PERFORMATIVE SUBMODULAR DECISION MAKING

Anonymous authors
Paper under double-blind review

ABSTRACT

Modern decision-making systems increasingly operate under feedback, where deployed decisions alter future data distributions. Many such systems involve combinatorial subset selection, for which submodularity provides a useful diminishing-returns structure and efficient approximation methods. However, decision-dependent distributions can invalidate the structural and algorithmic guarantees of classical stochastic submodular optimization. We study this setting through *performative submodular decision making*, where the selected set both determines the utility and induces its evaluation distribution. We show that the resulting performative objective may lose monotonicity and submodularity even when $f(\cdot, \xi)$ is monotone and submodular for every fixed ξ, and that repeated best-response retraining may decrease deployed performance or cycle. To recover tractable structure, we decouple the deployed set from the candidate set being optimized. For any fixed deployment, the resulting frozen objective remains monotone and submodular, leading to approximate performative stability as a self-consistent solution concept. Under distribution-sensitivity conditions, approximate stability yields a global approximation guarantee for the original performative objective, up to an additive feedback-dependent penalty. We further develop a finite-sample feedback-safe retraining method that, with high probability, improves the performative objective along every accepted redeployment and provides a terminal quality guarantee upon rejection. Experiments on decision-dependent weighted coverage illustrate structural violations, retraining instability, and the reliability–conservativeness tradeoff of certification.

1 INTRODUCTION

Combinatorial decision making is a central problem in machine learning and optimization: a decision maker often needs to select a subset of actions, items, sensors, resources, or individuals from a large ground set. Submodular maximization provides a standard framework for such subset-selection problems through diminishing-returns and admits efficient approximation algorithms under cardinality and related constraints (Nemhauser et al., 1978; Krause & Golovin, 2014). This structure underlies applications including influence maximization, sensor placement, data summarization and resource allocation (Kempe et al., 2003; Krause et al., 2008; Bilmes, 2022). Classical stochastic submodular optimization, however, typically assumes that the uncertainty distribution is fixed, or independent of the selected set. This assumption fails in deployed systems where decisions influence future data distributions. For example, in recommendation, a selected set of items may provide diminishing-return coverage of user interests, while repeated exposure can reshape future preference distributions. Consequently, an improvement measured under the current preference distribution need not translate into improved performance after deployment. Such feedback is central to performative prediction (Perdomo et al., 2020; Mendler-Dünner et al., 2020) and more broadly motivates decision-dependent optimization (Drusvyatskiy & Xiao, 2023).

Motivated by this tension between submodular structure and performative feedback, we study *performative submodular decision making*, where the selected set induces the distribution under which it is evaluated. We consider

$$\max_{S \in \mathcal{I}} F(S), \qquad F(S) := \mathbb{E}_{\xi \sim \mathcal{D}(S)}[f(S, \xi)], \tag{1}$$

where $\mathcal{I}$ denotes the feasible family. Although $f(\cdot, \xi)$ is normalized, monotone, and submodular for every realization ξ, the decision-dependent distribution can destroy these properties at the level of F.

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

This gives rise to two fundamental challenges. First, the decision S simultaneously determines the selected elements and the distribution under which they are evaluated. As a result, the performative objective F may lose monotonicity and submodularity, preventing a direct application of classical greedy approximation guarantees. Second, retraining creates a dynamic mismatch: a candidate is optimized under the current deployment distribution, but once deployed it changes the distribution under which it is evaluated. Thus, retraining may decrease the deployed objective or lead to cycles.

Our main contributions are summarized as follows.

- **Performative submodular framework and structural characterization.** We formulate *performative submodular decision making*, where the selected set also determines its evaluation distribution. We show that this coupling fundamentally differs from classical stochastic submodular optimization: the performative objective may lose monotonicity and submodularity, and repeated best-response retraining may decrease deployed performance or cycle. By decoupling deployment from optimization, we recover a frozen objective that remains monotone and submodular for every fixed deployment, providing a tractable approximate best-response framework (Section 4).
- **From performative stability to global approximation.** We introduce approximate performative stability as a self-consistent solution concept and establish a stability-to-optimality connection. An α-approximately performatively stable solution achieves an α-approximation to the performative optimum up to an explicit feedback-dependent additive error, which vanishes in the fixed-distribution case and scales with the distribution sensitivity and the diameter of the feasible family (Section 5).
- **Certified retraining under feedback.** We develop an a posteriori certified greedy algorithm and a finite-sample feedback-safe greedy algorithm with guarantees. Under a cardinality constraint, the former returns a $(1-1/e, \tau)$-approximately performatively stable terminal solution, together with a global approximation guarantee for the performative objective. The latter ensures, with high probability, that every accepted redeployment improves the performative objective, while rejection yields a terminal quality guarantee (Section 6).

Experiments on decision-dependent weighted coverage illustrate structural violations, retraining instability, and the reliability–conservativeness tradeoff of certification (Section 7).

2 Related work

Performative and decision-dependent optimization. Performative prediction formalizes optimization when deploying a model changes the data distribution (Perdomo et al., 2020). Repeated risk minimization and stochastic gradient methods have been analyzed under distribution sensitivity, convexity, and smoothness (Mendler-Dünner et al., 2020; Drusvyatskiy & Xiao, 2023), with saddle-point and constrained extensions in (Wood & Dall'Anese, 2023; Wang et al., 2025). Subsequent work has considered neural models, distributional robustness, convergence limitations, and explicit distribution dynamics (Mofakhami et al., 2023; Jia et al., 2025; Khorsandi et al., 2025; He et al., 2025). Recent zeroth-order methods further extend the framework to nonsmooth nonconvex objectives (Liu et al., 2026); see (Hardt & Mendler-Dünner, 2025) for a survey.

Endogenous uncertainty in stochastic programming. Decision-dependent uncertainty has also been studied in stochastic programming under the names *endogenous uncertainty* and *decision-dependent random elements*; see Hellemo et al. (2018). Existing work includes structured stochastic programs, combinatorial models, and decision-dependent ambiguity formulations (Bhuiyan, 2021; Pantuso & Hewitt, 2025; Pantuso, 2025; Yu & Basciftci, 2026).

Stochastic and adaptive submodular maximization. Classical greedy achieves a $(1 - 1/e)$-approximation for monotone submodular maximization under a cardinality constraint (Nemhauser et al., 1978). Adaptive submodularity extends submodular optimization to sequential decision making under uncertainty, but assumes an exogenous state distribution rather than one altered by the selected set (Golovin & Krause, 2011).

Taken together, these lines of work leave a gap at the intersection of set-valued decisions, decision-dependent distributions, and submodular approximation. Our framework addresses this intersection

Table 1: Representative comparison with related optimization frameworks. Entries summarize typical formulations and guarantees in the cited literature. Here, DDD indicates whether the data distribution depends on the deployed decision.

Framework	Decision	DDD	Solution notion
Performative prediction (Perdomo et al., 2020)	Vector-valued	Yes	Performative stability
Endogenous uncertainty (Hellemo et al., 2018)	Problem-specific	Yes	Closed-loop optimum
Submodular maximization (Nemhauser et al., 1978)	Set-valued	No	Approx. optimizer
This work	Set-valued	Yes	Approx. performative stability & global approximation

by recovering submodular structure under a frozen deployment and relating approximate performative stability to global performative approximation. Table 1 summarizes this positioning.

3 PRELIMINARIES

In this section, we introduce the basic notation and formulate the stochastic submodular maximization problem with decision-dependent distributions.

Submodular decision making. Submodular maximization models set-valued decision problems with diminishing returns. Submodular functions provide a standard model for this structure. Let $V = \{1, \dots, n\}$ be a finite ground set, and let $\mathcal{I} \subseteq 2^V$ denote a downward-closed family of feasible sets. A common example is the cardinality constraint $\mathcal{I}_k := \{S \subseteq V : |S| \leq k\}$. For a set function $h : 2^V \to \mathbb{R}$, $S \subseteq V$ and $e \in V \setminus S$, we denote the marginal gain of adding e to S by

$$\Delta_e h(S) := h(S \cup \{e\}) - h(S).$$

The function h is called submodular if it satisfies the diminishing-returns property: $\Delta_e h(A) \geq \Delta_e h(B)$, for all $A \subseteq B \subseteq V,\ e \notin B$. It is normalized if $h(\emptyset) = 0$, and monotone if $h(A) \leq h(B)$ whenever $A \subseteq B \subseteq V$. For two sets $S, T \subseteq V$, their symmetric difference is denoted by $S \triangle T := (S \setminus T) \cup (T \setminus S)$.

Classical greedy algorithm. For monotone submodular maximization under a cardinality constraint $\mathcal{I}_k$, the classical greedy algorithm repeatedly selects the element with the largest current marginal gain. Starting from $S_0 = \emptyset$, it updates

$$e_t \in \arg\max_{e \in V \setminus S_t} \Delta_e h(S_t), \qquad S_{t+1} = S_t \cup \{e_t\}, \quad t = 0, \dots, k-1.$$

Submodularity yields diminishing marginal returns along the greedy sequence, leading to the classical approximation guarantee $h(S_k) \geq (1 - 1/e) \max_{|S| \leq k} h(S)$ for normalized monotone submodular functions (Nemhauser et al., 1978).

Problem statement. We consider the performative submodular maximization problem introduced in equation 1, where each feasible set $S \in \mathcal{I}$ induces a distribution $\mathcal{D}(S)$. Here $f(S, \xi)$ denotes the utility of selecting S under state ξ. We assume sample access to $\mathcal{D}(S)$, without requiring its explicit form. For every fixed $\xi \in \Omega$, $f(\cdot, \xi)$ is assumed to be normalized, monotone, and submodular. The key distinction is that S is both the decision being evaluated and the source of its evaluation distribution. To separate these roles, we define the decoupled objective

$$G(S; T) := \mathbb{E}_{\xi \sim \mathcal{D}(T)}[f(S, \xi)], \tag{2}$$

where T is the deployed set inducing the distribution and S is the candidate set evaluated under that distribution. For fixed T, we refer to $G(\cdot; T)$ as the *frozen objective*. The performative objective is recovered on the diagonal: $F(S) = G(S; S)$. We denote the optimal performative value by $\mathrm{OPT} := \max_{S \in \mathcal{I}} F(S)$. We measure solution quality through guarantees of the form

$$F(\widehat{S}) \geq \alpha\, \mathrm{OPT} - \varepsilon, \qquad \alpha \in (0, 1], \quad \varepsilon \geq 0.$$

When $\varepsilon = 0$, this reduces to an α-approximation. For a fixed deployment T, the frozen objective separates optimization over the candidate set from the induced distribution, which motivates the approximate best-response formulation developed in Section 4.

4 STRUCTURAL AND DYNAMIC CHALLENGES

We identify two difficulties induced by performative feedback. First, although $f(\cdot, \xi)$ is monotone and submodular for every fixed ξ, the performative objective F need not preserve these properties. Second, even after freezing the deployment distribution to recover submodularity, repeated best-response updates may remain unstable. Proofs are deferred to Appendix A.

4.1 STRUCTURAL LOSS AND APPROXIMATE STABILITY

Although $f(\cdot, \xi)$ is normalized, monotone, and submodular for every realization ξ, the performative objective F need not preserve these properties. The difficulty is that the same set S both determines the selected elements and induces the evaluation distribution. For $e \notin S$,

$$\Delta_e F(S) = \underbrace{G(S \cup \{e\}; S) - G(S; S)}_{\text{fixed-distribution marginal gain}} + \underbrace{G(S \cup \{e\}; S \cup \{e\}) - G(S \cup \{e\}; S)}_{\text{distribution feedback term}}.$$

The first term is a standard marginal gain under the fixed distribution $\mathcal{D}(S)$, for which the underlying submodular structure is preserved. The second term captures the additional change induced by shifting the distribution from $\mathcal{D}(S)$ to $\mathcal{D}(S \cup \{e\})$, and is not controlled by the pointwise submodularity of f. Hence, pointwise submodularity alone is insufficient to guarantee that the performative objective F remains submodular.

Proposition 4.1 (Loss of monotonicity and submodularity)**.** *Even if $f(\cdot, \xi)$ is monotone and submodular for every fixed $\xi \in \Omega$, the performative objective $F(S)$ need not be submodular or monotone.*

This loss of structure arises because the decision and its induced distribution vary simultaneously. Fixing the deployed decision T in the decoupled objective $G(S; T)$ holds the distribution fixed while optimizing S, thereby recovering the underlying submodular structure.

Proposition 4.2 (Frozen objectives preserve submodularity)**.** *Suppose that $f(\cdot, \xi)$ is normalized, monotone, and submodular for every $\xi \in \Omega$. Then, for every fixed $T \in \mathcal{I}$, the frozen objective $G(\cdot; T)$ is normalized, monotone, and submodular.*

Thus, for each fixed deployment T, optimizing $G(\cdot; T)$ is a standard submodular maximization problem. For $\alpha \in (0, 1]$ and $\varepsilon \geq 0$, define the (α, ε)-approximate best-response set

$$\mathrm{BR}_{\alpha,\varepsilon}(T) := \left\{ S \in \mathcal{I} : G(S; T) \geq \alpha \max_{R \in \mathcal{I}} G(R; T) - \varepsilon \right\}.$$

Because redeploying a response may change the induced distribution, a response to T need not remain approximately optimal after deployment. This motivates a self-consistency condition in which the deployed set is itself an approximate best response under its own induced distribution.

Definition 4.1 ((α, ε)-approximate performative stability)**.** For $\alpha \in (0, 1]$ and $\varepsilon \geq 0$, a set $S^* \in \mathcal{I}$ is called (α, ε)-approximately performatively stable if $S^* \in \mathrm{BR}_{\alpha,\varepsilon}(S^*)$, or equivalently,

$$G(S^*; S^*) \geq \alpha \max_{S \in \mathcal{I}} G(S; S^*) - \varepsilon.$$

When $\varepsilon = 0$, we simply say that S^* is α-approximately performatively stable.

4.2 INSTABILITY OF REPEATED BEST RESPONSES

Repeated optimization of the frozen objective does not necessarily yield stable deployment dynamics. Given the current deployment S_t, suppose that a response S_{t+1} improves the frozen objective:

$$G(S_{t+1}; S_t) \geq G(S_t; S_t).$$

This improvement is evaluated under the distribution induced by S_t. After redeployment, however, the relevant objective is $F(S_{t+1}) = G(S_{t+1}; S_{t+1})$. Hence,

$$G(S_{t+1}; S_t) \geq G(S_t; S_t) \quad \text{does not imply} \quad F(S_{t+1}) \geq F(S_t).$$

Hence, repeated frozen best-response updates may decrease the performative objective or exhibit cyclic behavior.

Proposition 4.3 (Repeated best responses may cycle)**.** *There exist instances for which every frozen objective $G(\cdot; T)$ is a monotone submodular function, yet exact best-response dynamics $S_{t+1} \in \mathrm{BR}_{1,0}(S_t)$ may decrease the performative objective and enter a nontrivial cycle. Moreover, an exact performatively stable set need not exist.*

Appendix A.3 gives a concrete two-cycle example exhibiting both performance degradation and failure of exact stability.

5 Sensitivity Analysis and Stability-to-Optimality Guarantees

In this section, building on the notion of approximate performative stability, we establish a connection between stability and global performance. We first quantify the sensitivity of the induced distribution to changes in the deployed set, and then show that an approximately performatively stable solution achieves a global approximation guarantee for the original performative objective, up to an additive error controlled by the feedback sensitivity. Proofs are deferred to Appendix B.

5.1 Distributional Sensitivity

We first quantify how changes in the deployed set affect the induced distribution.

Assumption 5.1 (Set-Lipschitz distribution map)**.** There exists $\lambda \geq 0$ such that for all $S, T \in \mathcal{I}$,

$$W_1(\mathcal{D}(S), \mathcal{D}(T)) \leq \lambda |S \triangle T|,$$

where W_1 denotes the Wasserstein-1 distance on the state space (Ω, d_Ω).

The constant λ measures the sensitivity of the induced distribution to changes in the deployed set. When $\lambda = 0$, the distribution is independent of the decision, and the problem reduces to stochastic submodular maximization under a fixed distribution. A Lipschitz common-latent-source representation provides a sufficient condition for Assumption 5.1; see Appendix B.1.

We also assume that the utility is Lipschitz continuous with respect to the state variable.

Assumption 5.2 (State-Lipschitz utility)**.** There exists $L_\xi \geq 0$ such that, for all $S \in \mathcal{I}$ and $\xi, \xi' \in \Omega$, we have $|f(S, \xi) - f(S, \xi')| \leq L_\xi d_\Omega(\xi, \xi')$.

Combining Assumptions 5.1 and 5.2, define $\rho := L_\xi \lambda$. The following lemma controls the variation of the decoupled objective with respect to the deployed set.

Lemma 5.1 (Distributional sensitivity of values)**.** *Suppose that Assumptions 5.1 and 5.2 hold. Then, for any $A, S, T \in \mathcal{I}$, we have $|G(A; S) - G(A; T)| \leq \rho |S \triangle T|$.*

Lemma 5.1 shows that changing the deployment from S to T perturbs the value of any fixed candidate A by at most a term proportional to the set distance. The parameter $\rho = L_\xi \lambda$ therefore captures the combined effect of distributional sensitivity and state sensitivity, and will control the additive error in the stability-to-optimality guarantees below.

5.2 Stability-to-Optimality Guarantees

We next connect approximate performative stability to global performance under the original performative objective. Lemma 5.1 controls the distributional mismatch needed to compare a stable solution with a performative optimizer, leading to the following guarantee.

Theorem 5.1 (From stability to performative optimality)**.** *Suppose that Assumptions 5.1 and 5.2 hold. If $\widehat{S} \in \mathcal{I}$ is α-approximately performatively stable, then*

$$F(\widehat{S}) \geq \alpha \max_{R \in \mathcal{I}} F(R) - \alpha \rho \Delta_{\mathcal{I}},$$

where $\Delta_{\mathcal{I}} := \max_{S,T \in \mathcal{I}} |S \triangle T|$.

The bound separates the approximation factor α, inherited from approximate stability, from the additive feedback penalty $\alpha \rho \Delta_{\mathcal{I}}$. In particular, when $\mathrm{OPT} > 0$, defining $\kappa := \rho \Delta_{\mathcal{I}} / \mathrm{OPT}$, gives

the equivalent relative guarantee $F(\widehat{S}) \geq \alpha(1-\kappa)\,\mathrm{OPT}$. Thus, the classical α-approximation is recovered when $\rho = 0$, while stronger feedback induces a degradation controlled by the feedback-to-optimum ratio κ. The same argument extends to (α, ε)-approximate performative stability with an additional additive error ε; under the cardinality constraint $\mathcal{I}_k$, using $\Delta_{\mathcal{I}_k} \leq 2k$ and $\rho = L_\xi \lambda$, we obtain

$$F(\widehat{S}) \geq \alpha\,\mathrm{OPT} - 2\alpha k L_\xi \lambda - \varepsilon.$$

The stability-to-optimality argument only requires the value-sensitivity bound in Lemma 5.1. For the weighted-coverage softmax model, Proposition D.1 derives a direct model-specific constant ρ_{wc}, which can replace $\rho = L_\xi \lambda$ in the resulting guarantee.

6 Retraining Dynamics and Stability Certification

We address the dynamic challenge in Section 4 with two retraining methods: an a posteriori certification scheme for terminal stability and a feedback-safe algorithm for deployment-level safety. Proofs are deferred to Appendix C.

6.1 Repeated Best Response with A Posteriori Certificate

Given a deployment $T \in \mathcal{I}$, consider the frozen best-response problem

$$\max_{S \in \mathcal{I}} G(S; T).$$

By Proposition 4.2, $G(\cdot; T)$ is normalized, monotone, and submodular for every fixed T. We consider an α-approximate best-response oracle $\mathcal{A}$ satisfying $\mathcal{A}(T) \in \mathrm{BR}_{\alpha,0}(T)$. At iteration t, let $R_t = \mathcal{A}(S_t)$ be the approximate best response to the current deployment S_t. We define the frozen improvement gap

$$\mathrm{Gap}_t := G(R_t; S_t) - G(S_t; S_t). \tag{3}$$

For a tolerance $\tau \geq 0$, if $\mathrm{Gap}_t \leq \tau$, the procedure terminates and returns S_t; otherwise, R_t is redeployed and $S_{t+1} = R_t$. Under the cardinality constraint $\mathcal{I}_k$, we instantiate the oracle with classical greedy: $\mathcal{A}_{\mathrm{gr}}(T) := \mathrm{Greedy}(G(\cdot; T))$. By Proposition 4.2 and the greedy guarantee, $\mathcal{A}_{\mathrm{gr}}$ is a $(1 - 1/e)$-approximate best-response oracle.

Theorem 6.1 (Greedy terminal guarantee)**.** *Suppose that Assumptions 5.1 and 5.2 hold. Under the cardinality constraint $\mathcal{I}_k$, if the procedure with $\mathcal{A} = \mathcal{A}_{\mathrm{gr}}$ terminates at S_t with $\mathrm{Gap}_t \leq \tau$, then S_t is $\big(1 - 1/e, \tau\big)$-approximately performatively stable. Consequently,*

$$F(S_t) \geq \left(1 - \frac{1}{e}\right)\mathrm{OPT} - 2k\left(1 - \frac{1}{e}\right)L_\xi \lambda - \tau.$$

Under sample-only access, the greedy response can be computed from empirical marginal gains; Appendix C.1 provides a high-probability approximation guarantee for this sample-based greedy oracle. In practice, the stopping gap can be estimated using independent Monte Carlo samples.

6.2 Feedback-Safe Repeated Best Response

A candidate that improves the frozen objective may still decrease the performative objective after redeployment. We therefore introduce a feedback-safe acceptance rule that compares the frozen improvement with a certificate of the distribution-shift loss.

For a candidate response R to the current deployment T, define the frozen improvement

$$q(R, T) := G(R; T) - F(T). \tag{4}$$

Since $q(R, T)$ is evaluated under $\mathcal{D}(T)$, it need not equal the deployed improvement $F(R) - F(T)$. Let $C(R, T) \geq 0$ be any certificate satisfying

$$G(R; R) \geq G(R; T) - C(R, T), \tag{5}$$

Obviously, $C(R, T)$ upper bounds the possible loss in evaluating R after the distribution shifts from $\mathcal{D}(T)$ to $\mathcal{D}(R)$. We accept R only if $q(R, T) > C(R, T) + \tau$. Then, we have

$$F(R) \geq F(T) + q(R, T) - C(R, T) > F(T) + \tau.$$

Algorithm 1 Finite-Sample Feedback-Safe Repeated Greedy

Input: initial deployment $T_0 \in \mathcal{I}_k$, certificate C, margin $\tau \geq 0$, sample size N_{gr}, sample size N, confidence levels $\delta_{\mathrm{gr}}, \delta_{\mathrm{cert}} \in (0,1)$, and maximum number of tests K.
for $t = 0, 1, \ldots, K-1$ **do**
  Compute $R_t \leftarrow \widehat{\mathcal{A}}_{\mathrm{gr}}(T_t)$.
  Draw $\xi_t^1, \ldots, \xi_t^N \overset{\text{i.i.d.}}{\sim} \mathcal{D}(T_t)$, independently of the samples used to compute R_t.
  Compute $\widehat{q}_N(R_t, T_t)$ according to equation 6.
  if $\widehat{q}_N(R_t, T_t) > C(R_t, T_t) + \tau + \varepsilon_N(\delta_{\mathrm{cert}})$ **then**
    Deploy R_t and set $T_{t+1} \leftarrow R_t$.
  else
    Stop and return T_t and the rejected candidate R_t.
  end if
end for
Return T_K.

Hence, every accepted redeployment improves F by more than τ and accepted redeployments cannot form a nontrivial cycle. If $0 \leq f(S, \xi) \leq M$ and $\tau > 0$, there are at most $\lfloor (M - F(T_0))/\tau \rfloor$ accepted redeployments. Under Assumptions 5.1 and 5.2, Lemma 5.1 yields the valid global certificate

$$C_{\mathrm{global}}(R, T) = \rho |R \triangle T|.$$

If only an upper bound $\bar{\rho} \geq \rho$ is known, ρ can be replaced by $\bar{\rho}$. Since this global certificate does not exploit problem-specific structure, it can be conservative. Proposition D.1 derives a tighter problem-dependent certificate C_{loc} for the weighted-coverage softmax model used in our experiments.

Finite-sample implementation. Under sample-only access, greedy responses are computed from empirical marginal gains. As shown in Appendix C.1, with probability at least $1 - \delta_{\mathrm{gr}}$, the sample-based greedy response satisfies

$$G(R_t; T_t) \geq \left(1 - \frac{1}{e}\right) \max_{S \in \mathcal{I}_k} G(S; T_t) - \eta_{\mathrm{gr}},$$

where η_{gr} is defined in Lemma C.1. Separately, we estimate the frozen improvement $q(R_t, T_t)$ using an independent certification batch. Let $\xi_t^1, \ldots, \xi_t^N \overset{\text{i.i.d.}}{\sim} \mathcal{D}(T_t)$ be independent of the samples used to construct R_t, and define

$$\widehat{q}_N(R_t, T_t) := \frac{1}{N} \sum_{i=1}^{N} \big(f(R_t, \xi_t^i) - f(T_t, \xi_t^i) \big). \tag{6}$$

For $0 \leq f(S, \xi) \leq M$, define $\varepsilon_N(\delta_{\mathrm{cert}}) := M\sqrt{2 \log(2K/\delta_{\mathrm{cert}})/N}$. We accept R_t only if $\widehat{q}_N(R_t, T_t) > C(R_t, T_t) + \tau + \varepsilon_N(\delta_{\mathrm{cert}})$ as summarized in Algorithm 1.

Let $\delta_{\mathrm{gr}} + \delta_{\mathrm{cert}} \leq \delta$ and define

$$\Gamma_t := \left(1 - \frac{1}{e}\right) \rho \Delta_{\mathcal{I}_k} + \eta_{\mathrm{gr}} + C(R_t, T_t) + \tau + 2\varepsilon_N(\delta_{\mathrm{cert}}).$$

Theorem 6.2 (Finite-sample feedback-safe greedy retraining)**.** *Suppose that $0 \leq f(S, \xi) \leq M$, C satisfies equation 5, and $\mathcal{I} = \mathcal{I}_k$. Then, with probability at least $1 - \delta_{\mathrm{cert}}$, all accepted redeployments in Algorithm 1 satisfy*

$$F(T_{t+1}) > F(T_t) + \tau.$$

Moreover, under Assumptions 5.1 and 5.2, with probability at least $1 - \delta_{\mathrm{gr}} - \delta_{\mathrm{cert}}$, if the algorithm terminates at some $t < K$, then

$$F(T_t) \geq \left(1 - \frac{1}{e}\right) \mathrm{OPT} - \Gamma_t.$$

Remark 6.1. The bound retains the $1 - 1/e$ greedy factor up to Γ_t, which captures performative feedback, finite-sample response error, and deployment certification. Tighter problem-dependent certificates directly reduce this terminal error.

7 EXPERIMENTS

We use a controlled decision-dependent weighted coverage problem to study two questions: how decision dependence affects the structural properties of the performative objective, and how different retraining procedures behave under performative feedback.

Decision-dependent weighted coverage. We consider a weighted coverage problem with item set $V = \{1, \dots, n\}$ and m topics. Given a topic-weight vector $\xi \in \Delta_m$, where Δ_m denotes the probability simplex, the utility of a selected set S is

$$f(S, \xi) = \sum_{j=1}^{m} \xi_j (1 - \prod_{e \in S} (1 - a_{ej})),$$

where $a_{ej} \in [0, 1]$ denotes the coverage strength of item e for topic j. For every fixed ξ, $f(\cdot, \xi)$ is normalized, monotone, and submodular. To model decision dependence, we generate the topic-weight vector as $\xi(S, \zeta) = \mathrm{softmax}\,(\theta_0 + \gamma B \mathbf{1}_S + \sigma\zeta)$ for $\zeta \sim \mathcal{N}(0, I_m)$, where θ_0 specifies the base topic preferences, B controls how the deployed set shifts these preferences, and σ controls exogenous noise. Thus, each deployed set S induces a distribution $\mathcal{D}(S)$ over ξ, and $\gamma \geq 0$ controls the strength of decision dependence, with $\gamma = 0$ recovering the fixed-distribution setting.

Experimental protocols. The structural diagnostic and the retraining comparison use separate configurations tailored to their respective purposes. Structural violations are evaluated under an adversarial-feedback configuration with $\sigma = 0.2$, while the retraining comparison uses aligned feedback with $\sigma = 0.5$. All methods within the retraining comparison are evaluated under the same protocol. Full settings and implementation details are provided in Appendix E.

Compared methods. We compare four methods. (i) STATIC GREEDY is a fixed-distribution baseline that optimizes once under the initial deployment distribution. (ii) NAIVE REPEATED GREEDY repeatedly recomputes a greedy solution under the distribution induced by the current deployment, without any stability or safety check. (iii) A POSTERIORI CERTIFIED GREEDY is the greedy instantiation of the certified repeated-best-response procedure in Section 6.1, which terminates according to the a posteriori stability certificate. (iv) FINITE-SAMPLE FEEDBACK-SAFE GREEDY is the greedy instantiation of Algorithm 1, which accepts an update only when its certified improvement dominates the possible loss due to distribution shift.

7.1 STRUCTURAL VIOLATIONS UNDER DECISION DEPENDENCE

To quantify the structural effect of decision dependence, we evaluate F under the adversarial-feedback configuration described above. We randomly test monotonicity and diminishing-returns inequalities of F over feasible sets and report the fraction of violated inequalities; formal definitions and sampling procedures are given in Appendix E.3.

Figure 1 shows that both violation rates are zero at $\gamma = 0$, consistent with the fixed-distribution setting. As decision dependence increases, structural violations emerge. The sampled diminishing-returns violation rate increases steadily with γ, reaching about 31% at $\gamma = 6$. The monotonicity violation rate is smaller and non-monotone, peaking at about 6.2% around $\gamma = 2$. Thus, both structural properties can be violated under performative feedback, although diminishing-returns violations are more pronounced in this experiment. As a robustness check, exhaustive enumeration on small-scale instances (Appendix E.4) finds no violations at $\gamma = 0$, while both types of violations occur in every tested instance for $\gamma > 0$.

7.2 RELIABILITY AND CONSERVATIVENESS OF CERTIFIED RETRAINING

We compare terminal quality, measured by the final performative objective and the frozen-improvement gap in equation 3, and pathwise reliability, measured by the F-decrease and nontrivial-cycle rates. These denote the fractions of runs with at least one decrease in F and with a revisited deployment after an intervening step, respectively; fixed points are excluded.

As shown in Table 2, naive repeated greedy and the a posteriori certified method achieve similar final objectives and near-zero terminal stability gaps at $\gamma = 2.0$, with small negative values reflecting

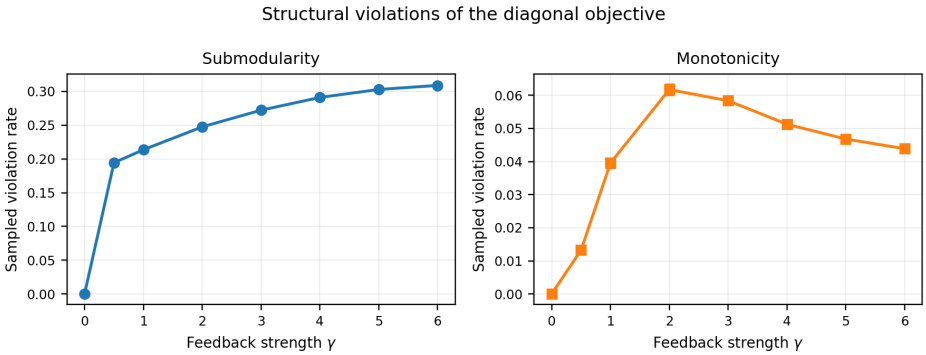

Figure 1: Sampled diminishing-returns and monotonicity violation rates of F under adversarial feedback as a function of feedback strength γ. Curves show means over 50 independent instances.

Table 2: Final deployment quality and reliability at $\gamma = 2.0$. Values are means over 50 instances with standard errors in parentheses; decrease and cycle columns report trajectory-level rates.

Method	F_{final}	Stability gap	F-decrease	Nontrivial cycle
Static Greedy	0.8495 (0.0030)	0.0082 (0.0014)	–	–
Naive Repeated Greedy	0.8824 (0.0058)	−0.0001 (0.0001)	0.42	0.10
A Posteriori Certified Greedy (Ours)	0.8833 (0.0057)	−0.0005 (0.0002)	0.20	0.00
Finite-Sample Feedback-Safe Greedy (Ours)	0.8495 (0.0030)	0.0082 (0.0014)	0.00	0.00

Monte Carlo error. Their trajectories differ: naive repeated greedy exhibits an F-decrease in 42% of runs and a nontrivial cycle in 10%, whereas the certified method has a 20% decrease rate and no observed nontrivial cycle, consistent with its terminal rather than pathwise guarantee.

The finite-sample feedback-safe method exhibits no F-decrease or nontrivial cycle. With the global certificate C_{global}, it rejects the first candidate in every run and therefore coincides with Static Greedy. Using the tighter problem-dependent certificate C_{loc} with a cold-start initialization, certified redeployments occur in all runs for $\gamma \leq 1$ and in 96%–100% of runs for $\gamma = 2$, with no observed safety violation. Full certificate and sample-size diagnostics are reported in Appendix E.4.4.

8 Conclusion

In this paper, we studied submodular decision making under decision-dependent distributions. We showed that performative feedback can destroy monotonicity and submodularity even when the utility is monotone and submodular for every fixed realization, and that repeated optimization of frozen objectives may decrease deployed performance or lead to cycles. By decoupling deployment from optimization, we recovered a monotone submodular frozen objective and introduced approximate performative stability as a self-consistent solution concept. Under distributional sensitivity, we established a stability-to-optimality guarantee with an explicit feedback-dependent degradation. We further developed retraining procedures for terminal stability and deployment-level safety, including a finite-sample feedback-safe greedy method with high-probability guarantees on accepted redeployments and terminal quality. Experiments on decision-dependent weighted coverage illustrate structural violations, retraining instability, and the safety–conservativeness tradeoff of certification.

Several limitations remain. The guarantees require valid sensitivity bounds or deployment certificates and can be conservative under strong feedback. Our analysis focuses mainly on cardinality-constrained monotone submodular maximization, while broader constraints and non-monotone objectives remain open. Extending the framework beyond the controlled weighted-coverage model is another important direction.

AI USE STATEMENT

Generative AI tools were used for literature discovery, research ideation, feedback on theoretical arguments and experimental design, and assistance with drafting and polishing the manuscript. All AI-assisted content, including mathematical claims, proofs, references, and experimental analyses, was independently verified and revised by the authors. The authors take full responsibility for the final content of the paper.

REPRODUCIBILITY STATEMENT

All assumptions and proofs are provided in the main text and appendix. The retraining procedures and certification rules are specified explicitly, and the appendix reports experimental settings, instance generation, evaluation metrics, and hyperparameters needed to reproduce the experiments.

REFERENCES

Tanveer Hossain Bhuiyan. *Models, Theoretical Properties, and Solution Approaches for Stochastic Programming with Endogenous Uncertainty*. PhD thesis, University of Tennessee, Knoxville, 2021. URL https://trace.tennessee.edu/utk_graddiss/6629/.

Jeff Bilmes. Submodularity in machine learning and artificial intelligence. *arXiv preprint arXiv:2202.00132*, 2022.

Dmitriy Drusvyatskiy and Lin Xiao. Stochastic optimization with decision-dependent distributions. *Mathematics of Operations Research*, 48(2):954–998, 2023.

Daniel Golovin and Andreas Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011.

Moritz Hardt and Celestine Mendler-Dünner. Performative prediction: Past and future. *Statistical Science*, 40(3), 2025.

Zhaolin He, Saverio Bolognani, Florian Dörfler, and Michael Muehlebach. Decision-dependent stochastic optimization: The role of distribution dynamics. *arXiv preprint arXiv:2503.07324*, 2025.

Lars Hellemo, Paul I. Barton, and Asgeir Tomasgard. Decision-dependent probabilities in stochastic programs with recourse. *Computational Management Science*, 15:369–395, 2018.

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.

Zhen Jia, Yifan Wang, Rui Dong, and Grani A. Hanasusanto. Distributionally robust performative optimization. In *Advances in Neural Information Processing Systems*, 2025.

David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 137–146, 2003.

Payam Khorsandi, Rahul Gupta, Maryam Mofakhami, Simon Lacoste-Julien, and Gauthier Gidel. Tight lower bounds and improved convergence in performative prediction. In *Advances in Neural Information Processing Systems*, 2025.

Andreas Krause and Daniel Golovin. Submodular function maximization. *Tractability*, 3(71-104): 3, 2014.

Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9(2), 2008.

Chen Liu, Zhaolin Wan, Haishan Ye, and John C. S. Lui. Stochastic non-smooth non-convex optimization with decision-dependent distributions. *arXiv preprint arXiv:2605.06549*, 2026.

Celestine Mendler-Dünner, Juan Perdomo, Tijana Zrnic, and Moritz Hardt. Stochastic optimization for performative prediction. In *Advances in Neural Information Processing Systems*, volume 33, pp. 4929–4939, 2020.

Maryam Mofakhami, Ioannis Mitliagkas, and Gauthier Gidel. Performative prediction with neural networks. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 11079–11093, 2023.

George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher. An analysis of approximations for maximizing submodular set functions—i. *Mathematical Programming*, 14:265–294, 1978.

Giacomo Pantuso. Solution of stochastic facility location problems with combinatorially many decision-dependent distributions. *Journal of Combinatorial Optimization*, 50:42, 2025.

Giacomo Pantuso and M. Hewitt. The l-shaped method for stochastic programs with decision-dependent uncertainty. *Mathematical Programming*, 2025. doi: 10.1007/s10107-025-02246-9.

Juan Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7599–7609, 2020.

Zhen Wang, Chen Liu, Thomas Parisini, Michael M. Zavlanos, and Karl H. Johansson. Constrained optimization with decision-dependent distributions. *IEEE Transactions on Automatic Control*, 70 (8):5176–5189, 2025. doi: 10.1109/TAC.2025.3540441.

Kevin Wood and Emiliano Dall'Anese. Stochastic saddle point problems with decision-dependent distributions. *SIAM Journal on Optimization*, 33(3):1943–1967, 2023. doi: 10.1137/22M1488077.

Xin Yu and Berk Basciftci. Distributionally robust optimization with multimodal decision-dependent ambiguity sets. *Mathematical Programming*, 2026. doi: 10.1007/s10107-026-02337-1.

Part I

Appendix

A Proofs for Propositions in Sec. 4

A.1 Proof of Proposition 4.1

Proof. Let $V = \{a, b\}$ and $\Omega = \{0, 1\}$. Define $f(S, \xi) = \xi|S|$. For every fixed $\xi \in \Omega$, the function $f(\cdot, \xi)$ is nonnegative, normalized, monotone, and submodular. We first show that F need not be submodular. Consider the decision-dependent distribution map

$$\mathcal{D}(\emptyset) = \mathcal{D}(\{a\}) = \mathcal{D}(\{b\}) = \delta_0, \qquad \mathcal{D}(\{a, b\}) = \delta_1,$$

where δ_i denotes the Dirac distribution at $i \in \{0, 1\}$. Then

$$F(\emptyset) = 0, \qquad F(\{a\}) = 0, \qquad F(\{b\}) = 0, \qquad F(\{a, b\}) = 2.$$

Hence $F(\{a\}) + F(\{b\}) = 0 < 2 = F(\{a, b\}) + F(\emptyset)$, which violates the submodularity inequality.

We next show that F need not be monotone. Using the same function $f(S, \xi) = \xi|S|$, define instead

$$\mathcal{D}(\{a\}) = \delta_1, \qquad \mathcal{D}(\{a, b\}) = \delta_0.$$

Then $F(\{a\}) = 1 > 0 = F(\{a, b\})$, which violates monotonicity. Therefore, pointwise normalized monotone submodularity of $f(\cdot, \xi)$ does not imply either submodularity or monotonicity of the induced objective F. □

A.2 Proof of Proposition 4.2

Proof. For fixed T, $G(\cdot; T)$ is the expectation of normalized, monotone, submodular functions $f(\cdot, \xi)$ under the fixed distribution $\mathcal{D}(T)$. These properties are preserved under nonnegative linear combinations, and hence under expectation. □

A.3 An example for Proposition 4.3

Example A.1 (A two-cycle of exact best responses)**.** Let $V = \{a, b\}$, $k = 1$, and $\Omega = [-1, 1]$. Define

$$f(S, x) = \left(\frac{3 + x}{2}\right) \mathbf{1}_{\{a \in S\}} + (1 - x)\mathbf{1}_{\{b \in S\}}.$$

For every fixed $x \in [-1, 1]$, $f(\cdot, x)$ is normalized, monotone, and modular, and hence submodular. Define $\mathcal{D}(\emptyset) = \delta_0, \mathcal{D}(\{a\}) = \delta_{-1}$, and $\mathcal{D}(\{b\}) = \delta_1$. Under the distribution induced by $\{a\}$, we have

$$f(\{a\}, -1) = 1, \qquad f(\{b\}, -1) = 2,$$

so the unique best response to $\{a\}$ is $\{b\}$. Conversely, under the distribution induced by $\{b\}$,

$$f(\{a\}, 1) = 2, \qquad f(\{b\}, 1) = 0,$$

so the unique best response to $\{b\}$ is $\{a\}$. Hence repeated exact best-response dynamics follow the two-cycle

$$\{a\} \longleftrightarrow \{b\}.$$

Moreover, we note that

$$F(\{a\}) = f(\{a\}, -1) = 1, \qquad F(\{b\}) = f(\{b\}, 1) = 0.$$

Thus, the exact best-response update

$$\{a\} \to \{b\}$$

strictly decreases the performative objective, even though $\{b\}$ is optimal under the frozen distribution induced by $\{a\}$. Therefore, the performative objective is not monotone along the best-response trajectory.

Finally, under $\mathcal{D}(\emptyset) = \delta_0$, we have

$$f(\{a\}, 0) = \frac{3}{2}, \qquad f(\{b\}, 0) = 1 > f(\emptyset, 0) = 0,$$

so $\emptyset$ is not a best response to itself. Since neither $\{a\}$ nor $\{b\}$ is a best response to itself, no feasible set is an exact best response to itself.

B Several proofs and examples in Sec. 5

B.1 A Sufficient Condition for Distributional Sensitivity

We give a common-latent-source sufficient condition for Assumption 5.1. Suppose that there exists a base random variable $\zeta \sim P_0$ on a measurable space $\mathcal{Z}$, together with a measurable map

$$M : 2^V \times \mathcal{Z} \to \Omega,$$

such that $M(S, \zeta) \sim \mathcal{D}(S)$ for every $S \in \mathcal{I}$. Assume further that, for every fixed ζ,

$$d_\Omega(M(S, \zeta), M(T, \zeta)) \le \lambda |S \triangle T|, \qquad \forall S, T \in \mathcal{I}.$$

Then Assumption 5.1 holds because the coupling induced by the common latent variable gives

$$W_1(\mathcal{D}(S), \mathcal{D}(T)) \le \mathbb{E}_{\zeta \sim P_0}\big[d_\Omega(M(S, \zeta), M(T, \zeta))\big] \le \lambda |S \triangle T|.$$

B.2 Stability-to-Optimality Proof in Sec. 5

First, we introduce the distributional sensitivity of values and marginal gains.

Lemma B.1 (Distributional sensitivity of values and marginal gains)**.** *Suppose that Assumptions 5.1 and 5.2 hold. Then, for any $A, S, T \in \mathcal{I}$, we have*

$$|G(A; S) - G(A; T)| \le \rho |S \triangle T|.$$

Moreover, for any $A \in \mathcal{I}$, $e \notin A$ with $A \cup \{e\} \in \mathcal{I}$, and any $S, T \in \mathcal{I}$, it holds

$$|\Delta_e G(A; S) - \Delta_e G(A; T)| \le 2\rho |S \triangle T|,$$

where $\Delta_e G(A; T) := G(A \cup \{e\}; T) - G(A; T)$.

Proof. We first prove the value sensitivity bound. For any fixed $A \in \mathcal{I}$, Assumption 5.2 implies that the function $\xi \mapsto f(A, \xi)$ is L_ξ-Lipschitz. By the Kantorovich–Rubinstein duality for the Wasserstein-1 distance,

$$\begin{aligned} |G(A; S) - G(A; T)| &= \big|\mathbb{E}_{\xi \sim \mathcal{D}(S)}[f(A, \xi)] - \mathbb{E}_{\xi \sim \mathcal{D}(T)}[f(A, \xi)]\big| \\ &\le L_\xi W_1(\mathcal{D}(S), \mathcal{D}(T)). \end{aligned}$$

Using Assumption 5.1, we obtain

$$|G(A; S) - G(A; T)| \le L_\xi \lambda |S \triangle T| = \rho |S \triangle T|.$$

We next prove the marginal-gain sensitivity bound. For any $A \in \mathcal{I}$ and $e \notin A$ with $A \cup \{e\} \in \mathcal{I}$, define

$$h_{A,e}(\xi) := f(A \cup \{e\}, \xi) - f(A, \xi).$$

Since both $\xi \mapsto f(A \cup \{e\}, \xi)$ and $\xi \mapsto f(A, \xi)$ are L_ξ-Lipschitz, $h_{A,e}$ is $2L_\xi$-Lipschitz:

$$\begin{aligned} |h_{A,e}(\xi) - h_{A,e}(\xi')| &\le |f(A \cup \{e\}, \xi) - f(A \cup \{e\}, \xi')| \\ &\quad + |f(A, \xi) - f(A, \xi')| \\ &\le 2L_\xi d_\Omega(\xi, \xi'). \end{aligned}$$

Applying Kantorovich–Rubinstein duality again gives

$$\begin{aligned} |\Delta_e G(A; S) - \Delta_e G(A; T)| &= \big|\mathbb{E}_{\xi \sim \mathcal{D}(S)}[h_{A,e}(\xi)] - \mathbb{E}_{\xi \sim \mathcal{D}(T)}[h_{A,e}(\xi)]\big| \\ &\le 2L_\xi W_1(\mathcal{D}(S), \mathcal{D}(T)) \\ &\le 2L_\xi \lambda |S \triangle T| = 2\rho |S \triangle T|. \end{aligned}$$

This completes the proof. □

Proof of Theorem 5.1

Proof. Let $S^* \in \arg\max_{R\in\mathcal{I}} F(R)$. Since $\widehat{S}$ is α-approximately performatively stable, there is

$$G(\widehat{S};\widehat{S}) \geq \alpha \max_{R\in\mathcal{I}} G(R;\widehat{S}) \geq \alpha G(S^*;\widehat{S}).$$

Using $F(\widehat{S}) = G(\widehat{S};\widehat{S})$, we obtain

$$F(\widehat{S}) \geq \alpha G(S^*;\widehat{S}).$$

By Lemma 5.1, applied with $A = S^*$, $S = \widehat{S}$, and $T = S^*$, we have

$$G(S^*;\widehat{S}) \geq G(S^*;S^*) - \rho|S^* \triangle \widehat{S}|.$$

Since

$$G(S^*;S^*) = F(S^*) = \max_{R\in\mathcal{I}} F(R),$$

and

$$|S^* \triangle \widehat{S}| \leq \Delta_{\mathcal{I}},$$

we conclude that

$$\begin{aligned} F(\widehat{S}) \geq \alpha G(S^*;\widehat{S}) &\geq \alpha G(S^*;S^*) - \alpha\rho|S^* \triangle \widehat{S}| \\ &\geq \alpha \max_{R\in\mathcal{I}} F(R) - \alpha\rho\Delta_{\mathcal{I}}. \end{aligned}$$

This completes the proof. □

C Algorithm analysis in Sec. 6

C.1 Sample-Based Greedy Responses

Under sample-only access, the frozen objective $G(\cdot;T)$ and its marginal gains are not available exactly. We therefore consider a sample-based greedy oracle based on empirical marginal gains.

Fix a deployment T. Starting from $S_0 = \emptyset$, at greedy round $\ell = 0, \ldots, k-1$, draw

$$\xi_\ell^1, \ldots, \xi_\ell^{N_{\mathrm{gr}}} \overset{\text{i.i.d.}}{\sim} \mathcal{D}(T),$$

independently of the samples used in previous rounds. For each $e \in V \setminus S_\ell$, define

$$\widehat{\Delta}_{e,N_{\mathrm{gr}}}(S_\ell;T) := \frac{1}{N_{\mathrm{gr}}} \sum_{i=1}^{N_{\mathrm{gr}}} \left(f(S_\ell \cup \{e\}, \xi_\ell^i) - f(S_\ell, \xi_\ell^i) \right),$$

and select

$$e_\ell \in \arg\max_{e\in V\setminus S_\ell} \widehat{\Delta}_{e,N_{\mathrm{gr}}}(S_\ell;T), \qquad S_{\ell+1} = S_\ell \cup \{e_\ell\}.$$

We denote the resulting response by $\widehat{\mathcal{A}}_{\mathrm{gr}}(T) := S_k$.

Lemma C.1 (Sample-based greedy response)**.** *Suppose that $0 \leq f(S,\xi) \leq M$ and that $f(\cdot,\xi)$ is normalized, monotone, and submodular for every ξ. Let $\mathcal{I} = \mathcal{I}_k$, and consider at most K adaptive calls to $\widehat{\mathcal{A}}_{\mathrm{gr}}$. For $\delta_{\mathrm{gr}} \in (0,1)$, define*

$$\varepsilon_{\mathrm{gr}} := M\sqrt{\frac{\log(2Kkn/\delta_{\mathrm{gr}})}{2N_{\mathrm{gr}}}}, \qquad \eta_{\mathrm{gr}} := 2k\varepsilon_{\mathrm{gr}}.$$

Then, with probability at least $1-\delta_{\mathrm{gr}}$, every returned response $R_t = \widehat{\mathcal{A}}_{\mathrm{gr}}(T_t)$ satisfies

$$G(R_t;T_t) \geq \left(1 - \frac{1}{e}\right) \max_{S\in\mathcal{I}_k} G(S;T_t) - \eta_{\mathrm{gr}}.$$

Equivalently, we have $R_t \in \mathrm{BR}_{1-1/e,\,\eta_{\mathrm{gr}}}(T_t)$.

Proof. Condition on the history before a greedy round, including the current deployment T and set S_ℓ. Since $f(\cdot, \xi)$ is monotone and $0 \le f(S, \xi) \le M$,

$$0 \le f(S_\ell \cup \{e\}, \xi) - f(S_\ell, \xi) \le M.$$

Hence, Hoeffding's inequality (Hoeffding, 1963) gives

$$\Pr\left(\left|\widehat{\Delta}_{e,N_{\mathrm{gr}}}(S_\ell; T) - \Delta_e G(S_\ell; T)\right| > \varepsilon_{\mathrm{gr}}\right) \le 2\exp\left(-\frac{2N_{\mathrm{gr}}\varepsilon_{\mathrm{gr}}^2}{M^2}\right).$$

Across at most K oracle calls, k greedy rounds per call, and at most n candidate marginal gains per round, a union bound implies that, with probability at least $1 - \delta_{\mathrm{gr}}$,

$$\left|\widehat{\Delta}_{e,N_{\mathrm{gr}}}(S_\ell; T) - \Delta_e G(S_\ell; T)\right| \le \varepsilon_{\mathrm{gr}} \tag{7}$$

for every empirical marginal gain evaluated by the algorithm.

On this event, let e_ℓ be the element selected at round ℓ. Using equation 7,

$$\begin{aligned}
\Delta_{e_\ell} G(S_\ell; T) &\ge \widehat{\Delta}_{e_\ell,N_{\mathrm{gr}}}(S_\ell; T) - \varepsilon_{\mathrm{gr}} \\
&\ge \max_{e \in V \setminus S_\ell} \widehat{\Delta}_{e,N_{\mathrm{gr}}}(S_\ell; T) - \varepsilon_{\mathrm{gr}} \\
&\ge \max_{e \in V \setminus S_\ell} \Delta_e G(S_\ell; T) - 2\varepsilon_{\mathrm{gr}}.
\end{aligned}$$

Let $S^\star \in \arg\max_{S \in \mathcal{I}_k} G(S; T)$. By monotonicity and submodularity, it holds

$$\max_{e \in V \setminus S_\ell} \Delta_e G(S_\ell; T) \ge \frac{G(S^\star; T) - G(S_\ell; T)}{k}.$$

Therefore, we obtain

$$G(S_{\ell+1}; T) - G(S_\ell; T) \ge \frac{G(S^\star; T) - G(S_\ell; T)}{k} - 2\varepsilon_{\mathrm{gr}}.$$

Iterating this recursion for k rounds and using $G(\emptyset; T) = 0$ yields

$$G(S_k; T) \ge \left(1 - \left(1 - \frac{1}{k}\right)^k\right) G(S^\star; T) - 2k\varepsilon_{\mathrm{gr}}.$$

Since the inequality

$$1 - \left(1 - \frac{1}{k}\right)^k \ge 1 - \frac{1}{e},$$

and $\eta_{\mathrm{gr}} = 2k\varepsilon_{\mathrm{gr}}$ hold, we obtain

$$G(S_k; T) \ge \left(1 - \frac{1}{e}\right) G(S^\star; T) - \eta_{\mathrm{gr}},$$

which proves the claim. □

Proof of Theorem 6.1

Proof. Since $\mathcal{A}_{\mathrm{gr}}$ is a $(1 - 1/e)$-approximate best-response oracle,

$$G(R_t; S_t) \ge \left(1 - \frac{1}{e}\right) \max_{S \in \mathcal{I}_k} G(S; S_t).$$

At termination, $\mathrm{Gap}_t \le \tau$, and hence

$$G(S_t; S_t) \ge G(R_t; S_t) - \tau.$$

Combining the two inequalities shows that S_t is $\big(1 - 1/e, \tau\big)$-approximately performatively stable.

Applying the (α, ε)-extension of Theorem 5.1 with $\alpha = 1 - 1/e$ and $\varepsilon = \tau$, together with $\Delta_{\mathcal{I}_k} \le 2k$ and $\rho = L_\xi \lambda$, yields

$$F(S_t) \ge \left(1 - \frac{1}{e}\right) \mathrm{OPT} - 2k \left(1 - \frac{1}{e}\right) L_\xi \lambda - \tau.$$

□

Proof of Theorem 6.2

Proof. For each iteration $t < K$, condition on the current deployment T_t and the candidate $R_t = \widehat{\mathcal{A}}_{\mathrm{gr}}(T_t)$. Since the certification batch is independent of the samples used to construct R_t,

$$\mathbb{E}[\widehat{q}_N(R_t, T_t) \mid T_t, R_t] = q(R_t, T_t).$$

Moreover, $0 \le f(S, \xi) \le M$ implies

$$f(R_t, \xi) - f(T_t, \xi) \in [-M, M].$$

Hence, by Hoeffding's inequality (Hoeffding, 1963), we obtain

$$\Pr(|\widehat{q}_N(R_t, T_t) - q(R_t, T_t)| > \varepsilon \,|\, T_t, R_t) \le 2\exp\left(-\frac{N\varepsilon^2}{2M^2}\right).$$

Taking $\varepsilon = \varepsilon_N(\delta_{\mathrm{cert}}) := M\sqrt{\frac{2\log(2K/\delta_{\mathrm{cert}})}{N}}$ gives a failure probability at most δ_{cert}/K. A union bound over at most K certification tests therefore implies that, with probability at least $1 - \delta_{\mathrm{cert}}$, we have

$$|\widehat{q}_N(R_t, T_t) - q(R_t, T_t)| \le \varepsilon_N(\delta_{\mathrm{cert}}) \tag{8}$$

simultaneously for all executed iterations $t < K$.

Consider first an accepted candidate R_t. By the acceptance rule,

$$\widehat{q}_N(R_t, T_t) > C(R_t, T_t) + \tau + \varepsilon_N(\delta_{\mathrm{cert}}).$$

On the event equation 8, this yields

$$q(R_t, T_t) > C(R_t, T_t) + \tau.$$

Using equation 5, we have

$$\begin{aligned} F(T_{t+1}) = F(R_t) &= G(R_t; R_t) \\ &\ge G(R_t; T_t) - C(R_t, T_t) \\ &= F(T_t) + q(R_t, T_t) - C(R_t, T_t) \\ &> F(T_t) + \tau. \end{aligned}$$

Thus, all accepted redeployments satisfy the claimed safety guarantee simultaneously with probability at least $1 - \delta_{\mathrm{cert}}$.

Now suppose that the algorithm terminates at iteration $t < K$ by rejecting R_t. The rejection condition gives

$$\widehat{q}_N(R_t, T_t) \le C(R_t, T_t) + \tau + \varepsilon_N(\delta_{\mathrm{cert}}).$$

Together with equation 8, we obtain

$$q(R_t, T_t) \le C(R_t, T_t) + \tau + 2\varepsilon_N(\delta_{\mathrm{cert}}),$$

and hence, we have

$$F(T_t) \ge G(R_t; T_t) - C(R_t, T_t) - \tau - 2\varepsilon_N(\delta_{\mathrm{cert}}). \tag{9}$$

By the sample-based greedy guarantee in Appendix C.1, with probability at least $1 - \delta_{\mathrm{gr}}$, we have

$$G(R_t; T_t) \ge \left(1 - \frac{1}{e}\right) \max_{A \in \mathcal{I}_k} G(A; T_t) - \eta_{\mathrm{gr}}.$$

Let $S^* \in \arg\max_{S \in \mathcal{I}_k} F(S)$. Under Assumptions 5.1 and 5.2, Lemma 5.1 gives

$$\begin{aligned} \max_{A \in \mathcal{I}_k} G(A; T_t) &\ge G(S^*; T_t) \\ &\ge G(S^*; S^*) - \rho|S^* \triangle T_t| \\ &\ge \mathrm{OPT} - \rho\Delta_{\mathcal{I}_k}. \end{aligned}$$

Combining this inequality with equation 9 yields

$$\begin{aligned} F(T_t) \ge\ & \left(1 - \frac{1}{e}\right)\mathrm{OPT} - \left(1 - \frac{1}{e}\right)\rho\Delta_{\mathcal{I}_k} \\ & - \eta_{\mathrm{gr}} - C(R_t, T_t) - \tau - 2\varepsilon_N(\delta_{\mathrm{cert}}) \\ =\ & \left(1 - \frac{1}{e}\right)\mathrm{OPT} - \Gamma_t. \end{aligned}$$

Finally, a union bound over the sample-based greedy and certification events gives probability at least $1 - \delta_{\mathrm{gr}} - \delta_{\mathrm{cert}}$, completing the proof. □

D Additional Technical Results

The global sensitivity bound in Section 5 does not exploit the structure of the weighted-coverage softmax model. The following result gives a sharper model-specific bound, which yields both a problem-dependent deployment certificate and an explicit feedback-sensitivity constant.

Proposition D.1 (Model-specific sensitivity for weighted coverage)**.** *Consider the decision-dependent weighted-coverage model*

$$f(S,\xi)=\sum_{j=1}^{m}\xi_j c_j(S),\qquad c_j(S)=1-\prod_{e\in S}(1-a_{ej}),$$

with $\xi(S,\zeta)=\mathrm{softmax}\big(\theta_0+\gamma B\mathbf{1}_S+\sigma\zeta\big)$*. Then, for any* $A,S,T\in\mathcal{I}$*,*

$$|G(A;S)-G(A;T)|\le\frac{\gamma}{4}\,\mathrm{osc}(c(A))\,\mathrm{osc}\big(B(\mathbf{1}_S-\mathbf{1}_T)\big),\tag{10}$$

where $\mathrm{osc}(v):=\max_j v_j-\min_j v_j$*. In particular, for any* $R,T\in\mathcal{I}$*, we obtain*

$$G(R;R)\ge G(R;T)-C_{\mathrm{loc}}(R,T),$$

where $C_{\mathrm{loc}}(R,T):=\frac{\gamma}{4}\,\mathrm{osc}(c(R))\,\mathrm{osc}\big(B(\mathbf{1}_T-\mathbf{1}_R)\big)$*. Hence,* C_{loc} *is a valid certificate for the feedback-safe acceptance rule. Moreover, defining*

$$\beta_B:=\max_{j,\ell,e}|B_{je}-B_{\ell e}|,\qquad \rho_{\mathrm{wc}}:=\frac{\gamma}{4}\beta_B,$$

we have

$$|G(A;S)-G(A;T)|\le\rho_{\mathrm{wc}}|S\triangle T|.$$

Proof. Fix $A,S,T\in\mathcal{I}$ and a realization ζ. Let

$$z(t)=\theta_0+\gamma B\big((1-t)\mathbf{1}_T+t\mathbf{1}_S\big)+\sigma\zeta,\qquad t\in[0,1],$$

and let $p(t)=\mathrm{softmax}(z(t))$. Define

$$h(t):=c(A)^\top p(t).$$

Using the Jacobian of the softmax map, we obtain

$$h'(t)=\gamma\,\mathrm{Cov}_{p(t)}\big(c(A),B(\mathbf{1}_S-\mathbf{1}_T)\big).$$

For any probability vector p and vectors u,v, the Cauchy–Schwarz inequality together with the standard range bound on the variance gives

$$|\,\mathrm{Cov}_p(u,v)|\le\sqrt{\mathrm{Var}_p(u)\,\mathrm{Var}_p(v)}\le\frac{1}{4}\,\mathrm{osc}(u)\,\mathrm{osc}(v).$$

Therefore, we obtain

$$|h'(t)|\le\frac{\gamma}{4}\,\mathrm{osc}(c(A))\,\mathrm{osc}\big(B(\mathbf{1}_S-\mathbf{1}_T)\big).$$

Integrating over $t\in[0,1]$ and then taking expectation over ζ gives equation 10. Furthermore, setting $A=R$ and $S=R$ yields the stated local certificate. Finally, since $c_j(A)\in[0,1]$, we have

$$\mathrm{osc}(c(A))\le 1.$$

Moreover, note that $\mathrm{osc}\big(B(\mathbf{1}_S-\mathbf{1}_T)\big)\le\beta_B|S\triangle T|$, which gives

$$|G(A;S)-G(A;T)|\le\frac{\gamma}{4}\beta_B|S\triangle T|.$$

□

E Additional Experimental Details

E.1 Decision-Dependent Weighted Coverage

We provide additional implementation details for the decision-dependent weighted-coverage benchmark introduced in Section 7. The generation of the coverage matrix A, base preference vector θ_0, feedback matrix B, and all experimental parameters is specified below.

E.2 Experimental Protocol

General settings. For each feedback strength γ, we generate 50 independent instances and report means and standard errors across instance seeds. Unless otherwise stated, all methods use the same instances, initialization, and evaluation samples. Expectations in F and G are estimated by Monte Carlo sampling over the latent variable ζ. Common latent samples are used for within-instance comparisons, while evaluation samples are independent of the certification samples used by the feedback-safe method.

The parameters in Table 3 correspond to the main retraining comparison under aligned feedback. The structural diagnostic and adversarial-feedback stress test use the configurations specified in their respective subsections.

Table 3: Parameters for the main retraining comparison.

Parameter	Value
Number of items n	20
Number of topics m	8
Cardinality budget k	5
Noise level σ	0.5
Feedback strengths γ	$\{0, 0.5, 1.0, 2.0, 3.0\}$
Monte Carlo samples N_{MC}	2000
Independent instances	50
Certificate tolerance τ	10^{-4}
Numerical tolerance tol	10^{-8}
Maximum retraining iterations $T_{\max}$	20
Feedback-safe confidence level δ	0.05
Feedback-safe certification samples N	2000
Global certificate parameter ρ	1
Utility bound M	1

Instance generation. For each instance, item centers c_e are sampled independently and uniformly from the m topics. Before clipping, the coverage matrix $A = (a_{ej})$ is generated as

$$a_{ej} = a_{\max} \exp(-\lambda_{\mathrm{cov}} d_{\mathrm{circ}}(c_e, j)) + \eta_{ej},$$

where d_{circ} denotes circular topic distance. We use

$$a_{\max} = 0.85, \qquad \lambda_{\mathrm{cov}} = 1.75, \qquad \eta_{ej} \sim \mathcal{N}(0, 0.02^2),$$

and clip a_{ej} to $[0, 1]$. The initial feedback matrix satisfies $\widetilde{B}_{je} \sim \mathcal{N}(0, 0.05^2)$. We then impose one of the following feedback structures, where $U(a, b)$ denotes a uniform random variable on $[a, b]$.

Feedback mechanisms. We consider two feedback modes. In the aligned setting, deployment shifts future topic weights toward the selected item's primary coverage region; in the adversarial setting, it shifts weight away from that region.

For aligned feedback, an item e with center topic c_e modifies $\widetilde{B}_{\cdot e}$ as

$$\begin{aligned}
\widetilde{B}_{c_e,e} &\leftarrow \widetilde{B}_{c_e,e} + U(0.75, 1.15), \\
\widetilde{B}_{c_e+1,e} &\leftarrow \widetilde{B}_{c_e+1,e} - U(0.55, 0.95), \\
\widetilde{B}_{c_e+2,e} &\leftarrow \widetilde{B}_{c_e+2,e} + U(0.20, 0.45), \\
\widetilde{B}_{c_e-1,e} &\leftarrow \widetilde{B}_{c_e-1,e} - U(0.10, 0.35).
\end{aligned}$$

This setting is used in the main retraining comparison.

For adversarial feedback, define $o_e = \big(c_e + \max\{2, \lfloor m/2 \rfloor\}\big) \bmod m$, and modify

$$\begin{aligned}
\widetilde{B}_{c_e,e} &\leftarrow \widetilde{B}_{c_e,e} - U(0.95, 1.35),\\
\widetilde{B}_{c_e+1,e} &\leftarrow \widetilde{B}_{c_e+1,e} - U(0.35, 0.70),\\
\widetilde{B}_{c_e-1,e} &\leftarrow \widetilde{B}_{c_e-1,e} - U(0.20, 0.50),\\
\widetilde{B}_{o_e,e} &\leftarrow \widetilde{B}_{o_e,e} + U(0.95, 1.35),\\
\widetilde{B}_{o_e+1,e} &\leftarrow \widetilde{B}_{o_e+1,e} + U(0.35, 0.70).
\end{aligned}$$

This setting is used in the structural diagnostic and adversarial-feedback stress test. Topic indices are interpreted modulo m.

For both modes, each column is normalized as

$$B_{\cdot e} = \frac{\widetilde{B}_{\cdot e}}{\max\{\|\widetilde{B}_{\cdot e}\|_2, 10^{-12}\}},$$

and $\theta_0 \sim \mathcal{N}(0, 0.15^2 I_m)$.

Initialization and stopping. In the main comparison, all retraining methods start from $S_0 = \textsc{Greedy}(G(\cdot; \varnothing))$. Naive repeated greedy immediately redeploys each greedy response, while the a posteriori method stops when

$$G(R_t; S_t) - G(S_t; S_t) \le \tau$$

or after $T_{\max}$ iterations. The feedback-safe method deploys a candidate R_t only if

$$\widehat{q}_N(R_t, T_t) > C(R_t, T_t) + \tau + \epsilon_N(\delta).$$

For the main comparison, we use

$$C_{\text{global}}(R, T) = \rho |R \triangle T|, \qquad \epsilon_N(\delta) = M\sqrt{\frac{2\log(2T_{\max}/\delta)}{N}}.$$

A rejected candidate terminates the procedure and is not counted as a deployed update.

Supplemental feedback-safe diagnostics. Appendix E.4.4 compares the global and local certificates under static-greedy initialization with $N = 10{,}000$, and also considers a cold start with $T_0 = \varnothing$, the local certificate, and a pair-dependent Hoeffding radius for $N \in \{500, 2000, 10000\}$. All other experimental settings follow the main protocol.

E.3 Evaluation Metrics

We evaluate structural violations, deployment reliability, and terminal quality. Throughout, $\mathrm{tol} > 0$ denotes a numerical tolerance used to avoid counting negligible Monte Carlo or floating-point errors as genuine violations.

Structural violations. The submodularity violation rate is the fraction of sampled feasible triples (A, B, e), with $A \subseteq B$ and $e \notin B$, satisfying

$$\Delta_e F(A) < \Delta_e F(B) - \mathrm{tol}.$$

The monotonicity violation rate is the fraction of sampled feasible pairs $A \subseteq B$ satisfying

$$F(A) > F(B) + \mathrm{tol}.$$

Deployment reliability. The F-decrease rate is the fraction of runs containing at least one deployed transition satisfying

$$F(S_{t+1}) < F(S_t) - \mathrm{tol}.$$

The nontrivial-cycle rate is the fraction of runs that revisit a previous deployment after at least one intervening step; fixed points are excluded. For the adversarial-feedback stress test, we additionally report

$$\mathrm{MaxDrop} = \max_t \big(F(S_t) - F(S_{t+1})\big)_+.$$

Terminal quality. We report the final performative value $F(S_{\text{final}})$ and the greedy stability gap

$$\text{Gap}_{\text{gr}} = G(R_{\text{gr}}; S_{\text{final}}) - G(S_{\text{final}}; S_{\text{final}}),$$

where $R_{\text{gr}} = \text{Greedy}\big(G(\cdot; S_{\text{final}})\big)$. For feedback-safe retraining, we additionally report the number of accepted redeployments.

E.4 Additional Results

E.4.1 Exhaustive Structural Check

To complement the sampled diagnostic in the main text, we perform an exhaustive check on smaller instances with $n = 10$, $m = 6$, and $k = 3$, using $N_{\text{MC}} = 20{,}000$ samples to evaluate F. For each instance, we enumerate all feasible submodularity and monotonicity conditions over 10 independent instances for each γ. Table 4 shows no violations at $\gamma = 0$, whereas every tested instance exhibits both types of violations for $\gamma > 0$. The exhaustive results therefore support the sampled structural diagnostic reported in the main text.

Table 4: Structural violation rates over all feasible inequalities. All feasible monotonicity and diminishing-returns inequalities are enumerated, with F evaluated using $N_{\text{MC}} = 20{,}000$ Monte Carlo samples.

γ	Submod. exists	Mono. exists	Submod. violation rate	Mono. violation rate
0	0/10	0/10	0.000	0.000
0.5	10/10	10/10	0.325	0.037
1	10/10	10/10	0.324	0.101
2	10/10	10/10	0.373	0.146
3	10/10	10/10	0.381	0.111
4	10/10	10/10	0.399	0.088
5	10/10	10/10	0.415	0.077
6	10/10	10/10	0.421	0.074

E.4.2 Adversarial-Feedback Stress Test

Figure 2 shows substantially stronger instability under adversarial feedback: already at $\gamma = 0.5$, about 96% of runs exhibit an F-decrease, with frequent nontrivial cycles and larger one-step performance drops.

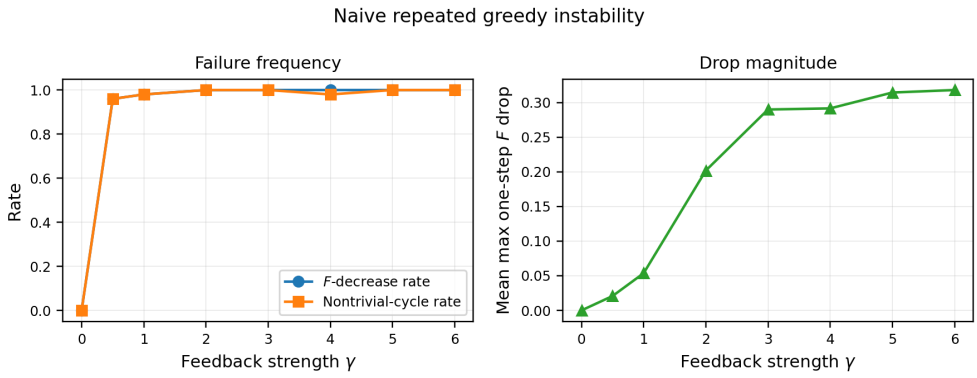

Figure 2: Instability of naive repeated greedy under the adversarial-feedback stress test. Left: fractions of runs exhibiting at least one F-decrease or a nontrivial cycle, with fixed points excluded. Right: mean maximum one-step performance drop $\max_t (F(S_t) - F(S_{t+1}))_+$.

E.4.3 Full Reliability Results Across Feedback Strengths

Table 5 reports deployment reliability across the tested feedback strengths. Under the global certificate, the feedback-safe method rejects every candidate and therefore has zero deployed decreases and cycles. Naive repeated greedy frequently exhibits F-decreases and occasional cycles, whereas the a posteriori method has lower observed decrease rates and no observed nontrivial cycles.

Table 5: Deployment reliability across feedback strengths. The F-decrease rate is the fraction of runs with at least one deployed decrease in F, and the nontrivial-cycle rate is the fraction of runs containing a cycle of period greater than one.

γ	Method	F-decrease rate	Nontrivial cycle rate
0.5	Naive Repeated Greedy	0.42	0.08
0.5	A Posteriori Certified Greedy	0.14	0.00
0.5	Finite-Sample Feedback-Safe Greedy	0.00	0.00
1.0	Naive Repeated Greedy	0.36	0.06
1.0	A Posteriori Certified Greedy	0.16	0.00
1.0	Finite-Sample Feedback-Safe Greedy	0.00	0.00
2.0	Naive Repeated Greedy	0.42	0.10
2.0	A Posteriori Certified Greedy	0.20	0.00
2.0	Finite-Sample Feedback-Safe Greedy	0.00	0.00
3.0	Naive Repeated Greedy	0.36	0.14
3.0	A Posteriori Certified Greedy	0.20	0.00
3.0	Finite-Sample Feedback-Safe Greedy	0.00	0.00

E.4.4 Additional Diagnostics for Feedback-Safe Retraining

The main experiments show that finite-sample feedback-safe retraining can be conservative under static-greedy initialization. We therefore compare the global and local certificates and then consider a cold-start setting in which larger improvements are available.

Global versus local certificates. For the weighted-coverage softmax model, we use the problem-dependent local certificate established in Proposition D.1,

$$C_{\mathrm{loc}}(R,T) = \frac{\gamma}{4}\operatorname{osc}(c(R))\operatorname{osc}(B(\mathbf{1}_T - \mathbf{1}_R)).$$

It upper-bounds $G(R;T) - G(R;R)$ and can be substantially tighter than the global certificate

$$C_{\mathrm{global}}(R,T) = \rho|R \triangle T|.$$

Table 6 compares the two certificates using the same instances, candidate responses, and certification samples. The local certificate is substantially tighter; for example, at $\gamma = 2$, it decreases from 3.24 to 0.229. However, the estimated frozen improvement remains small relative to the certification threshold, and neither certificate accepts a redeployment. This indicates that the conservativeness observed in the main experiment reflects both the certificate size and the near-stability of the static-greedy initialization.

Cold-start safe retraining. The preceding diagnostic suggests that the lack of accepted redeployments is partly due to the near-stable static-greedy initialization. We therefore consider a cold-start setting with $T_0 = \varnothing$, using the local certificate C_{loc} and a pair-dependent Hoeffding radius based on the range of $f(R,\xi) - f(T,\xi)$. The instance generation, greedy response oracle, and evaluation protocol remain unchanged.

Table 7 shows that certified redeployments are possible when starting farther from a frozen best response. For $\gamma \leq 1$, one update is accepted in every run, while for $\gamma = 2$ the mean number of accepted updates is 0.96 at $N = 500$ and 1.00 for $N \geq 2000$. At $\gamma = 3$, accepted updates become rare. No empirical safety violation is observed in the tested runs.

Table 6: Global and local certificate diagnostics under static-greedy initialization with $N = 10{,}000$. Values are averages over 50 instances using the same candidate responses and certification samples. Here $\epsilon_N = 0.0366$ for all settings.

γ	$\widehat{q}_N$	C_{global}	C_{loc}
0.5	0.00005	2.04	0.0186
1.0	0.00153	2.32	0.0381
2.0	0.00816	3.24	0.2289
3.0	0.02654	3.96	0.9163

Table 7: Cold-start feedback-safe retraining with the local certificate and pair-dependent confidence radius. "Accepted updates" is the mean number of accepted redeployments per run, and "Zero-update rate" is the fraction of runs with no accepted redeployment. Results are averaged over 50 instances.

γ	N	ϵ_N	Accepted updates	Zero-update rate	Safety violation rate
0.0	500	0.0506	1.00	0.00	0.00
0.0	2000	0.0253	1.00	0.00	0.00
0.0	10000	0.0113	1.00	0.00	0.00
0.5	500	0.0534	1.00	0.00	0.00
0.5	2000	0.0267	1.00	0.00	0.00
0.5	10000	0.0119	1.00	0.00	0.00
1.0	500	0.0523	1.00	0.00	0.00
1.0	2000	0.0261	1.00	0.00	0.00
1.0	10000	0.0117	1.00	0.00	0.00
2.0	500	0.0553	0.96	0.04	0.00
2.0	2000	0.0277	1.00	0.00	0.00
2.0	10000	0.0124	1.00	0.00	0.00
3.0	500	0.0512	0.02	0.98	0.00
3.0	2000	0.0256	0.04	0.96	0.00
3.0	10000	0.0114	0.04	0.96	0.00