mathcal{A}_S^c, S) \leq 2L|\mathcal{A}|\alpha_2(|S|/|\mathcal{A}|).$$

We will again instantiate UNBIASED-PREDICTION to make decision calibrated predictions.

Theorem 34 *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent will run Algorithm 3 to compete with the benchmark class \mathcal{A}_S^c over every subsequence $S \in \mathcal{S}$. Each agent will maintain a candidate action set $\{U_t\}_{t \in [T]}$. There is an instantiation of UNBIASED-PREDICTION (Noarov et al., 2023)—which we call DECISION-CALIBRATION-SUBSEQUENCE—producing predictions $p_1, \dots, p_T \in \mathcal{Y}$ satisfying, for any sequence of outcomes $y_1, \dots, y_T \in \mathcal{Y}$, the following holds with probability at least $1 - \delta$ for any $(u, \mathbf{c}) \in \mathcal{N}$, $a \in \mathcal{A}$, and $S \in \mathcal{S}$:*

$$\begin{aligned} & \left\| \sum_{t=1}^T \mathbb{1}[t \in S, \text{CBR}_{U_t}^u(p_t) = a](p_t - y_t) \right\|_{\infty} \\ & \leq O \left(\ln(d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T) + \sqrt{\ln(d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T) \cdot T^{u, \mathbf{c}, S}(a)} + \sqrt{\ln \frac{d|\mathcal{A}||\mathcal{N}||\mathcal{S}|}{\delta} \cdot |S|} \right). \end{aligned}$$

Substituting the above bounds into Theorem 33 and combining with Theorem 28, we arrive at the following corollary bounding the cumulative constraint violation and the \mathcal{A}_S^c -constrained swap regret over every subsequence $S \in \mathcal{S}$.

Corollary 35 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent will run Algorithm 3 to compete with the benchmark class \mathcal{A}_S^c over every subsequence $S \in \mathcal{S}$. The sequence of predictions p_1, \dots, p_T outputted by DECISION-CALIBRATION-SUBSEQUENCE ensures that with probability at least $1 - \delta$, the cumulative constraint violation and the \mathcal{A}_S^c -constrained swap regret of any agent over any subsequence $S \in \mathcal{S}$ are bounded by:*

$$\begin{aligned} \text{CCV}(S) &\leq |\mathcal{A}||\mathcal{S}|, \\ \text{Reg}_{\text{swap}}(u, \mathcal{A}_S^c, S) &\leq O\left(L|\mathcal{A}|\sqrt{|S| \ln \frac{d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T}{\delta}}\right). \end{aligned}$$

Since constrained swap regret is stronger than constrained external regret, the same regret bound also holds for $\text{Reg}_{\text{ext}}(u, \mathcal{A}_S^c, S)$.

Appendix E. Proof of Theorem 33

Similarly to the proof of Theorem 14, for any subsequence $S \in \mathcal{S}$ and action modification rule $\phi : \mathcal{A} \rightarrow \mathcal{A}_S^c$, we can decompose the regret against ϕ into three parts as:

$$\begin{aligned} &\sum_{t \in S} (u(\phi(a_t), y_t) - u(a_t, y_t)) \\ &= \sum_{t \in S} (u(\phi(a_t), y_t) - u(\phi(a_t), p_t)) + \sum_{t \in S} (u(\phi(a_t), p_t) - u(a_t, p_t)) + \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)) \end{aligned}$$

By Proposition 27 and the fact that $\phi(a_t) \in \mathcal{A}_S^c$, $\phi(a_t) \in U_t$ for every $t \in S$. Since agents U_t -constrained best response to the prediction p_t in round t , $a_t = \arg \max_{b \in U_t} u(b, p_t)$. This implies that:

$$u(\phi(a_t), p_t) \leq u(a_t, p_t).$$

So, the second part in the decomposition is upper bounded by 0.

Thus, it suffices to bound the other two parts using decision calibration.

Lemma 36 *If the sequence of predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \alpha_2)$ -decision calibrated, then for any $(u, c) \in \mathcal{N}$ and $S \in \mathcal{S}$:*

$$\left| \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)) \right| \leq L|\mathcal{A}|\alpha_2(|S|/|\mathcal{A}|).$$

Proof Using the linearity of u , we can write:

$$\begin{aligned}
 \left| \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (u(a, p_t) - u(a, y_t)) \right| \\
 &= \left| \sum_{a \in \mathcal{A}} \left(u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} \left| u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (p_t - y_t) \right\|_{\infty} \\
 &\leq L \sum_{a \in \mathcal{A}} \alpha_2(T^{u, \mathbf{c}, S}(a)).
 \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , and the third inequality follows from $(\mathcal{N}, \mathcal{S}, \alpha_2)$ -decision calibration. By concavity of α_2 and the fact that $\sum_{a \in \mathcal{A}} T^{u, \mathbf{c}, S}(a) = \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] = |S|$, this expression is at most:

$$L|\mathcal{A}|\alpha_2(|S|/|\mathcal{A}|).$$

■

Lemma 37 *If the sequence of predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \alpha_2)$ -decision calibrated, then for any $(u, \mathbf{c}) \in \mathcal{N}$ and $S \in \mathcal{S}$:*

$$\left| \sum_{t \in S} (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| \leq L|\mathcal{A}|\alpha_2(|S|/|\mathcal{A}|).$$

Proof Using the linearity of u , we can write:

$$\begin{aligned}
 \left| \sum_{t \in S} (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (u(\phi(a), p_t) - u(\phi(a), y_t)) \right| \\
 &= \left| \sum_{a \in \mathcal{A}} \left(u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} \left| u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right| \\
 &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (p_t - y_t) \right\|_{\infty} \\
 &\leq L \sum_{a \in \mathcal{A}} \alpha_2(T^{u, \mathbf{c}, S}(a)) \\
 &\leq L|\mathcal{A}|\alpha_2(|S|/|\mathcal{A}|)
 \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , the third inequality follows from $(\mathcal{N}, \mathcal{S}, \alpha_2)$ -decision calibration, and the fourth inequality follows from concavity of α_2 . \blacksquare

We can now complete the proof of Theorem 33.

Proof Fix any $(u, \mathbf{c}) \in \mathcal{N}$, $S \in \mathcal{S}$, and $\phi : \mathcal{A} \rightarrow \mathcal{A}_S^c$. Applying Lemmas 36 and 37 to the decomposition of the regret against ϕ , we have that:

$$\begin{aligned} \sum_{t \in S} (u(a, y_t) - u(a_t, y_t)) &= \sum_{t \in S} (u(a, y_t) - u(a, p_t)) + \sum_{t \in S} (u(a, p_t) - u(a_t, p_t)) + \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)) \\ &\leq 2L|\mathcal{A}|\alpha_2(|S|/|\mathcal{A}|). \end{aligned}$$

\blacksquare

Appendix F. Omitted Results from Appendix D

In this section, we derive regret bounds for the benchmark classes $\{\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}\}_{S \in \mathcal{S}}$. By Corollary 30, with probability at least $1 - \delta$ over the randomness of the outcomes y_1, \dots, y_T , $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]} \subseteq \widehat{\mathcal{A}}_{t,S}^{\mathbb{E}[\mathbf{c}]} \subseteq U_t$ for every agent, every $t \in [T]$, and every $S \in \mathcal{S}_t$. We will condition on this event throughout this section. Therefore, all the guarantees we give in this section hold with probability at least $1 - \delta$.

Similarly to Definition 32, we will need the following notion of conditional unbiasedness.

Definition 38 ($(\mathcal{N}, \mathcal{S}, \beta_2)$ -Decision Calibration) *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent will run Algorithm 4 to compete with the benchmark class $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$ over every subsequence $S \in \mathcal{S}$, where the threshold parameter for each subsequence S is set to $4\sqrt{|S| \ln \frac{|\mathcal{A}||\mathcal{N}||\mathcal{S}|^2 J |S|}{\delta}}$. Each agent will maintain a candidate action set $\{U_t\}_{t \in [T]}$. Let $\beta_2 : \mathbb{R} \rightarrow \mathbb{R}$. We say that a sequence of predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \beta_2)$ -decision calibrated with respect to a sequence of outcomes y_1, \dots, y_T if for every $S \in \mathcal{S}$, $a \in \mathcal{A}$, and $(u, \mathbf{c}) \in \mathcal{N}$:*

$$\left\| \sum_{t=1}^T \mathbb{1}[t \in S, \text{CBR}_{U_t}^u(p_t) = a](p_t - y_t) \right\|_{\infty} \leq \beta_2(T^{u, \mathbb{E}[\mathbf{c}], S}(a)),$$

where $T^{u, \mathbb{E}[\mathbf{c}], S}(a) = \sum_{t=1}^T \mathbb{1}[t \in S, \text{CBR}_{U_t}^u(p_t) = a]$.

Similarly to Theorem 20 in the single-sequence case, decision calibration leads to the constrained swap regret bound.

Theorem 39 *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent runs Algorithm 4 to compete with the benchmark class $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$ over every subsequence $S \in \mathcal{S}$, where the threshold parameter for each subsequence S is set to $4\sqrt{|S| \ln \frac{|\mathcal{A}||\mathcal{N}||\mathcal{S}|^2 J |S|}{\delta}}$. If the sequence of*

predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \beta_2)$ -decision calibrated, then the $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$ -constrained swap regret of any agent over any subsequence $S \in \mathcal{S}$ is bounded by:

$$\text{Reg}_{\text{swap}}(u, \mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}, S) \leq 2L|\mathcal{A}|\beta_2(|S|/|\mathcal{A}|).$$

Similarly to the proof of Theorem 20, for any subsequence $S \in \mathcal{S}$ and action modification rule $\phi : \mathcal{A} \rightarrow \mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$, we can decompose the regret against ϕ into three parts as:

$$\begin{aligned} & \sum_{t \in S} (u(\phi(a_t), y_t) - u(a_t, y_t)) \\ &= \sum_{t \in S} (u(\phi(a_t), y_t) - u(\phi(a_t), p_t)) + \sum_{t \in S} (u(\phi(a_t), p_t) - u(a_t, p_t)) + \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)). \end{aligned}$$

Conditioning on the event in Corollary 30 and the fact that $\phi(a_t) \in \mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$, $\phi(a_t) \in U_t$ for every $t \in S$. Since agents U_t -constrained best response to the prediction p_t in round t , $a_t = \arg \max_{b \in U_t} u(b, p_t)$. This implies that:

$$u(\phi(a_t), p_t) \leq u(a_t, p_t).$$

So, the second part in the decomposition is upper bounded by 0.

Thus, it suffices to bound the other two parts using decision calibration.

Lemma 40 *If the sequence of predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \beta_2)$ -decision calibrated, then for any $(u, \mathbf{c}) \in \mathcal{N}$ and $S \in \mathcal{S}$:*

$$\left| \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)) \right| \leq L|\mathcal{A}|\beta_2(|S|/|\mathcal{A}|).$$

Proof Using the linearity of u , we can write:

$$\begin{aligned} \left| \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (u(a, p_t) - u(a, y_t)) \right| \\ &= \left| \sum_{a \in \mathcal{A}} \left(u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} \left| u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(a, \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (p_t - y_t) \right\|_{\infty} \\ &\leq L \sum_{a \in \mathcal{A}} \beta_2(T^{u, \mathbb{E}[\mathbf{c}], S}(a)) \\ &\leq L|\mathcal{A}|\beta_2(|S|/|\mathcal{A}|). \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , the third inequality follows from $(\mathcal{N}, \mathcal{S}, \beta_2)$ -decision calibration, and the fourth inequality follows from concavity of β_2 . \blacksquare

Lemma 41 *If the sequence of predictions p_1, \dots, p_T is $(\mathcal{N}, \mathcal{S}, \beta_2)$ -decision calibrated, then for any $(u, \mathbf{c}) \in \mathcal{N}$ and $S \in \mathcal{S}$:*

$$\left| \sum_{t \in S} (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| \leq L|\mathcal{A}|\beta_2(|S|/|\mathcal{A}|).$$

Proof Using the linearity of u , we can write:

$$\begin{aligned} \left| \sum_{t \in S} (u(\phi(a_t), p_t) - u(\phi(a_t), y_t)) \right| &= \left| \sum_{a \in \mathcal{A}} \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (u(\phi(a), p_t) - u(\phi(a), y_t)) \right| \\ &= \left| \sum_{a \in \mathcal{A}} \left(u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} \left| u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] p_t \right) - u \left(\phi(a), \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] y_t \right) \right| \\ &\leq \sum_{a \in \mathcal{A}} L \left\| \sum_{t=1}^T \mathbb{1}[t \in S, a_t = a] (p_t - y_t) \right\|_{\infty} \\ &\leq L \sum_{a \in \mathcal{A}} \beta_2(T^{u, \mathbb{E}[\mathbf{c}], S}(a)) \\ &\leq L|\mathcal{A}|\beta_2(|S|/|\mathcal{A}|). \end{aligned}$$

where the first inequality follows from the triangle inequality, the second inequality follows from L -Lipschitzness of u , the third inequality follows from $(\mathcal{N}, \mathcal{S}, \beta_2)$ -decision calibration, and the fourth inequality follows from concavity of β_2 . \blacksquare

We can now complete the proof of Theorem 39.

Proof Fix any $(u, \mathbf{c}) \in \mathcal{N}$, $S \in \mathcal{S}$, and $\phi : \mathcal{A} \rightarrow \mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$. Applying Lemmas 40 and 41 to the decomposition of the regret against ϕ , we have that:

$$\begin{aligned} \sum_{t \in S} (u(a, y_t) - u(a_t, y_t)) &= \sum_{t \in S} (u(a, y_t) - u(a, p_t)) + \sum_{t \in S} (u(a, p_t) - u(a_t, p_t)) + \sum_{t \in S} (u(a_t, p_t) - u(a_t, y_t)) \\ &\leq 2L|\mathcal{A}|\beta_2(|S|/|\mathcal{A}|). \end{aligned}$$

\blacksquare

We will again instantiate UNBIASED-PREDICTION to make decision calibrated predictions.

Theorem 42 *Let \mathcal{S} be a collection of subsequences. Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. Each agent will run Algorithm 4 to compete with the benchmark class $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$ over every subsequence $S \in \mathcal{S}$, where the threshold parameter for each subsequence S is set to $4\sqrt{|S| \ln \frac{|\mathcal{A}||\mathcal{N}||S|^2 J |S|}{\delta}}$. Each agent will maintain a candidate action set $\{U_t\}_{t \in [T]}$. There is*

an instantiation of UNBIASED-PREDICTION (Noarov et al., 2023) —which we call DECISION-CALIBRATION-SUBSEQUENCE— producing predictions $p_1, \dots, p_T \in \mathcal{Y}$ satisfying, with probability at least $1 - \delta$, for any $(u, \mathbf{c}) \in \mathcal{N}$, $a \in \mathcal{A}$, and $S \in \mathcal{S}$:

$$\begin{aligned} & \left\| \sum_{t=1}^T \mathbb{1}[t \in S, \text{CBR}_{U_t}^u(p_t) = a](p_t - y_t) \right\|_{\infty} \\ & \leq O \left(\ln(d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T) + \sqrt{\ln(d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T) \cdot T^{u, \mathbb{E}[\mathbf{c}], S}(a)} + \sqrt{\ln \frac{d|\mathcal{A}||\mathcal{N}||\mathcal{S}|}{\delta} \cdot |S|} \right). \end{aligned}$$

Substituting the above bounds into Theorem 39 and combining with Theorem 31, we arrive at the following corollary bounding the cumulative constraint violation and the $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$ -constrained swap regret over every subsequence $S \in \mathcal{S}$. To obtain an overall failure probability of at most δ , we instantiate the regret guarantee and the cumulative-constraint-violation guarantee each with failure probability $\delta/2$, and then apply a union bound.

Corollary 43 *Let \mathcal{N} be a set of agents, where each agent is equipped with a utility function $u : \mathcal{A} \times \mathcal{Y} \rightarrow [0, 1]$ and J constraint functions $\{c_j : \mathcal{A} \times \mathcal{Y} \rightarrow [-1, 1]\}_{j \in [J]}$. The utility functions are linear and L -Lipschitz in the second argument. Each agent will run Algorithm 4 to compete with the benchmark class $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$ over every subsequence $S \in \mathcal{S}$, where the threshold parameter for each subsequence S is set to $4\sqrt{|S| \ln \frac{2|\mathcal{A}||\mathcal{N}||\mathcal{S}|^2 J |S|}{\delta}}$. The sequence of predictions p_1, \dots, p_T outputted by DECISION-CALIBRATION-SUBSEQUENCE ensures that with probability at least $1 - \delta$, the cumulative constraint violation and the $\mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}$ -constrained swap regret of any agent over any subsequence $S \in \mathcal{S}$ are bounded by:*

$$\begin{aligned} \text{CCV}(S) & \leq \sum_{S' \in \mathcal{S}} 4|\mathcal{A}| \sqrt{|S'| \ln \frac{2|\mathcal{A}||\mathcal{N}||\mathcal{S}|^2 J |S'|}{\delta}} + |\mathcal{A}||S|, \\ \text{Reg}_{\text{swap}}(u, \mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}, S) & \leq O \left(L|\mathcal{A}| \sqrt{|S| \ln \frac{d|\mathcal{A}||\mathcal{N}||\mathcal{S}|T}{\delta}} \right). \end{aligned}$$

Since constrained swap regret is stronger than constrained external regret, the same regret bound also holds for $\text{Reg}_{\text{ext}}(u, \mathcal{A}_S^{\mathbb{E}[\mathbf{c}]}, S)$.

Appendix G. Unbiased Prediction Algorithm

In this section we present the UNBIASED-PREDICTION algorithm of Noarov et al. (2023), from which our guarantees in Theorems 17, 34, 23, and 42 follow.

We first introduce several notations and concepts from Noarov et al. (2023). Let $\Pi = \{(x, p, y) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Y}\}$ denote the set of possible realized triples at each round. An interaction over T rounds produces a transcript $\pi_T \in \Pi^T$. We write $\pi_T^{<t}$ as the prefix of the first $t - 1$ triples in π_T , for any $t \leq T$. We write $\Pi^* = \cup_{T=1}^{\infty} \Pi^T$ for the space of all transcripts. An event $E \in \mathcal{E}$ is a mapping from transcripts, contexts and predictions to $[0, 1]$, i.e. $E : \Pi^* \times \mathcal{X} \times \mathcal{Y} \rightarrow [0, 1]$.

The UNBIASED-PREDICTION algorithm makes predictions that are unbiased conditional on a collection of events \mathcal{E} . The algorithm's conditional bias guarantee depends logarithmically on the number of events:

Theorem 44 (Noarov et al., 2023) For a collection of events \mathcal{E} and convex prediction/outcome space $\mathcal{Y} \subseteq [0, 1]^d$, Algorithm 5 produces predictions $\psi_1, \dots, \psi_T \in \Delta\mathcal{Y}$ such that for any sequence of outcomes $y_1, \dots, y_T \in \mathcal{Y}$ chosen by the adversary:

$$\left\| \sum_{t=1}^T \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \leq O\left(\ln(d|\mathcal{E}|T) + \sqrt{\ln(d|\mathcal{E}|T) \cdot n_T(E)}\right).$$

where $n_T(E) = \sum_{t=1}^T E(\pi_T^{\leq t}, x_t, p_t)$. The algorithm can be implemented with per-round running time scaling polynomially in d and $|\mathcal{E}|$.

Algorithm 5 UNBIASED-PREDICTION

for $t = 1$ **to** T **do**

 Observe x_t

 Define the distribution $q_t \in \Delta[2d|\mathcal{E}|]$ such that for $E \in \mathcal{E}, i \in [d], \sigma \in \{\pm 1\}$,

$$q_t^{E,i,\sigma} \propto \exp\left(\frac{\eta}{2} \sum_{s=1}^{t-1} \sigma \cdot \mathbb{E}_{p_s \sim \psi_s} [E(\pi_T^{\leq s}, x_s, p_s)(p_s^i - y_s^i)]\right);$$

 Output the solution to the minmax problem:

$$\psi_t \leftarrow \arg \min_{\psi'_t \in \Delta\mathcal{Y}} \max_{y \in \mathcal{Y}} \mathbb{E}_{p_t \sim \psi'_t} \left[\sum_{E,i,\sigma} q_t(E,i,\sigma) \cdot \sigma \cdot E(\pi_T^{\leq s}, x_s, p_s) \cdot (p_s^i - y_s^i) \right];$$

end

Theorem 44 is stated as expected error bounds over randomized predictions $\psi_t \in \Delta\mathcal{Y}$. In the following Corollary 45, we state the guarantees based on realized predictions p_t that are sampled from ψ_t . Our guarantees in Theorems 17 and 34 directly follow from Corollary 45.

Corollary 45 For a collection of events \mathcal{E} and convex prediction/outcome space $\mathcal{Y} \subseteq [0, 1]^d$, Algorithm 5 produces predictions $p_1, \dots, p_T \in \mathcal{Y}$ such that for any sequence of outcomes $y_1, \dots, y_T \in \mathcal{Y}$ chosen by the adversary, with probability at least $1 - \delta$:

$$\left\| \sum_{t=1}^T E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) \right\|_{\infty} \leq O\left(\ln(d|\mathcal{E}|T) + \sqrt{\ln(d|\mathcal{E}|T) \cdot n_T(E)} + \sqrt{\ln(d|\mathcal{E}|/\delta) \cdot T}\right).$$

where $n_T(E) = \sum_{t=1}^T E(\pi_T^{\leq t}, x_t, p_t)$. The algorithm can be implemented with per-round running time scaling polynomially in d and $|\mathcal{E}|$.

Proof Fix any $m \in [d]$ and $E \in \mathcal{E}$. Consider the sequence $\{E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m}) - \mathbb{E}_{p_t \sim \psi_t}[E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m})]\}_{t=1}^T$, where $p_{t,m}$ and $y_{t,m}$ are the m -th coordinate of p_t and y_t , respectively. It is a sequence of martingale differences, since for any $t \in [T]$:

$$\mathbb{E} \left[E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m}) - \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m})] \mid \sigma(\pi_T^{\leq t}) \right] = 0.$$

By Azuma-Hoeffding inequality (Lemma 47), we have that with probability at least $1 - \frac{\delta}{d|\mathcal{E}|}$:

$$\left| \sum_{t=1}^T E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m}) - \sum_{t=1}^T \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m})] \right| \leq 2\sqrt{2 \ln(2d|\mathcal{E}|/\delta) \cdot T}.$$

Using the union bound over all $m \in [d]$ and $E \in \mathcal{E}$, we have that with probability at least $1 - \delta$, for any $m \in [d]$ and any $E \in \mathcal{E}$:

$$\left\| \sum_{t=1}^T E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) - \sum_{t=1}^T \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \leq 2\sqrt{2 \ln(2d|\mathcal{E}|/\delta) \cdot T}.$$

As a result,

$$\begin{aligned} \left\| \sum_{t=1}^T E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) \right\|_{\infty} &\leq \left\| \sum_{t=1}^T \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \\ &\quad + \left\| \sum_{t=1}^T E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) - \sum_{t=1}^T \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \\ &\leq O\left(\ln(d|\mathcal{E}|T) + \sqrt{\ln(d|\mathcal{E}|T) \cdot n_T(E)}\right) + 2\sqrt{2 \ln(2d|\mathcal{E}|/\delta) \cdot T} \\ &= O\left(\ln(d|\mathcal{E}|T) + \sqrt{\ln(d|\mathcal{E}|T) \cdot n_T(E)} + \sqrt{\ln(d|\mathcal{E}|/\delta) \cdot T}\right). \end{aligned}$$

■

We further extend the guarantee to the multi-subsequence setting in Corollary 46. Our guarantees in Theorems 23 and 42 directly follow from Corollary 46.

Corollary 46 *Let \mathcal{S} be a collection of subsequences. For a collection of events \mathcal{E} and convex prediction/outcome space $\mathcal{Y} \subseteq [0, 1]^d$, Algorithm 5 instantiated with the collection of events $\{\mathbb{1}[t \in S] \cdot E\}_{S \in \mathcal{S}, E \in \mathcal{E}}$ produces predictions $p_1, \dots, p_T \in \mathcal{Y}$ such that for any sequence of outcomes $y_1, \dots, y_T \in \mathcal{Y}$ chosen by the adversary, with probability at least $1 - \delta$, for any $S \in \mathcal{S}$:*

$$\left\| \sum_{t \in S} E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) \right\|_{\infty} \leq O\left(\ln(d|\mathcal{E}||\mathcal{S}|T) + \sqrt{\ln(d|\mathcal{E}||\mathcal{S}|T) \cdot n_S(E)} + \sqrt{\ln(d|\mathcal{E}||\mathcal{S}|/\delta) \cdot |S|}\right).$$

where $n_S(E) = \sum_{t \in S} E(\pi_T^{\leq t}, x_t, p_t)$. The algorithm can be implemented with per-round running time scaling polynomially in d , $|\mathcal{E}|$, and $|\mathcal{S}|$.

Proof Fix any $m \in [d]$, $E \in \mathcal{E}$, and $S \in \mathcal{S}$. Consider the sequence $\{E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m}) - \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m})]\}_{t=1}^T$, where $p_{t,m}$ and $y_{t,m}$ are the m -th coordinate of p_t and y_t , respectively. It is a sequence of martingale differences, since for any $t \in [T]$:

$$\mathbb{E} \left[E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m}) - \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m})] \mid \sigma(\pi_T^{\leq t}, x_t) \right] = 0.$$

The subsequence of these terms corresponding to rounds $s \in S$, i.e., $\{E(\pi_T^{\leq s}, x_s, p_s)(p_{s,m} - y_{s,m}) - \mathbb{E}_{p_s \sim \psi_s}[E(\pi_T^{\leq s}, x_s, p_s)(p_{s,m} - y_{s,m})]\}_{s \in S: s \leq t}$, is also a martingale difference sequence, because the selection rule is predictable with respect to the filtration $\sigma(\pi_T^{\leq s}, x_s)$.

By Azuma-Hoeffding inequality (Lemma 47), we have that with probability at least $1 - \frac{\delta}{d|\mathcal{E}||S|}$:

$$\left| \sum_{t \in S} E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m}) - \sum_{t \in S} \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_{t,m} - y_{t,m})] \right| \leq 2\sqrt{2 \ln(2d|\mathcal{E}||S|/\delta) \cdot |S|}.$$

Using the union bound over all $m \in [d]$, $E \in \mathcal{E}$, and $S \in \mathcal{S}$, we have that with probability at least $1 - \delta$, for any $m \in [d]$, any $E \in \mathcal{E}$, and any $S \in \mathcal{S}$:

$$\left\| \sum_{t \in S} E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) - \sum_{t \in S} \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \leq 2\sqrt{2 \ln(2d|\mathcal{E}||S|/\delta) \cdot |S|}.$$

By Theorem 44, for any $E \in \mathcal{E}$ and any $S \in \mathcal{S}$:

$$\begin{aligned} \left\| \sum_{t \in S} \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} &= \left\| \sum_{t=1}^T \mathbb{E}_{p_t \sim \psi_t} [\mathbb{1}[t \in S] E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \\ &\leq O\left(\ln(d|\mathcal{E}||S|T) + \sqrt{\ln(d|\mathcal{E}||S|T) \cdot n_S(E)}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} \left\| \sum_{t \in S} E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) \right\|_{\infty} &\leq \left\| \sum_{t \in S} \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \\ &\quad + \left\| \sum_{t \in S} E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t) - \sum_{t \in S} \mathbb{E}_{p_t \sim \psi_t} [E(\pi_T^{\leq t}, x_t, p_t)(p_t - y_t)] \right\|_{\infty} \\ &\leq O\left(\ln(d|\mathcal{E}||S|T) + \sqrt{\ln(d|\mathcal{E}||S|T) \cdot n_S(E)}\right) + 2\sqrt{2 \ln(2d|\mathcal{E}||S|/\delta) \cdot |S|} \\ &= O\left(\ln(d|\mathcal{E}||S|T) + \sqrt{\ln(d|\mathcal{E}||S|T) \cdot n_S(E)} + \sqrt{\ln(d|\mathcal{E}||S|/\delta) \cdot |S|}\right). \end{aligned}$$

■

Theorem 17 instantiates Algorithm 5 with the collection of events $\mathcal{E} = \{\mathbb{1}[\text{CBR}_{\hat{\mathcal{A}}_t^c}^u(p_t) = a]\}_{(u,c) \in \mathcal{N}, a \in \mathcal{A}}$. Thus, $|\mathcal{E}| = |\mathcal{A}||\mathcal{N}|$, from which our guarantees follow.

Similarly, Theorem 23 instantiates Algorithm 5 with the collection of events $\mathcal{E} = \{\mathbb{1}[\text{CBR}_{\hat{\mathcal{A}}_t^{\mathbb{E}[c]}}^u(p_t) = a]\}_{(u,c) \in \mathcal{N}, a \in \mathcal{A}}$. Thus, $|\mathcal{E}| = |\mathcal{A}||\mathcal{N}|$ as well.

In the multi-subsequence setting, Theorems 34 and 42 instantiate Algorithm 5 with the collection of events $\mathcal{E} = \{\mathbb{1}[\text{CBR}_{U_t}^u(p_t) = a]\}_{(u,c) \in \mathcal{N}, a \in \mathcal{A}}$. Thus, $|\mathcal{E}| = |\mathcal{A}||\mathcal{N}|$.

We note that $\hat{\mathcal{A}}_t^c$, $\hat{\mathcal{A}}_t^{\mathbb{E}[c]}$, and U_t are constructed based on the actions' performances in the first $t - 1$ rounds. Therefore, $\mathbb{1}[\text{CBR}_{\hat{\mathcal{A}}_t^c}^u(p_t) = a]$, $\mathbb{1}[\text{CBR}_{\hat{\mathcal{A}}_t^{\mathbb{E}[c]}}^u(p_t) = a]$, and $\mathbb{1}[\text{CBR}_{U_t}^u(p_t) = a]$ are indeed functions of $\pi_T^{\leq t}$ and p_t , hence they qualify for the events in Theorem 44.

Appendix H. Azuma-Hoeffding's Inequality

Lemma 47 *If X_1, \dots, X_T is a martingale difference sequence, and for every t , with probability 1, $|X_t| \leq M_t$. Then with probability at least $1 - 2e^{-\frac{\epsilon^2}{2\sum_{t=1}^T M_t^2}}$:*

$$\left| \sum_{t=1}^T X_t \right| \leq \epsilon.$$