STRONGER, TRACTABLE NOTIONS OF REGRET FOR ONLINE SUBMODULAR MAXIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Online Submodular Maximization (OSM) is typically evaluated via static regret, which only guarantees competitiveness with the best *fixed* action in hindsight. While fundamental, this notion fails to capture more refined notions of rational behavior that arise naturally in combinatorial decision spaces. We revisit OSM through the lens of Φ-regret, a family of stronger regret notions defined with respect to sets of action transformations. We identify rich yet tractable classes of transformations that encode natural operations on sets, including permutations and conditional trades, and show how to achieve sublinear Φ-regret in polynomial time. Our algorithms combine classical machinery from online convex optimization with rounding schemes from combinatorial optimization.

1 INTRODUCTION

In online learning, optimization problems with submodular objectives arise naturally in sequential decision-making tasks where actions correspond to selecting subsets of items and rewards exhibit diminishing returns. This structure captures a wide range of practical settings in which the marginal benefit of selecting an additional element decreases due to redundancy, overlap, or saturation effects. Representative applications include influence maximization, facility location, caching networks, and recommender systems (Kempe et al., 2003; Karimi et al., 2017; Krause & Golovin, 2014; Ioannidis & Yeh, 2016).

In this setting, an online submodular maximization (OSM) algorithm receives a sequence of (potentially adversarially chosen) monotone submodular objective functions $f_1, \dots, f_T : 2^{[n]} \to \mathbb{R}_{\geq 0}$ over T rounds. At each round, the algorithm must pick a set $S_t \subseteq [n]$ before the corresponding objective f_t is revealed; after that, f_t is revealed and the algorithm receives reward $f_t(S_t)$. We assume that the decision maker is constrained to only picking sets S_t with cardinality $|S_t| \leq k, k \in \mathbb{N}$. The standard performance metric for an online submodular maximization algorithm is its α-*static regret*

$$\mathcal{R}_\alpha \coloneqq \alpha \max_{|S| \leq k} \sum_{t=1}^{T} f_t(S) - \sum_{t=1}^{T} f_t(S_t). \tag{1}$$

Under this metric, a good submodular maximization algorithm would strive to keep its α-static regret sublinear with high probability. Typically, the parameter α is taken to be $1 - 1/e \approx 0.63$, which is known to be unimprovable in general due to complexity barriers (Feige, 1998).

Sublinear static regret only guarantees that, in hindsight, the learner would not have been better off replacing its entire sequence of decisions by a single fixed set. That said, it would seem natural to demand of a good learning algorithm other conditions, such as

i) never regretting a *trade* from element a to element b, i.e., replacing a by b whenever a is selected and b is not; and

ii) never regretting *swapping* a and b, i.e., exchanging them whenever exactly one of them is selected.

More generally, one could demand that a good learning algorithm learn to never regret certain specified deterministic transformations that could have been applied in hindsight to the decision produced. These more sophisticated notions of regret typically go under the name of Φ-regret (Gordon et al., 2008). In this context, Φ specifies a set of transformations with respect to which the learner is

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.

evaluated. The corresponding relevant notion of regret is therefore (α, Φ)-regret

$$\mathcal{R}_\alpha^\Phi := \alpha \max_{\phi \in \Phi} \sum_{t=1}^{T} f_t(\phi(S_t)) - \sum_{t=1}^{T} f_t(S_t). \tag{2}$$

In recent years, in the adjacent literature on learning in games, these more sophisticated regret notions have received increased attention, due in part to their connections with equilibrium concepts (Dagan et al., 2024; Peng & Rubinstein, 2024; Farina & Pipis, 2023). In the context of online submodular maximization, however, their algorithmic tractability has remained unclear because one must simultaneously handle optimization over a continuous relaxation and the transfer of guarantees back to integral set decisions.

In this paper, we show for the first time that it is possible to achieve sublinear (α, Φ)-regret in polynomial time for rich classes of transformations Φ that include both examples (i) and (ii). Our approach works by optimizing over a continuous relaxation of the problem and then translating the resulting guarantees back to feasible set decisions.

2 Related Work and Techniques

This paper combines modern advances in the theory of Φ-regret minimization with specific transformation classes for online submodular maximization. As we will show, the structural properties of these classes allow us to sidestep hardness barriers that arise in more general Φ-regret domains.

Φ-Regret in Convex Games. A standard starting point for Φ-regret minimization is the reduction of Gordon et al. (2008), which identifies two core algorithmic ingredients: no-static-regret learning over the transformation class Φ and fixed-point computation for maps in the relevant convex hull. This reduction underlies efficient algorithms for several equilibrium concepts in games, including correlated and extensive-form correlated equilibria (Foster & Vohra, 1997; Hart & Mas-Colell, 2000; Celli et al., 2020; Farina & Pipis, 2023). More broadly, the literature on Φ-regret in convex and game-theoretic settings is extensive, and has clarified the close connection between transformation-based regret notions and equilibrium computation.

At the same time, Gordon's reduction quickly encounters computational barriers outside highly structured settings. Even when fixed points are tractable, the induced static-regret problem over the transformation class may already be computationally intractable; and for nonlinear transformation classes, the fixed-point step itself typically becomes hard as well. Indeed, already for general linear endomorphisms of a convex compact domain, the fixed-point ingredient can be handled efficiently while the induced static-regret problem remains intractable. The difficulty only becomes more pronounced once Φ contains nonlinear maps.

Recent work has developed ways around these barriers, either by relaxing the fixed-point requirement via in-expectation notions (Zhang et al., 2024) or by using semi-separation-style constructions that allow outputs outside the target transformation class (Daskalakis et al., 2025; Dann et al., 2024; Arunachaleswaran et al., 2025; Zhang et al., 2025). These advances substantially expand the range of domains in which efficient Φ-regret minimization is possible, but they also highlight how delicate tractability becomes once the transformation class is nonlinear or combinatorially rich. In particular, they suggest that efficient Φ-regret minimization in new domains is unlikely to follow from black-box reductions alone, and instead requires exploiting problem-specific structure.

Our contribution is to identify natural transformation classes for online submodular maximization for which this stronger notion of regret remains tractable. In particular, we focus on affine and quadratic multilinear transformations that encode operations such as permutations and conditional trades. We show that these transformations are non-expansive, which yields efficient approximate fixed-point computation, and that the induced static-regret problems over the corresponding relaxed transformation classes can be solved directly in polynomial time without semi-separation or ellipsoid-based machinery. Thus, rather than treating Φ-regret minimization in OSM as an instance of a general-purpose framework, we exploit the specific algebraic structure of the transformation classes that arise in this domain.

Online Submodular Maximization (OSM). OSM has received significant attention in recent years as a principled framework for sequential decision making with combinatorial structure. Prior work has

studied OSM under a variety of feedback models (full-information, bandit), constraints (cardinality, general matroid constraints), and reward models (stochastic, adversarial) (Streeter et al., 2009; Harvey et al., 2020; Niazadeh et al., 2021; Matsuoka et al., 2021; Si Salem et al., 2024). There has also been recent work on the online maximization of continuous DR-submodular functions (Chen et al., 2018; Zhang et al., 2019; 2022).

Most closely related to our work is Si Salem et al. (2024), which introduced the rounding-augmented online convex optimization (RAOCO) framework for OSM and showed how to obtain sublinear regret guarantees by combining online optimization over a fractional relaxation with randomized rounding. Our approach builds on this relax-and-round template and combines it with the auxiliary-function construction of Zhang et al. (2022). The main new ingredient is that we use this framework not only to compete with the best fixed set in hindsight, but to handle richer transformation-based comparators while preserving polynomial-time tractability.

3 Preliminaries and Background

Notation. We use bold letters to denote vectors (e.g., $\boldsymbol{x}, \boldsymbol{y}$) and non-bold letters to denote scalars. The i-th coordinate of a vector $\boldsymbol{x}$ is denoted by x_i. We use $\mathbf{0}$ and $\mathbf{1}$ to denote the all-zeros and all-ones vectors, respectively. For a positive integer n, we write $[n] = \{1, \dots, n\}$. We identify a set $S \subseteq [n]$ with its indicator vector $\boldsymbol{x} = \mathbf{1}_S \in \{0,1\}^n$. We denote by $\mathcal{Y} = \mathrm{conv}(\mathcal{X})$ the convex hull of $\mathcal{X}$.

3.1 Online Convex Optimization

In Online Convex Optimization (OCO), the learner repeatedly chooses points $\boldsymbol{x}_t$ from a compact convex set $\mathcal{X}$ and then observes a concave reward function $g_t : \mathcal{X} \to \mathbb{R}$. The standard performance measure is the static regret

$$\mathcal{R}_{\mathcal{X}} := \max_{\boldsymbol{x} \in \mathcal{X}} \sum_{t=1}^{T} g_t(\boldsymbol{x}) - \sum_{t=1}^{T} g_t(\boldsymbol{x}_t). \tag{3}$$

More generally, given a set of transformations Φ from $\mathcal{X}$ to itself, the corresponding Φ-regret is

$$\mathcal{R}_{\mathcal{X}}^{\Phi} := \max_{\phi \in \Phi} \sum_{t=1}^{T} g_t(\phi(\boldsymbol{x}_t)) - \sum_{t=1}^{T} g_t(\boldsymbol{x}_t). \tag{4}$$

Static regret is the special case of Φ-regret in which Φ is the class of constant transformations on $\mathcal{X}$. More precisely, define $\mathcal{C} := \{\phi_{\boldsymbol{x}} : \mathcal{X} \to \mathcal{X} \mid \phi_{\boldsymbol{x}}(\boldsymbol{y}) = \boldsymbol{x} \text{ for all } \boldsymbol{y} \in \mathcal{X},\ \boldsymbol{x} \in \mathcal{X}\}$. Then, $\mathcal{R}_{\mathcal{X}}^{\mathcal{C}} = \mathcal{R}_{\mathcal{X}}$.

3.2 Online Submodular Maximization

Let $\mathcal{V} = [n]$ be a ground set. A set function $f : 2^{\mathcal{V}} \to \mathbb{R}_{\geq 0}$ is *submodular* if $f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B)$ for all $A \subseteq B$ and $v \in \mathcal{V} \setminus B$, and *monotone* if $f(A) \leq f(B)$ for all $A \subseteq B$. Furthermore, we assume that f is normalized, i.e., $f(\emptyset) = 0$.

In Online Submodular Maximization (OSM), at each round $t = 1, \dots, T$, the learner selects a feasible set $\boldsymbol{x}_t \in \mathcal{X}$, then observes a monotone submodular reward function $f_t : \{0,1\}^n \to \mathbb{R}_{\geq 0}$ and receives reward $f_t(\boldsymbol{x}_t)$. Throughout the paper we focus on cardinality constraints,

$$\mathcal{X} = \left\{ \boldsymbol{x} \in \{0,1\}^n : \sum_{i=1}^{n} x_i \leq k \right\}, \qquad \mathcal{Y} = \mathrm{conv}(\mathcal{X}) = \left\{ \boldsymbol{y} \in [0,1]^n : \sum_{i=1}^{n} y_i \leq k \right\}.$$

Since offline monotone submodular maximization is approximable in polynomial time only up to the factor $1 - 1/e$ (Feige, 1998), online algorithms are evaluated via α-regret with $\alpha = 1 - 1/e$:

$$\mathcal{R}_{\alpha,\mathcal{X}} := \alpha \max_{\boldsymbol{x} \in \mathcal{X}} \sum_{t=1}^{T} f_t(\boldsymbol{x}) - \sum_{t=1}^{T} f_t(\boldsymbol{x}_t). \tag{5}$$

More generally, for a set of transformations Φ on $\mathcal{X}$, we define the (α, Φ)-regret as

$$\mathcal{R}^{\Phi}_{\alpha,\mathcal{X}} := \alpha \max_{\phi\in\Phi} \sum_{t=1}^{T} f_t(\phi(\boldsymbol{x}_t)) - \sum_{t=1}^{T} f_t(\boldsymbol{x}_t). \tag{6}$$

Multilinear extension. Given a monotone submodular function $f : \{0,1\}^n \to \mathbb{R}_{\geq 0}$, let $R(\boldsymbol{y})$ denote the random set obtained by including each element $i \in [n]$ independently with probability y_i. The *multilinear extension* of f (Calinescu et al., 2011) is defined as

$$F(\boldsymbol{y}) := \mathbb{E}[f(R(\boldsymbol{y}))],\ \boldsymbol{y} \in [0,1]^n. \tag{7}$$

In particular, $F(\boldsymbol{x}) = f(\boldsymbol{x})$ for every integral point $\boldsymbol{x} \in \{0,1\}^n$. To optimize submodular rewards online, we work over the fractional domain $\mathcal{Y}$ via the multilinear extension, and then round back to an integral set.

Relax-and-round template. A standard approach to online submodular maximization is the following *Relax-and-Round* template (Algorithm 1). At each round t, the learner first computes a fractional point $\boldsymbol{y}_t \in \mathcal{Y}$ using an online convex optimization (OCO) policy $\mathcal{P}_{\mathcal{Y}}$, and then rounds it to an integral action $\boldsymbol{x}_t = \Xi(\boldsymbol{y}_t) \in \mathcal{X}$, where $\Xi : \mathcal{Y} \to \mathcal{X}$ is a rounding procedure. Thus the algorithm operates over the fractional domain $\mathcal{Y}$, where optimization is tractable, and subsequently translates the resulting decisions back to feasible sets of cardinality at most k.

The key property required from the rounding step is that it approximately preserves the fractional objective in expectation. In particular, we use a rounding procedure based on the monotone contention resolution scheme of Kashaev & Santiago (2023). For monotone submodular objectives, this rounding satisfies

$$\mathbb{E}[f_t(\boldsymbol{x}_t)] \geq \beta\, F_t(\boldsymbol{y}_t), \qquad \beta = 1 - \frac{e^{-k}k^k}{k!}. \tag{8}$$

Thus, if the fractional sequence $(\boldsymbol{y}_t)_{t=1}^T$ enjoys a suitable regret guarantee with respect to the multilinear values F_t, then the rounded integral sequence inherits a corresponding guarantee up to the factor β. For completeness, the rounding procedure Ξ is described in Appendix A.

OSM via Auxiliary-Function OCO. For general monotone submodular functions, however, the multilinear extension F_t is not concave, so it cannot be optimized directly using standard OCO methods. Following Zhang et al. (2022), we therefore work with the auxiliary function

$$\Psi_t(\boldsymbol{y}) := \int_0^1 \frac{e^{z-1}}{z} F_t(z\boldsymbol{y})\, dz.$$

Although F_t is not concave, Ψ_t satisfies a "concave-like" inequality: for all $\boldsymbol{y}, \boldsymbol{u} \in \mathcal{Y}$,

$$\langle \nabla\Psi_t(\boldsymbol{y}), \boldsymbol{u} - \boldsymbol{y}\rangle \geq \left(1 - \frac{1}{e}\right) F_t(\boldsymbol{u}) - F_t(\boldsymbol{y}).$$

This inequality plays the role of concavity in the analysis. In particular, if one runs an OCO method on the linearized rewards $\ell_t(\boldsymbol{u}) := \langle \nabla\Psi_t(\boldsymbol{y}_t), \boldsymbol{u}\rangle$, then sublinear static regret on the sequence $(\ell_t)_{t=1}^T$ yields a $(1 - 1/e)$-approximate regret guarantee with respect to the multilinear objectives F_t.

Combining this auxiliary-function reduction with the Relax-and-Round template above yields a no-$\beta(1 - 1/e)$-regret algorithm for online submodular maximization. More details on the auxiliary-function construction are given in Appendix B, and the resulting high-probability regret bound for the relax-and-round scheme is proved in Appendix B.3.

4 Stronger Regret Notions for Online Submodular Maximization

Standard analyses of Online Submodular Maximization (OSM) focus on *static regret*, which benchmarks performance against the best fixed feasible set in hindsight. While fundamental, this notion is quite weak. For one, it does not prevent the learner from making systematic structural mistakes even when simple transformations of past actions would have improved reward. This motivates the study of Φ-regret, which compares the learner's sequence against transformed versions of its own actions.

In this section, we introduce transformation classes for OSM, using permutations and conditional trades as motivating examples. We then define two tractable classes that include these transformations and constant maps, strengthening static regret by capturing structured deviations of the learner's actions. Appendix E provides explicit examples separating the resulting regret notions.

Algorithm 1: Relax-and-Round

Input: OCO policy $\mathcal{P}_{\mathcal{Y}}$; rounding procedure $\Xi : \mathcal{Y} \to \mathcal{X}$

for $t = 1, 2, \ldots, T$ **do**
- $\boldsymbol{y}_t \leftarrow \mathcal{P}_{\mathcal{Y}}(\ell_1, \ldots, \ell_{t-1})$
- $\boldsymbol{x}_t \leftarrow \Xi(\boldsymbol{y}_t)$
- play $\boldsymbol{x}_t$ and observe f_t
- construct the linear reward $\ell_t(\boldsymbol{u}) \leftarrow \langle \nabla \Psi_t(\boldsymbol{y}_t), \boldsymbol{u} \rangle$

4.1 Item Swaps and Permutations

We begin with the simplest nontrivial transformations: *item swaps* and, more generally, *permutations*.

Item Swap Transformations. Fix two distinct items $i \neq j$. An *item swap transformation* exchanges the mass assigned to coordinates i and j while leaving all other coordinates unchanged. Formally, for $\boldsymbol{x} \in \{0,1\}^n$, define

$$\phi_{i \leftrightarrow j}(\boldsymbol{x})_r = \begin{cases} x_j, & r = i, \\ x_i, & r = j, \\ x_r, & r \notin \{i, j\}. \end{cases}$$

Permutation Transformations. More generally, any *permutation* π of $[n]$ induces a transformation

$$\phi_\pi(\boldsymbol{x})_i = x_{\pi(i)}.$$

Minimizing regret against permutations ensures that the learner cannot benefit from relabeling items in hindsight. Equivalently, these transformations have the form $\phi_A(\boldsymbol{x}) = A\boldsymbol{x}$, where $A \in \{0,1\}^{n \times n}$ is a permutation matrix, i.e., each row and each column contains exactly one 1. We refer to regret against this collection as *permutation regret*. Permutation maps are examples of the affine transformations defined below.

4.2 Conditional Item Trades

Permutation maps relabel items unconditionally. A learner may also wish to replace item i with item j only when i is selected and j is not. Such conditional deviations can be expressed by quadratic multilinear transformations.

Conditional Trade Transformations. As a motivating example, we introduce a simple quadratic transformation that captures conditional item trades. Fix distinct indices $i \neq j$. For any $\boldsymbol{x} \in \{0,1\}^n$, define

$$\phi_{i \to j}(\boldsymbol{x})_r = \begin{cases} x_i x_j, & r = i, \\ 1 - (1 - x_i)(1 - x_j), & r = j, \\ x_r, & r \notin \{i, j\}. \end{cases}$$

This transformation swaps x_i and x_j if and only if $x_i = 1$ and $x_j = 0$; otherwise, both coordinates remain unchanged.

This transformation expresses a conditional preference: the learner may regret selecting item i in the absence of item j, but not the reverse. The quadratic terms encode the dependence on the joint presence of items i and j. We refer to regret against these transformations as *conditional trade regret*.

4.3 Affine and Quadratic Transformation Classes

For $\ell \in \{1, 2\}$, let $\widetilde{\Phi}_\ell$ denote the class of maps $\phi : [0,1]^n \to [0,1]^n$ whose coordinates are multilinear polynomials of degree at most ℓ, allowing a constant term. Writing $b := \phi(\mathbf{0})$, we require $b \in \mathcal{Y}$, so that $\boldsymbol{y} \mapsto b$ is a feasible constant transformation. We further require that there exists $\rho \in [0,1]$ such that:

- **Range:** $\phi(\boldsymbol{y}) \in [0,1]^n, \ \forall \boldsymbol{y} \in [0,1]^n$.
- **Monotonicity:** $\boldsymbol{z} \leq \boldsymbol{y}$ implies $\phi(\boldsymbol{z}) \leq \phi(\boldsymbol{y}), \ \forall \boldsymbol{z}, \boldsymbol{y} \in [0,1]^n$.

- **Affine total mass:** $\mathbf{1}^\top \phi(\boldsymbol{y}) = \mathbf{1}^\top b + \rho\, \mathbf{1}^\top \boldsymbol{y},\ \forall \boldsymbol{y} \in [0,1]^n$.
- **Budget:** $\mathbf{1}^\top b + \rho k \le k$.

The range condition bounds each output coordinate between zero and one. The affine total mass and budget conditions also ensure that the total output mass does not exceed k: for every $\boldsymbol{y} \in \mathcal{Y}$,

$$\mathbf{1}^\top \phi(\boldsymbol{y}) \le \mathbf{1}^\top b + \rho k \le k.$$

Consequently, $\phi(\mathcal{Y}) \subseteq \mathcal{Y}$. Although $b \in \mathcal{Y}$ also follows from the range and budget conditions, we state it explicitly to emphasize that the constant term is itself a feasible fractional action.

The degree-one class $\widetilde{\Phi}_1$ consists of affine transformations $\phi(\boldsymbol{y}) = b + A\boldsymbol{y}$. The degree-two class $\widetilde{\Phi}_2$ additionally permits quadratic multilinear terms. Both classes contain every constant map $\phi(\boldsymbol{x}) = c$ with $c \in \mathcal{Y}$, realized by $b = c$ and $\rho = 0$. The polynomial extensions of permutations and conditional trades to the cube belong to $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$, respectively, with $b = 0$ and $\rho = 1$. Sections 5 and 6.1 provide polynomial-size coefficient descriptions and efficient fixed-point routines, respectively.

Integral Transformations. To define regret for the integral decision problem, we retain the transformations that preserve integrality. For $\ell \in \{1, 2\}$, let

$$\Phi_\ell := \left\{ \phi \in \widetilde{\Phi}_\ell \;\middle|\; \phi(\{0,1\}^n) \subseteq \{0,1\}^n \right\}.$$

Every $\phi \in \Phi_\ell$ maps feasible actions to feasible actions: for $\boldsymbol{x} \in \mathcal{X}$, the output $\phi(\boldsymbol{x})$ is integral by definition and belongs to $\mathcal{Y}$ because $\phi \in \widetilde{\Phi}_\ell$. Hence $\phi(\boldsymbol{x}) \in \mathcal{X}$.

We refer to regret against Φ_1 on $\mathcal{X}$ as *affine regret* and regret against Φ_2 as *quadratic regret*. Both classes contain every constant map to a feasible action, so both regret notions include static regret.

Remark. It is natural to ask whether the framework developed here can be extended to higher-degree multilinear transformations. At present, the main obstacle is that, beyond the quadratic case, we do not have a polynomial-size explicit description of the admissible coefficient sets. Hence any such extension may require oracle-based or semi-separation techniques, which we leave to future work.

5 General Formulation and Description of Transformation Sets

We represent the transformations in $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$ through their constant, linear, and quadratic coefficients. Define the column vector of pairwise monomials $\mathbf{q}(\boldsymbol{x}) := (x_p x_q)_{1 \le p < q \le n}$ and the feature maps $\mathbf{z}_1(\boldsymbol{x}) = (1, \boldsymbol{x})$ and $\mathbf{z}_2(\boldsymbol{x}) = (1, \boldsymbol{x}, \mathbf{q}(\boldsymbol{x}))$. Their dimensions are $p_1 = 1 + n$ and $p_2 = 1 + n + \binom{n}{2}$, respectively. For $\ell \in \{1, 2\}$, every transformation in $\widetilde{\Phi}_\ell$ can be represented as

$$\phi_\Theta(\boldsymbol{x}) = \Theta \mathbf{z}_\ell(\boldsymbol{x}), \qquad \Theta \in \mathbb{R}^{n \times p_\ell}. \tag{9}$$

We write $\Theta = [\mathbf{b} \;\; A]$ for affine maps and $\Theta = [\mathbf{b} \;\; A \;\; Q]$ for quadratic maps, where $\mathbf{b} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, and $Q \in \mathbb{R}^{n \times \binom{n}{2}}$. Thus $\phi_\Theta(\boldsymbol{x}) = \mathbf{b} + A\boldsymbol{x}$ in the affine case and $\phi_\Theta(\boldsymbol{x}) = \mathbf{b} + A\boldsymbol{x} + Q\mathbf{q}(\boldsymbol{x})$ in the quadratic case.

Let $\mathcal{P}_\ell := \left\{ \Theta \in \mathbb{R}^{n \times p_\ell} \;\middle|\; \phi_\Theta \in \widetilde{\Phi}_\ell \right\}$ denote the admissible coefficient domain. In Appendix C.1, we give a polyhedral description of $\mathcal{P}_1$ with $\mathcal{O}(n^2)$ variables and constraints. In Appendix C.2, we describe $\mathcal{P}_2$ using $\mathcal{O}(n^3)$ variables and linear constraints.

6 Efficient Algorithms for Affine and Quadratic Transformations

We now instantiate the Relax-and-Round template (Algorithm 1) for the affine and quadratic transformation classes. Our construction specifies the fractional policy $\mathcal{P}_\mathcal{Y}$ using an online learner over the coefficient domain $\mathcal{P}_\ell$ ($\ell \in \{1, 2\}$) and an approximate fixed-point routine. Algorithm 2 gives the complete procedure, including rounding and the coefficient update.

The policy maintains a coefficient matrix $\Theta_t \in \mathcal{P}_\ell$, representing the transformation $\phi_{\Theta_t}(\boldsymbol{y}) = \Theta_t \mathbf{z}_\ell(\boldsymbol{y})$. At round t, it computes an approximate fixed point $\boldsymbol{y}_t \in \mathcal{Y}$ of this transformation.

This is the fractional action returned by $\mathcal{P}_{\mathcal{Y}}$ in Algorithm 1. The algorithm then rounds $\boldsymbol{y}_t$ to a feasible integral action $\boldsymbol{x}_t = \Xi(\boldsymbol{y}_t)$ and plays $\boldsymbol{x}_t$. After observing f_t, it constructs the surrogate reward $\ell_t(\boldsymbol{u}) = \langle \nabla\Psi_t(\boldsymbol{y}_t), \boldsymbol{u}\rangle$ and updates Θ_t by Online Gradient Ascent (OGA) on the induced transformation reward $r_t(\Theta) = \ell_t(\Theta \mathbf{z}_\ell(\boldsymbol{y}_t))$.

This policy construction follows the reduction of Φ-regret to static regret of Gordon et al. (2008). For our transformation classes, both required operations are efficient: anchored iteration computes approximate fixed points, and the polyhedral coefficient descriptions support projected OGA. Section 6.1 establishes the fixed-point guarantee, Section 6.2 describes the update and its surrogate-regret guarantee, and Section 6.3 transfers this guarantee to the integral actions.

6.1 Non-Expansiveness and Fixed Points

Monotonicity and affine total mass imply ℓ_1-non-expansiveness. The range and budget conditions additionally ensure that the transformations map $\mathcal{Y}$ into itself, allowing us to compute approximate fixed points by anchored iteration.

Lemma 6.1. *Let $\ell \in \{1, 2\}$ and $\phi \in \widetilde{\Phi}_\ell$, with associated parameter $\rho \in [0, 1]$. Then*

$$\|\phi(\boldsymbol{x}) - \phi(\boldsymbol{y})\|_1 \le \rho\|\boldsymbol{x} - \boldsymbol{y}\|_1 \le \|\boldsymbol{x} - \boldsymbol{y}\|_1, \ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{Y}.$$

The proof is given in Appendix D.1. It uses the fact that the Jacobian of ϕ is entrywise nonnegative and each of its columns sums to ρ.

Anchored contraction and approximate fixed points. Let $\psi \in \widetilde{\Phi}_\ell$ and fix an anchor $\boldsymbol{x}_0 \in \mathcal{Y}$. For $\mu \in (0, 1)$, define

$$T_\mu(\boldsymbol{x}) := (1 - \mu)\boldsymbol{x}_0 + \mu\psi(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathcal{Y}.$$

Since $\mathcal{Y}$ is convex and ψ maps $\mathcal{Y}$ into itself, T_μ also maps $\mathcal{Y}$ into itself. Moreover, non-expansiveness of ψ gives $\|T_\mu(\boldsymbol{x}) - T_\mu(\boldsymbol{y})\|_1 \le \mu\|\boldsymbol{x} - \boldsymbol{y}\|_1$. Thus T_μ is a contraction. The following lemma bounds the number of iterations needed to obtain an approximate fixed point of ψ.

Lemma 6.2. *Let $\psi : \mathcal{Y} \to \mathcal{Y}$ be non-expansive in ℓ_1, and suppose that $\mathcal{Y}$ has ℓ_1-diameter at most D. Set $\epsilon = 1/\sqrt{T}$ and $\mu = 2D\sqrt{T}/(2D\sqrt{T} + 1)$. Starting from $\boldsymbol{x}^0 = \boldsymbol{x}_0 \in \mathcal{Y}$, consider the iteration*

$$\boldsymbol{x}^{s+1} = T_\mu(\boldsymbol{x}^s) = (1 - \mu)\boldsymbol{x}_0 + \mu\psi(\boldsymbol{x}^s).$$

After $s = \mathcal{O}(D\sqrt{T}\log(2 + D\sqrt{T}))$ iterations, the iterate satisfies $\|\psi(\boldsymbol{x}^s) - \boldsymbol{x}^s\|_1 \le \frac{1}{\sqrt{T}}$.

The proof is in Appendix D.2. For affine transformations $\psi(\boldsymbol{y}) = \mathbf{b} + A\boldsymbol{y}$, a fixed point can alternatively be computed by solving $(I - A)\boldsymbol{y} = \mathbf{b}$ subject to $\boldsymbol{y} \in \mathcal{Y}$. Algorithm 2 uses the anchored routine for both transformation classes. Lemma 6.2 guarantees termination of its inner loop with an ϵ-approximate fixed point. We use $D = 2k$, an upper bound on the ℓ_1-diameter of $\mathcal{Y}$.

6.2 Online Optimization over Transformation Classes

Having established an efficient fixed-point routine, we now describe the online learner over transformations. As described in Section 5, we identify each transformation $\phi_\Theta \in \widetilde{\Phi}_\ell$ with its coefficient matrix $\Theta \in \mathcal{P}_\ell$ and run Online Gradient Ascent (OGA) over this domain.

At round t, the learner maintains $\Theta_t \in \mathcal{P}_\ell$ and computes an ϵ-fixed point $\boldsymbol{y}_t$ of ϕ_{Θ_t}. This is the fractional action returned by the policy $\mathcal{P}_{\mathcal{Y}}$ in Algorithm 1. After playing the rounded action $\boldsymbol{x}_t = \Xi(\boldsymbol{y}_t)$ and observing f_t, let $\mathbf{g}_t = \nabla\Psi_t(\boldsymbol{y}_t)$ and define

$$r_t(\Theta) = \ell_t(\phi_\Theta(\boldsymbol{y}_t)) = \langle \mathbf{g}_t, \Theta\mathbf{z}_\ell(\boldsymbol{y}_t)\rangle.$$

This reward is linear in Θ, with gradient $\nabla_\Theta r_t(\Theta) = \mathbf{g}_t\mathbf{z}_\ell(\boldsymbol{y}_t)^\top$. The update is therefore

$$\Theta_{t+1} = \Pi_{\mathcal{P}_\ell}\left(\Theta_t + \eta\, \mathbf{g}_t\mathbf{z}_\ell(\boldsymbol{y}_t)^\top\right),$$

where $\Pi_{\mathcal{P}_\ell}$ denotes Euclidean projection in the Frobenius norm. Projection reduces to a convex quadratic program using the coefficient descriptions in Appendices C.1 and C.2.

Since $\mathcal{P}_\ell$ is compact and convex, standard OGA bounds yield $\mathcal{O}(\sqrt{T})$ static regret with an appropriate choice of the learning rate. Together with the approximate fixed-point routine, OGA yields $\mathcal{O}(\sqrt{T})$ $\widetilde{\Phi}_\ell$-regret on the linear rewards ℓ_t. The ϵ-approximate fixed points contribute an additional $O(\epsilon T)$ regret term, which is $O(\sqrt{T})$ for $\epsilon = 1/\sqrt{T}$.

Algorithm 2: Affine and Quadratic Φ-Regret Minimization

Input: Degree $\ell \in \{1, 2\}$; horizon T; step size $\eta > 0$; anchor $\boldsymbol{x}_0 \in \mathcal{Y}$; rounding procedure $\Xi : \mathcal{Y} \to \mathcal{X}$.

Initialize $\Theta_1 \in \mathcal{P}_\ell$;
Set $\epsilon = T^{-1/2}$, $D = 2k$, and $\mu = 2D/(2D + \epsilon)$;
for $t = 1, \ldots, T$ **do**
  $\boldsymbol{y} \leftarrow \boldsymbol{x}_0$;
  while $\|\Theta_t \mathbf{z}_\ell(\boldsymbol{y}) - \boldsymbol{y}\|_1 > \epsilon$ **do**
    $\boldsymbol{y} \leftarrow (1 - \mu)\boldsymbol{x}_0 + \mu\Theta_t \mathbf{z}_\ell(\boldsymbol{y})$;
  $\boldsymbol{y}_t \leftarrow \boldsymbol{y}$;
  Sample $\boldsymbol{x}_t = \Xi(\boldsymbol{y}_t)$ and play $\boldsymbol{x}_t$;
  Observe f_t and set $\mathbf{g}_t = \nabla\Psi_t(\boldsymbol{y}_t)$;
  $\Theta_{t+1} \leftarrow \Pi_{\mathcal{P}_\ell}\big(\Theta_t + \eta\, \mathbf{g}_t \mathbf{z}_\ell(\boldsymbol{y}_t)^\top\big)$;

6.3 No Φ-regret for Online Submodular Maximization

Having constructed the fractional policy, we now establish the integral regret guarantee for Algorithm 2. This algorithm instantiates the Relax-and-Round template (Algorithm 1) with the policy described above. The fractional policy operates on the surrogate linear rewards $\ell_t(\boldsymbol{u}) = \langle \nabla\Psi_t(\boldsymbol{y}_t), \boldsymbol{u}\rangle$. The rounding procedure $\Xi : \mathcal{Y} \to \mathcal{X}$ uses the contention resolution scheme of Kashaev & Santiago (2023), described in Appendix A, to produce the integral action $\boldsymbol{x}_t = \Xi(\boldsymbol{y}_t)$. Algorithm 2 summarizes the complete procedure, combining the fractional policy constructed above with CRS rounding.

Since $\widetilde{\Phi}_\ell$ contains all constant maps to $\mathcal{Y}$, the same learner competes with both static comparators and structured transformations. Combining this fractional policy with rounding gives the following main guarantee against the integral class Φ_ℓ.

Theorem 6.3. *Fix $\ell \in \{1, 2\}$ and let $d_\ell = np_\ell$. Assume that $0 \le f_t(S) \le M$ for all $S \subseteq [n]$ and $t \in [T]$. Assume that P_Y achieves regret $R_{1,Y}^{\widetilde{\Phi}_\ell}$ on the surrogate linear rewards $\ell_t(u) = \langle \nabla\Psi_t(y_t), u\rangle$. Let $\beta = 1 - \frac{e^{-k}k^k}{k!}$ and $\alpha = \beta(1 - 1/e)$. Then, for any $p \in (0, 1)$, the Relax-and-Round algorithm (Algorithm 1) using the CRS rounding from Appendix A satisfies, with probability at least $1 - p$,*

$$R_{\alpha,X}^{\Phi_\ell} \le \beta R_{1,Y}^{\widetilde{\Phi}_\ell} + O\left(M\sqrt{T\left(d_\ell + \log\frac{1}{p}\right)}\right).$$

The proof is in Appendix C.4. Applying the theorem to the fractional policy constructed in Section 6.2, whose surrogate regret is $\mathcal{O}(\sqrt{T})$, shows that our complete algorithm, Algorithm 2, achieves sublinear (α, Φ_ℓ)-regret with high probability.

7 Experiments

We conclude with a small synthetic experiment illustrating the distinction between static regret and affine regret, and the performance of learners operating over our transformation classes. We make the code used for the experiments publicly available in an anonymized GitHub repository.

Setup. We consider a weighted-coverage instance with $n = 6$ items and cardinality budget $k = 2$. The items are partitioned into three pairs, and each item covers one distinct universe element. At each round, the elements corresponding to one pair have weight $1/2$, while all other elements have weight zero. The resulting objective is monotone submodular.

We evaluate all learners on integral actions obtained using the same contention resolution scheme (CRS, see Appendix A). We consider two reward schedules: a *static* schedule in which the same pair is active at every round, and a *cyclical* schedule in which the active pair rotates across time. We use $T = 10000$ rounds and report regret per round, averaged over 64 independent rounding runs. All reported regret values use $\alpha = 1$. Full instance construction, hyperparameters, and implementation details are given in Appendix G.

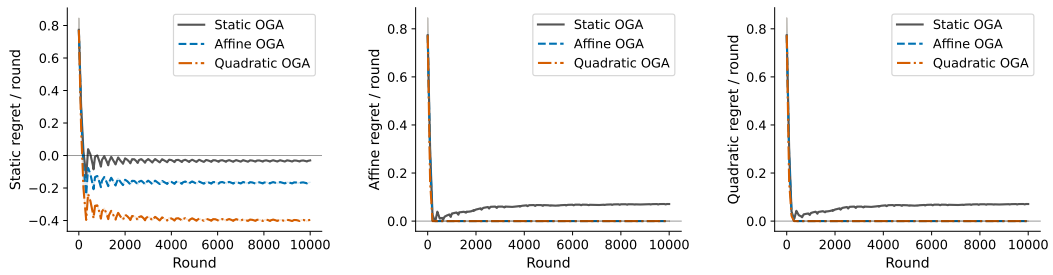


Figure 1: Static (left), affine (middle), and quadratic (right) regret per round on the cyclical weighted-coverage instance with $n = 6$, $k = 2$, and $T = 10000$. Curves show mean cumulative regret per round over 64 independent runs, with shaded regions indicating 95% confidence intervals.

Algorithms. We compare three algorithms. The first is standard OGA on the fractional action domain. The second is affine OGA (Algorithm 2, $\ell = 1$), which runs OGA over the coefficient domain $\mathcal{P}_1$ and obtains its fractional action as an approximate fixed point of the resulting transformation. The third is quadratic OGA (Algorithm 2, $\ell = 2$), which operates analogously over $\mathcal{P}_2$. All methods are initialized from the same uniform fractional point, use their respective theoretical learning rates $\eta = D/(G\sqrt{T})$, and round their fractional actions using the CRS described in Appendix A.

Results. Figure 1 shows the behavior on the cyclical instance. Static OGA achieves negative average static regret but incurs positive average affine and quadratic regret. In contrast, affine OGA and quadratic OGA achieve zero affine and quadratic regret at the final horizon in every sampled run, while also achieving negative static regret. Negative static regret means that the learner outperforms every fixed feasible action in hindsight. On the static instance, which is included in Appendix G, all three methods achieve low average static regret.

Thus good performance against static comparators need not imply good performance against affine deviations, whereas both transformation learners perform well under all three metrics on this instance. The comparison does not establish a separation between affine and quadratic regret.

8 Conclusions

We established the first sublinear Φ-regret guarantees for Online Submodular Maximization, extending its theoretical foundations beyond static regret. We identified broad and expressive transformation classes Φ, including affine and quadratic mappings, for which these stronger guarantees are achievable in polynomial time by combining learning over action transformations with appropriate rounding.

Limitations and Future Directions. Our results are restricted to the full-information setting and to cardinality constraints. Natural next steps are to extend the framework to bandit feedback and to more general matroid constraints. It would also be interesting to understand whether similar ideas can be extended to broader transformation classes, including higher-degree multilinear mappings, and to other combinatorial domains.

AI use statement

We used ChatGPT to improve the presentation throughout the paper and simplify the arguments used in Appendices E and F. Moreover, we used it for implementing the experiments in Section 7. Finally, we used it to proofread our work and pinpoint inconsistencies and minor errors. The authors are responsible for all content of the paper.

Ethics statement

The paper presents theoretical and experimental results and does not involve human subjects or sensitive personal data.

Reproducibility statement

All theoretical statements are accompanied by formal proofs. For the experiments we describe dataset construction, algorithmic details, and hyperparameter choices in Appendix G. We make the code used for the experiments publicly available in an anonymized GitHub repository.

References

Eshwar Ram Arunachaleswaran, Natalie Collina, Yishay Mansour, Mehryar Mohri, Jon Schneider, and Balasubramanian Sivan. Swap regret and correlated equilibria beyond normal-form games. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pp. 130–157, 2025.

Gruia Calinescu, Chandra Chekuri, Martin Pal, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011.

Andrea Celli, Alberto Marchesi, Gabriele Farina, and Nicola Gatti. No-regret learning dynamics for extensive-form correlated equilibrium. *Advances in Neural Information Processing Systems*, 33: 7722–7732, 2020.

Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pp. 783–792, 2011.

Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization. In *International Conference on Artificial Intelligence and Statistics*, pp. 1896–1905. PMLR, 2018.

Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. From external to swap regret 2.0: An efficient reduction for large action spaces. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pp. 1216–1222, 2024.

Christoph Dann, Yishay Mansour, Mehryar Mohri, Jon Schneider, and Balasubramanian Sivan. Rate-preserving reductions for blackwell approachability. *arXiv preprint arXiv:2406.07585*, 2024.

Constantinos Daskalakis, Gabriele Farina, Maxwell Fishelson, Charilaos Pipis, and Jon Schneider. Efficient learning and computation of linear correlated equilibrium in general convex games. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pp. 542–553, 2025.

Gabriele Farina and Charilaos Pipis. Polynomial-time linear-swap regret minimization in imperfect-information sequential games. *Advances in Neural Information Processing Systems*, 36:28255–28280, 2023.

Uriel Feige. A threshold of ln n for approximating set cover. *Journal of the ACM (JACM)*, 45(4): 634–652, 1998.

Dean P Foster and Rakesh V Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(589):40–55, 1997.

Geoffrey J Gordon, Amy Greenwald, and Casey Marks. No-regret learning in convex games. In *Proceedings of the 25th international conference on Machine learning*, pp. 360–367, 2008.

Theodore E Harris. A lower bound for the critical probability in a certain percolation process. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 56, pp. 13–20. Cambridge University Press, 1960.

Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

Nicholas Harvey, Christopher Liaw, and Tasuku Soma. Improved algorithms for online submodular maximization via first-order regret bounds. *Advances in Neural Information Processing Systems*, 33:123–133, 2020.

Stratis Ioannidis and Edmund Yeh. Adaptive caching networks with optimality guarantees. *ACM SIGMETRICS Performance Evaluation Review*, 44(1):113–124, 2016.

Mohammad Karimi, Mario Lucic, Hamed Hassani, and Andreas Krause. Stochastic submodular maximization: The case of coverage functions. *Advances in Neural Information Processing Systems*, 30, 2017.

Danish Kashaev and Richard Santiago. A simple optimal contention resolution scheme for uniform matroids. *Theoretical Computer Science*, 940:81–96, 2023.

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.

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

Tatsuya Matsuoka, Shinji Ito, and Naoto Ohsaka. Tracking regret bounds for online submodular optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 3421–3429. PMLR, 2021.

Rad Niazadeh, Negin Golrezaei, Joshua R Wang, Fransisca Susan, and Ashwinkumar Badanidiyuru. Online learning via offline greedy algorithms: Applications in market design and optimization. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 737–738, 2021.

Binghui Peng and Aviad Rubinstein. Fast swap regret minimization and applications to approximate correlated equilibria. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pp. 1223–1234, 2024.

Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003.

Tareq Si Salem, Gözde Özcan, Iasonas Nikolaou, Evimaria Terzi, and Stratis Ioannidis. Online submodular maximization via online convex optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 15038–15046, 2024.

Matthew Streeter, Daniel Golovin, and Andreas Krause. Online learning of assignments. *Advances in neural information processing systems*, 22, 2009.

Brian Zhang, Ioannis Anagnostides, Gabriele Farina, and Tuomas Sandholm. Efficient ϕ-regret minimization with low-degree swap deviations in extensive-form games. *Advances in Neural Information Processing Systems*, 37:125192–125230, 2024.

Brian Hu Zhang, Ioannis Anagnostides, Emanuel Tewolde, Ratip Emin Berker, Gabriele Farina, Vincent Conitzer, and Tuomas Sandholm. Learning and computation of Φ-equilibria at the frontier of tractability. *arXiv preprint arXiv:2502.18582*, 2025.

Mingrui Zhang, Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization: From full-information to bandit feedback. *Advances in Neural Information Processing Systems*, 32, 2019.

Qixin Zhang, Zengde Deng, Zaiyi Chen, Haoyuan Hu, and Yu Yang. Stochastic continuous submodular maximization: Boosting via non-oblivious function. In *International Conference on Machine Learning*, pp. 26116–26134. PMLR, 2022.

A CONTENTION RESOLUTION SCHEMES

Given some feasibility constraints (typically matroid constraints) over a ground set $\mathcal{V}$, a *Contention Resolution Scheme* (CRS) (Chekuri et al., 2011) is a randomized procedure that takes as input a random set $R \subseteq \mathcal{V}$, which is not necessarily feasible, and trims it down to a feasible set $\pi(R) \subseteq R$.

Most work on CRSs focuses on the following standard rounding pipeline. Given a fractional point $\boldsymbol{y} \in \mathcal{Y} \subseteq [0,1]^{\mathcal{V}}$, one first generates an independent random set $R(\boldsymbol{y})$ by including each element $e \in \mathcal{V}$ independently with probability y_e. Equivalently, one may view this step as generating a Bernoulli vector $\boldsymbol{x}^{\mathrm{ind}} \sim R(\boldsymbol{y})$. Since the independently rounded set may violate the feasibility constraints, a CRS π is then applied to obtain a feasible vector

$$\boldsymbol{x} = \pi(\boldsymbol{x}^{\mathrm{ind}}).$$

Formally, a CRS is said to be β-*balanced* if for every fractional point $\boldsymbol{y} \in \mathcal{Y}$ and every element $e \in \mathcal{V}$,

$$\mathbf{Pr}[e \in \pi(R(\boldsymbol{y})) \mid e \in R(\boldsymbol{y})] \geq \beta.$$

For cardinality constraints, there exists a polynomial-time monotone CRS with balancedness at least

$$\beta = 1 - \frac{e^{-k}k^k}{k!}.$$

This dimension-independent bound is asymptotically optimal as the ground-set size tends to infinity with k fixed.

We now briefly describe two contention resolution schemes for the uniform matroid, i.e., feasibility constraints of the form $\sum_{i\in\mathcal{V}} x_i \leq k$, proposed by Kashaev & Santiago (2023). Both schemes are optimal in the sense that they achieve the best possible balancedness for uniform matroids, but differ in their computational complexity.

Exponential-time optimal CRS. The first scheme is an optimal CRS that exactly attains the best possible balancedness for the uniform matroid. This CRS is defined via an explicit distribution over all feasible subsets of size at most k, constructed so as to exactly match the desired marginal probabilities. Concretely, given a fractional point $\boldsymbol{y} \in \mathcal{Y}$ and a realization $A \subseteq \mathcal{V}$ of the independently rounded set $R(\boldsymbol{y})$, the scheme proceeds as follows. If $|A| \leq k$, the set A is returned. Otherwise, when $|A| > k$, the scheme outputs a random subset $B \subseteq A$ with $|B| = k$, where each k-subset B is selected with probability

$$q_A(B) \;\propto\; 1 + \bar{y}(A \setminus B) - \bar{y}(B), \qquad \text{with } \bar{y}(S) := \frac{1}{|S|}\sum_{i\in S} y_i.$$

This distribution is chosen so that each element $i \in A$ is retained with the correct marginal probability, and it yields an optimal β-balanced CRS. A direct implementation enumerates all $\binom{|A|}{k}$ possible output subsets, which can require exponential time when k is part of the input. The marginal-based construction below avoids this enumeration.

Polynomial-time optimal CRS. The second scheme is a polynomial-time CRS that also achieves optimal balancedness. This construction avoids explicit enumeration of feasible subsets and instead relies on a carefully designed randomized selection process that trims an independently sampled set to a feasible one while preserving the required marginal guarantees. Specifically, for any realization A with $|A| > k$, the scheme defines marginal retention probabilities for each element $i \in A$ as

$$\mathbf{Pr}[i \in \pi(A)] = \frac{k - y_i}{|A|} \;+\; \frac{y(A \setminus \{i\})}{|A|(|A|-1)},$$

and ensures that the output set $\pi(A)$ is a random feasible subset of size at most k whose indicator vector has exactly these marginals. Since these marginals lie in the uniform matroid polytope, one can first decompose the marginal vector into a convex combination of vertices of the polytope, i.e., $\boldsymbol{z} = \sum_{j=1}^{m} \lambda_j \mathbf{1}_{S_j}$ with $S_j \subseteq \mathcal{V}$ and $|S_j| \leq k$, and then sample the set S_j with probability λ_j. Such a convex decomposition can be computed in polynomial time using standard polyhedral decomposition methods (see, e.g., Schrijver (2003), Corollary 40.4a). Importantly, this CRS achieves the same optimal balancedness guarantees as the exponential-time scheme. We refer the reader to Kashaev & Santiago (2023) for a more detailed analysis.

B Online Submodular Maximization and the Auxiliary Function

In this appendix, we recall the auxiliary-function construction used in the main text and explain how its gradient can be computed in practice. We also discuss how to cast Online Submodular Maximization to Online Convex Optimization via the auxiliary function.

Throughout, let $f_t : 2^{[n]} \to \mathbb{R}_{\geq 0}$ be monotone submodular, and let $F_t : [0,1]^n \to \mathbb{R}_{\geq 0}$ denote its multilinear extension. We assume $f_t(\emptyset) = 0$ for every t, and hence $F_t(\mathbf{0}) = 0$.

B.1 Definition

For general monotone submodular functions, the multilinear extension F_t is not concave. Following Zhang et al. (2022), we define

$$\Psi_t(\boldsymbol{y}) := \int_0^1 \frac{e^{z-1}}{z} F_t(z\boldsymbol{y})\, dz, \qquad \boldsymbol{y} \in \mathcal{Y}.$$

Because $F_t(\mathbf{0}) = 0$, the integrand is well defined near $z = 0$. Moreover, differentiating under the integral sign gives

$$\nabla\Psi_t(\boldsymbol{y}) = \int_0^1 e^{z-1} \nabla F_t(z\boldsymbol{y})\, dz.$$

B.2 Concave-like inequality

The usefulness of Ψ_t is that it satisfies the following concave-like inequality, which is a specialization of the auxiliary-function analysis of Zhang et al. (2022).

Lemma B.1. *For every round t and every $\boldsymbol{y}, \boldsymbol{u} \in \mathcal{Y}$,*

$$\langle \nabla\Psi_t(\boldsymbol{y}), \boldsymbol{u} - \boldsymbol{y} \rangle \geq \left(1 - \frac{1}{e}\right) F_t(\boldsymbol{u}) - F_t(\boldsymbol{y}).$$

Proof. More generally, Zhang et al. (2022) show that for a γ-weakly DR-submodular objective, the weight function $w(z) = e^{\gamma(z-1)}$ yields the bound

$$\langle \nabla\Psi(\boldsymbol{y}), \boldsymbol{u} - \boldsymbol{y} \rangle \geq \left(1 - e^{-\gamma}\right) F(\boldsymbol{u}) - F(\boldsymbol{y}).$$

In our setting, the multilinear extension of a monotone submodular function is DR-submodular, so $\gamma = 1$. Substituting $\gamma = 1$ gives

$$\langle \nabla\Psi_t(\boldsymbol{y}), \boldsymbol{u} - \boldsymbol{y} \rangle \geq \left(1 - \frac{1}{e}\right) F_t(\boldsymbol{u}) - F_t(\boldsymbol{y}),$$

which proves the claim. □

Summing the inequality above over t explains why we optimize the linearized rewards

$$\ell_t(\boldsymbol{u}) := \langle \nabla\Psi_t(\boldsymbol{y}_t), \boldsymbol{u} \rangle.$$

Indeed, applying Lemma B.1 with $\boldsymbol{y} = \boldsymbol{y}_t$ and an arbitrary comparator point $\boldsymbol{u} \in \mathcal{Y}$ gives

$$\langle \nabla\Psi_t(\boldsymbol{y}_t), \boldsymbol{u} - \boldsymbol{y}_t \rangle \geq \left(1 - \frac{1}{e}\right) F_t(\boldsymbol{u}) - F_t(\boldsymbol{y}_t).$$

Rearranging,

$$\left(1 - \frac{1}{e}\right) F_t(\boldsymbol{u}) \leq F_t(\boldsymbol{y}_t) + \ell_t(\boldsymbol{u}) - \ell_t(\boldsymbol{y}_t).$$

Now sum this inequality over all rounds $t = 1, \ldots, T$. For any fixed comparator $\boldsymbol{u} \in \mathcal{Y}$, we obtain

$$\left(1 - \frac{1}{e}\right) \sum_{t=1}^{T} F_t(\boldsymbol{u}) - \sum_{t=1}^{T} F_t(\boldsymbol{y}_t) \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{u}) - \sum_{t=1}^{T} \ell_t(\boldsymbol{y}_t).$$

The right-hand side is exactly the static regret of the OCO algorithm on the linear reward sequence $(\ell_t)_{t=1}^T$ with respect to comparator $\boldsymbol{u}$. Therefore, if the OCO method achieves sublinear static regret on the linearized rewards, then the left-hand side is also sublinear. In other words, the learner's cumulative reward on the multilinear objectives F_t is competitive, up to the factor $(1-1/e)$, with the best fixed fractional point in hindsight. This is precisely the sense in which optimizing the linearized rewards ℓ_t yields a $(1-1/e)$-approximate regret guarantee with respect to the original multilinear extensions F_t.

B.3 No $\beta(1-1/e)$-regret for Online Submodular Maximization

We now show how the auxiliary-function construction combines with a no-static-regret algorithm on the fractional domain and a rounding scheme to yield a high-probability no-$\beta(1-1/e)$-regret guarantee for online submodular maximization.

Let $\mathcal{P}_{\mathcal{Y}}$ be a policy over $\mathcal{Y}$, and let $\boldsymbol{y}_t = \mathcal{P}_{\mathcal{Y}}(\ell_1, \dots, \ell_{t-1})$ be the fractional point played at round t. Let $\Xi : \mathcal{Y} \to \mathcal{X}$ be the rounding procedure from Appendix A, and let $\boldsymbol{x}_t = \Xi(\boldsymbol{y}_t)$ be the corresponding integral action. Recall that

$$\ell_t(\boldsymbol{u}) = \langle \nabla \Psi_t(\boldsymbol{y}_t), \boldsymbol{u} \rangle.$$

Theorem B.2. *Assume that $0 \le f_t(S) \le M$ for all $S \subseteq [n]$ and all $t \in [T]$. Let $\mathcal{P}_{\mathcal{Y}}$ be a policy over $\mathcal{Y}$ whose static regret on the linear reward sequence $(\ell_t)_{t=1}^T$ is bounded by $\mathcal{R}^{\widetilde{\mathcal{C}}}_{1,\mathcal{Y}}$. Let $\beta = 1 - \frac{e^{-k}k^k}{k!}$ and $\alpha = \beta\left(1-\frac{1}{e}\right)$. Then, for any $p \in (0,1)$, the Relax-and-Round algorithm satisfies, with probability at least $1-p$,*

$$\mathcal{R}^{\mathcal{C}}_{\alpha,\mathcal{X}} \le \beta\, \mathcal{R}^{\widetilde{\mathcal{C}}}_{1,\mathcal{Y}} + M\sqrt{2T \log \frac{1}{p}}.$$

In particular, if $\mathcal{P}_{\mathcal{Y}}$ achieves $\mathcal{R}^{\widetilde{\mathcal{C}}}_{1,\mathcal{Y}} = O(\sqrt{T})$, then the resulting online submodular maximization algorithm achieves no-$\beta(1-1/e)$-regret with high probability.

Proof. Fix any comparator $\boldsymbol{u} \in \mathcal{Y}$. Summing the inequality from Lemma B.1 over $t = 1, \dots, T$ gives

$$\left(1-\frac{1}{e}\right)\sum_{t=1}^T F_t(\boldsymbol{u}) - \sum_{t=1}^T F_t(\boldsymbol{y}_t) \le \sum_{t=1}^T \ell_t(\boldsymbol{u}) - \sum_{t=1}^T \ell_t(\boldsymbol{y}_t).$$

Maximizing over $\boldsymbol{u} \in \mathcal{Y}$, we obtain

$$\left(1-\frac{1}{e}\right)\max_{\boldsymbol{u}\in\mathcal{Y}}\sum_{t=1}^T F_t(\boldsymbol{u}) - \sum_{t=1}^T F_t(\boldsymbol{y}_t) \le \mathcal{R}^{\widetilde{\mathcal{C}}}_{1,\mathcal{Y}}.$$

Since every integral action $\boldsymbol{x} \in \mathcal{X}$ also lies in $\mathcal{Y}$, and $F_t(\boldsymbol{x}) = f_t(\boldsymbol{x})$ for integral $\boldsymbol{x}$, it follows that

$$\left(1-\frac{1}{e}\right)\max_{\boldsymbol{x}\in\mathcal{X}}\sum_{t=1}^T f_t(\boldsymbol{x}) - \sum_{t=1}^T F_t(\boldsymbol{y}_t) \le \mathcal{R}^{\widetilde{\mathcal{C}}}_{1,\mathcal{Y}}.$$

Multiplying by β yields

$$\alpha \max_{\boldsymbol{x}\in\mathcal{X}}\sum_{t=1}^T f_t(\boldsymbol{x}) - \beta\sum_{t=1}^T F_t(\boldsymbol{y}_t) \le \beta\, \mathcal{R}^{\widetilde{\mathcal{C}}}_{1,\mathcal{Y}}.$$

On the other hand, by the rounding guarantee from Appendix A,

$$\mathbb{E}[f_t(\boldsymbol{x}_t) \mid \mathcal{H}_{t-1}] \ge \beta F_t(\boldsymbol{y}_t) \qquad \forall t \in [T].$$

Therefore the random variables

$$Z_t := \beta F_t(\boldsymbol{y}_t) - f_t(\boldsymbol{x}_t)$$

form a supermartingale difference sequence with respect to the filtration $(\mathcal{H}_t)_{t=0}^T$. Since $0 \le f_t(\boldsymbol{x}_t) \le M$ and $0 \le \beta F_t(\boldsymbol{y}_t) \le M$, we have $|Z_t| \le M$ for all t. Applying Azuma–Hoeffding (see Lemma C.5) gives

$$\mathbf{Pr}\left[\sum_{t=1}^{T} Z_t > M\sqrt{2T\log\frac{1}{p}}\right] \le p.$$

Equivalently, with probability at least $1-p$,

$$\beta\sum_{t=1}^{T} F_t(\boldsymbol{y}_t) - \sum_{t=1}^{T} f_t(\boldsymbol{x}_t) \le M\sqrt{2T\log\frac{1}{p}}.$$

Combining the last two displays, we conclude that with probability at least $1-p$,

$$\alpha\max_{\boldsymbol{x}\in\mathcal{X}}\sum_{t=1}^{T} f_t(\boldsymbol{x}) - \sum_{t=1}^{T} f_t(\boldsymbol{x}_t) \le \beta\,\mathcal{R}_{1,\mathcal{Y}}^{\widetilde{\mathcal{C}}} + M\sqrt{2T\log\frac{1}{p}},$$

which is exactly the claimed bound. □

B.4 Computing $\nabla\Psi_t$ in practice

The gradient of the auxiliary function is

$$\nabla\Psi_t(\boldsymbol{y}) = \int_0^1 e^{z-1}\nabla F_t(z\boldsymbol{y})\,dz.$$

Thus, computing $\nabla\Psi_t(\boldsymbol{y})$ reduces to evaluating gradients of the multilinear extension F_t at scaled points $z\boldsymbol{y}$ and integrating over $z \in [0,1]$.

When F_t and ∇F_t are tractable for the class of submodular functions under consideration, one may approximate the integral numerically using a one-dimensional quadrature rule. Since the integration is only over the scalar variable z, this step is straightforward once ∇F_t can be evaluated.

For many structured classes of monotone submodular functions, the multilinear extension admits a simple closed form, and so does its gradient. In such cases, $\nabla\Psi_t(\boldsymbol{y})$ can be computed directly from the integral representation above, without any sampling. This applies, for example, to coverage-type objectives and related models for which the multilinear extension is explicitly tractable.

More generally, one can use an unbiased stochastic estimator. Let Z be a random variable on $[0,1]$ with density

$$p(z) = \frac{e^{z-1}}{1-1/e}.$$

Then

$$\nabla\Psi_t(\boldsymbol{y}) = \left(1-\frac{1}{e}\right)\mathbb{E}\big[\nabla F_t(Z\boldsymbol{y})\big].$$

Therefore, if $\widehat{g}_t(\boldsymbol{x})$ is an unbiased estimator of $\nabla F_t(\boldsymbol{x})$, then

$$\widehat{h}_t(\boldsymbol{y}) := \left(1-\frac{1}{e}\right)\widehat{g}_t(Z\boldsymbol{y})$$

is an unbiased estimator of $\nabla\Psi_t(\boldsymbol{y})$.

For the multilinear extension, each coordinate of the gradient admits the representation

$$\frac{\partial F_t(\boldsymbol{x})}{\partial x_i} = \mathbb{E}[f_t(R_{-i}(\boldsymbol{x})\cup\{i\}) - f_t(R_{-i}(\boldsymbol{x}))]\,,$$

where $R_{-i}(\boldsymbol{x})$ includes each element $j \ne i$ independently with probability x_j. Hence, when only value-oracle access to f_t is available, one can estimate $\nabla F_t(\boldsymbol{x})$ by Monte Carlo sampling and thereby obtain an unbiased estimator of $\nabla\Psi_t(\boldsymbol{y})$ as well.

In the main text, the algorithms are stated using exact gradients for simplicity. A value-oracle implementation can instead average independent copies of the estimator above. Suppose $0 \le f_t(S) \le M$. For any $p \in (0,1)$, averaging

$$m = \max\left\{1, \left\lceil \frac{M^2 T}{2} \log \frac{2nT}{p} \right\rceil \right\}$$

copies per round gives estimates $\widehat{g}_t$ satisfying

$$\|\widehat{g}_t - \nabla \Psi_t(\boldsymbol{y}_t)\|_\infty \le T^{-1/2} \qquad \text{for all } t \in [T]$$

with probability at least $1-p$, by Hoeffding's inequality and a union bound. Since any two points in $\mathcal{Y}$ have ℓ_1-distance at most $2k$, replacing the estimated linear rewards by the true linear rewards adds at most $2k\sqrt{T}$ to the regret bound.

C Affine and Quadratic Transformations

C.1 Description of $\widetilde{\Phi}_1$

Recall that $z_1(\boldsymbol{x}) = (1, \boldsymbol{x})$ and $p_1 = 1 + n$. An affine map can be written as

$$\phi_\Theta(\boldsymbol{x}) = b + A\boldsymbol{x} = \Theta z_1(\boldsymbol{x}), \qquad \Theta = [b\ A] \in \mathbb{R}^{n \times p_1}.$$

Its coefficient matrix belongs to $\mathcal{P}_1$ if and only if there exists a scalar ρ satisfying

$$\begin{aligned}
&0 \le \rho \le 1, \\
&b_i \ge 0 && \forall i \in [n], \\
&A_{ij} \ge 0 && \forall i, j \in [n], \\
&\sum_{i=1}^n A_{ij} = \rho && \forall j \in [n], \\
&b_i + \sum_{j=1}^n A_{ij} \le 1 && \forall i \in [n], \\
&\sum_{i=1}^n b_i + k\rho \le k.
\end{aligned} \tag{10}$$

Equivalently,

$$\mathcal{P}_1 = \left\{ \Theta = [b\ A] \in \mathbb{R}^{n \times p_1} \;\middle|\; \exists \rho \text{ satisfying (10)} \right\}.$$

To verify this characterization, first observe that $A \ge 0$ is equivalent to coordinatewise monotonicity of ϕ_Θ. Under monotonicity, the minimum and maximum of each output coordinate over the cube occur at $\mathbf{0}$ and $\mathbf{1}$, respectively. Thus, the conditions $b_i \ge 0$ and $b_i + \sum_j A_{ij} \le 1$ are equivalent to $\phi_\Theta([0,1]^n) \subseteq [0,1]^n$. Moreover, the column-sum constraints are equivalent to the affine total-mass identity

$$\mathbf{1}^\top \phi_\Theta(\boldsymbol{x}) = \mathbf{1}^\top b + \rho\, \mathbf{1}^\top \boldsymbol{x} \qquad \forall \boldsymbol{x} \in [0,1]^n.$$

The budget constraint therefore ensures that, for every $\boldsymbol{x} \in \mathcal{Y}$,

$$\mathbf{1}^\top \phi_\Theta(\boldsymbol{x}) \le \mathbf{1}^\top b + \rho k \le k.$$

In particular, $b \in \mathcal{Y}$ and $\phi_\Theta(\mathcal{Y}) \subseteq \mathcal{Y}$.

The formulation has $n^2 + n + 1$ variables, including ρ, and $\mathcal{O}(n^2)$ linear constraints. Every feasible point satisfies $0 \le b_i \le 1$, $0 \le A_{ij} \le 1$, and $0 \le \rho \le 1$. Hence the extended feasible set is a bounded polyhedron, and its projection onto the coefficient coordinates Θ is a compact convex polytope. Consequently,

$$\widetilde{\Phi}_1 = \{\phi_\Theta : \boldsymbol{x} \mapsto \Theta z_1(\boldsymbol{x}) \mid \Theta \in \mathcal{P}_1\}.$$

Every feasible constant map is obtained by taking $b \in \mathcal{Y}$, $A = 0$, and $\rho = 0$. At $\rho = 1$, the budget constraint forces $b = 0$. The column sums of A then equal one and its row sums are at most one. Since the sum of all entries is n, every row sum also equals one, recovering the doubly stochastic matrices.

C.2 DESCRIPTION OF $\widetilde{\Phi}_2$

Recall the feature maps

$$q(\boldsymbol{x}) = (x_p x_q)_{1 \le p < q \le n}, \qquad z_2(\boldsymbol{x}) = (1, \boldsymbol{x}, q(\boldsymbol{x})), \qquad p_2 = 1 + n + \binom{n}{2}.$$

A multilinear map of degree at most two can be written as

$$\phi_\Theta(\boldsymbol{x}) = b + A\boldsymbol{x} + Qq(\boldsymbol{x}) = \Theta z_2(\boldsymbol{x}), \qquad \Theta = [b\ A\ Q] \in \mathbb{R}^{n \times p_2}.$$

We index the columns of Q by unordered pairs and write $Q_{i,\{p,q\}}$ for the coefficient of $x_p x_q$ in output coordinate i.

The coefficient domain $\mathcal{P}_2$ admits the following extended formulation. Introduce a scalar ρ and auxiliary variables $s_{i,h,j}$ for $i, j \in [n]$ and $h \in [n] \setminus \{j\}$, and impose

$$0 \le \rho \le 1, \tag{11a}$$

$$b_i \ge 0 \qquad \forall i \in [n], \tag{11b}$$

$$\sum_{i=1}^{n} A_{ij} = \rho \qquad \forall j \in [n], \tag{11c}$$

$$\sum_{i=1}^{n} Q_{i,\{p,q\}} = 0 \qquad \forall 1 \le p < q \le n, \tag{11d}$$

$$s_{i,h,j} \le 0 \qquad \forall i, j \in [n],\ h \ne j, \tag{11e}$$

$$s_{i,h,j} \le Q_{i,\{j,h\}} \qquad \forall i, j \in [n],\ h \ne j, \tag{11f}$$

$$A_{ij} + \sum_{h \ne j} s_{i,h,j} \ge 0 \qquad \forall i, j \in [n], \tag{11g}$$

$$b_i + \sum_{j=1}^{n} A_{ij} + \sum_{1 \le p < q \le n} Q_{i,\{p,q\}} \le 1 \qquad \forall i \in [n], \tag{11h}$$

$$\sum_{i=1}^{n} b_i + k\rho \le k. \tag{11i}$$

Then

$$\mathcal{P}_2 = \left\{ \Theta = [b\ A\ Q] \in \mathbb{R}^{n \times p_2} \;\middle|\; \exists \rho, s \text{ satisfying (11)} \right\}.$$

We next verify that this formulation is exact.

Matching the coefficients of the singleton and pairwise monomials shows that (11c)–(11d) are equivalent to the affine total-mass identity

$$\mathbf{1}^\top \phi_\Theta(\boldsymbol{x}) = \mathbf{1}^\top b + \rho\, \mathbf{1}^\top \boldsymbol{x} \qquad \forall \boldsymbol{x} \in [0,1]^n.$$

For monotonicity, each partial derivative is affine:

$$\frac{\partial \phi_{\Theta,i}(\boldsymbol{x})}{\partial x_j} = A_{ij} + \sum_{h \ne j} Q_{i,\{j,h\}} x_h.$$

Minimizing over the cube gives

$$\min_{\boldsymbol{x} \in [0,1]^n} \frac{\partial \phi_{\Theta,i}(\boldsymbol{x})}{\partial x_j} = A_{ij} + \sum_{h \ne j} \min\{0, Q_{i,\{j,h\}}\}.$$

If (11e)–(11g) hold, then $s_{i,h,j} \le \min\{0, Q_{i,\{j,h\}}\}$, so every partial derivative is nonnegative throughout the cube. Conversely, if ϕ_Θ is coordinatewise nondecreasing, the choice

$$s_{i,h,j} = \min\{0, Q_{i,\{j,h\}}\}$$

satisfies all three constraints. Thus these auxiliary constraints are equivalent to monotonicity.

Under monotonicity, the minimum and maximum of output coordinate i are

$$\phi_{\Theta,i}(\mathbf{0}) = b_i, \qquad \phi_{\Theta,i}(\mathbf{1}) = b_i + \sum_{j=1}^{n} A_{ij} + \sum_{1\le p<q\le n} Q_{i,\{p,q\}}.$$

Consequently, (11b) and (11h) are equivalent to the cube range condition. Finally, the affine total-mass identity and (11i) imply

$$\mathbf{1}^\top \phi_\Theta(\boldsymbol{x}) \le \mathbf{1}^\top b + \rho k \le k \qquad \forall \boldsymbol{x} \in \mathcal{Y}.$$

In particular, $b \in \mathcal{Y}$ and $\phi_\Theta(\mathcal{Y}) \subseteq \mathcal{Y}$. Together, these arguments establish

$$\widetilde{\Phi}_2 = \{\phi_\Theta : \boldsymbol{x} \mapsto \Theta z_2(\boldsymbol{x}) \mid \Theta \in \mathcal{P}_2\}.$$

The formulation includes every feasible constant map by setting $A = Q = 0$, $s = 0$, and $\rho = 0$. Setting $Q = 0$ recovers $\widetilde{\Phi}_1$, so $\widetilde{\Phi}_1 \subseteq \widetilde{\Phi}_2$.

There are np_2 coefficient variables, one variable ρ, and $n^2(n-1)$ auxiliary variables $s_{i,h,j}$. Hence the extended formulation has $\mathcal{O}(n^3)$ variables and linear constraints. To establish boundedness, let $J_\Theta(\boldsymbol{x})$ denote the Jacobian of a feasible map. Monotonicity and affine total mass give

$$J_{\Theta,ij}(\boldsymbol{x}) \ge 0, \qquad \sum_{i=1}^{n} J_{\Theta,ij}(\boldsymbol{x}) = \rho,$$

so every Jacobian entry lies in $[0, \rho]$. In particular,

$$0 \le A_{ij} = J_{\Theta,ij}(\mathbf{0}) \le \rho, \qquad Q_{i,\{p,q\}} = J_{\Theta,ip}(e_q) - J_{\Theta,ip}(\mathbf{0}) \in [-\rho, \rho].$$

The range condition also gives $0 \le b_i \le 1$. Since all $s_{i,h,j}$ are nonpositive, (11g) implies $-A_{ij} \le s_{i,h,j} \le 0$. Thus every variable in the extended formulation is bounded. Its feasible set is a bounded polyhedron, and its projection $\mathcal{P}_2$ onto the coefficient coordinates is a compact convex polytope.

Projection. Both coefficient descriptions support the Euclidean projection used by the online algorithm. For the affine class, given a target coefficient matrix $\widehat{\Theta} \in \mathbb{R}^{n\times p_1}$, the projection onto $\mathcal{P}_1$ is obtained by solving

$$\begin{aligned} \min_{\Theta,\rho} \quad & \frac{1}{2}\|\Theta - \widehat{\Theta}\|_F^2 \\ \text{subject to} \quad & (10). \end{aligned}$$

For the quadratic class, given $\widehat{\Theta} \in \mathbb{R}^{n\times p_2}$, the projection onto $\mathcal{P}_2$ is obtained by solving

$$\begin{aligned} \min_{\Theta,\rho,s} \quad & \frac{1}{2}\|\Theta - \widehat{\Theta}\|_F^2 \\ \text{subject to} \quad & (11). \end{aligned}$$

Both programs are convex quadratic programs of polynomial size and can be solved to prescribed accuracy in polynomial time.

C.3 Transfer Lemma

In this subsection, we establish a transfer inequality for coordinatewise nondecreasing multilinear transformations that map Boolean vectors to Boolean vectors. The argument uses monotonicity and submodularity of the objective, together with the fact that rounding only deletes elements.

Fix a round t, and condition throughout on the filtration $\mathcal{H}_{t-1}$. Let $\boldsymbol{y}_t \in \mathcal{Y}$ be the fractional point chosen by the OCO policy, let

$$\boldsymbol{x}_t^{\mathrm{ind}} \sim R(\boldsymbol{y}_t)$$

denote the independent Bernoulli rounding of $\boldsymbol{y}_t$, and let

$$\boldsymbol{x}_t = \pi(\boldsymbol{x}_t^{\mathrm{ind}})$$

be the output of a monotone contention resolution scheme. Then $\boldsymbol{x}_t \le \boldsymbol{x}_t^{\mathrm{ind}}$.

Let $f_t : 2^{[n]} \to \mathbb{R}_{\ge 0}$ be monotone submodular, let $F_t : [0,1]^n \to \mathbb{R}_{\ge 0}$ be its multilinear extension, and let $\phi : [0,1]^n \to [0,1]^n$ be a coordinatewise nondecreasing multilinear map such that $\phi(\{0,1\}^n) \subseteq \{0,1\}^n$.

We begin with the correlation inequality that we use.

Lemma C.1 (Harris/FKG inequality for product measures)**.** *Let $W = (W_1, \dots, W_m)$ have independent coordinates. If $g, h : \{0,1\}^m \to \mathbb{R}$ are coordinatewise nondecreasing, then*

$$\mathbb{E}[g(W)h(W)] \geq \mathbb{E}[g(W)]\,\mathbb{E}[h(W)].$$

Equivalently, if g is coordinatewise nondecreasing and q is coordinatewise nonincreasing, then

$$\mathbb{E}[g(W)q(W)] \leq \mathbb{E}[g(W)]\,\mathbb{E}[q(W)].$$

Proof. This is the classical Harris inequality for product measures; see, e.g., Harris (1960). The second statement follows by applying the first to g and $M - q$, where M is any constant such that $M - q \geq 0$. □

Lemma C.2 (Dominance of monotone images of product measures)**.** *Let $W = (W_1, \dots, W_m)$ have independent Bernoulli coordinates. Let*

$$Z = (\psi_1(W), \dots, \psi_k(W)) \in \{0,1\}^k,$$

where each $\psi_j : \{0,1\}^m \to \{0,1\}$ is coordinatewise nondecreasing. Let $f : 2^{[k]} \to \mathbb{R}$ be submodular, and let $F : [0,1]^k \to \mathbb{R}$ be its multilinear extension. Then

$$\mathbb{E}[f(Z)] \leq F(\mathbb{E}[Z]).$$

Proof. We argue by induction on k.

If $k = 1$, then F is affine and the claim is immediate.

Assume the result holds for output dimension $k - 1$. Write

$$Z = (U, Z_k), \qquad U = (Z_1, \dots, Z_{k-1}), \qquad p_k = \mathbb{E}[Z_k].$$

Define

$$\Delta_k(u) = f(u,1) - f(u,0), \qquad u \in \{0,1\}^{k-1}.$$

Since f is submodular, Δ_k is coordinatewise nonincreasing. Also,

$$f(Z) = f(U,0) + Z_k\,\Delta_k(U),$$

so taking expectations gives

$$\mathbb{E}[f(Z)] = \mathbb{E}[f(U,0)] + \mathbb{E}[Z_k \Delta_k(U)].$$

Since $Z_k = \psi_k(W)$ is coordinatewise nondecreasing in W, while $\Delta_k(U(W))$ is coordinatewise nonincreasing in W, Lemma C.1 gives

$$\mathbb{E}[Z_k \Delta_k(U)] \leq \mathbb{E}[Z_k]\,\mathbb{E}[\Delta_k(U)] = p_k\,\mathbb{E}[\Delta_k(U)].$$

Therefore,

$$\mathbb{E}[f(Z)] \leq \mathbb{E}[f(U,0)] + p_k\,\mathbb{E}[\Delta_k(U)].$$

For every $u \in [0,1]^{k-1}$,

$$F(u, p_k) = F(u,0) + p_k\big(F(u,1) - F(u,0)\big).$$

Since F agrees with f on $\{0,1\}^k$, we have

$$F(U, p_k) = f(U,0) + p_k\,\Delta_k(U).$$

Hence

$$\mathbb{E}[f(Z)] \leq \mathbb{E}[F(U, p_k)].$$

Now define

$$g(A) = (1 - p_k)f(A) + p_k f(A \cup \{k\}).$$

Then g is submodular, and its multilinear extension is

$$G(u) = F(u, p_k).$$

Since $U = (\psi_1(W), \dots, \psi_{k-1}(W))$ is again a coordinatewise nondecreasing image of the same Bernoulli vector W, the induction hypothesis applies to U and g. Therefore

$$\mathbb{E}[F(U, p_k)] = \mathbb{E}[G(U)] \leq G(\mathbb{E}[U]) = F(\mathbb{E}[U], p_k) = F(\mathbb{E}[Z]).$$

This completes the proof. □

We can now prove the transfer step from the integral action to the fractional benchmark.

Lemma C.3 (Transfer lemma)**.** *For every round t and every coordinatewise nondecreasing multilinear map ϕ such that $\phi(\{0,1\}^n) \subseteq \{0,1\}^n$,*

$$\mathbb{E}[f_t(\phi(\boldsymbol{x}_t)) \mid \mathcal{H}_{t-1}] \le F_t(\phi(\boldsymbol{y}_t)).$$

Proof. Set $S = \boldsymbol{x}_t^{\mathrm{ind}}$. Since $\boldsymbol{x}_t \le S$, and both f_t and ϕ are coordinatewise nondecreasing,

$$f_t(\phi(\boldsymbol{x}_t)) \le f_t(\phi(S))$$

Taking conditional expectations gives

$$\mathbb{E}[f_t(\phi(\boldsymbol{x}_t)) \mid \mathcal{H}_{t-1}] \le \mathbb{E}[f_t(\phi(S)) \mid \mathcal{H}_{t-1}].$$

Set $Z = \phi(S)$. By Lemma C.2,

$$\mathbb{E}[f_t(\phi(S)) \mid \mathcal{H}_{t-1}] = \mathbb{E}[f_t(Z) \mid \mathcal{H}_{t-1}] \le F_t(\mathbb{E}[Z \mid \mathcal{H}_{t-1}]).$$

Since the coordinates of S are independent Bernoulli random variables with means $\boldsymbol{y}_t$, and each coordinate of ϕ is multilinear,

$$\mathbb{E}[Z \mid \mathcal{H}_{t-1}] = \mathbb{E}[\phi(S) \mid \mathcal{H}_{t-1}] = \phi(\mathbb{E}[S \mid \mathcal{H}_{t-1}]) = \phi(\boldsymbol{y}_t).$$

Therefore

$$\mathbb{E}[f_t(\phi(\boldsymbol{x}_t)) \mid \mathcal{H}_{t-1}] \le F_t(\phi(\boldsymbol{y}_t)).$$

□

As an immediate consequence, for every fixed transformation ϕ in the class under consideration, the process

$$d_{t,\phi} = f_t(\phi(\boldsymbol{x}_t)) - F_t(\phi(\boldsymbol{y}_t))$$

is a supermartingale difference sequence with respect to $(\mathcal{H}_t)_{t\ge 0}$.

C.4 Proof of Theorem 6.3

Theorem C.4 (Restatement of Theorem 6.3)**.** *Fix $\ell \in \{1,2\}$ and let $d_\ell = np_\ell$. Assume that $0 \le f_t(S) \le M$ for all $S \subseteq [n]$ and $t \in [T]$. Assume that P_Y achieves regret $R_{1,Y}^{\widetilde{\Phi}_\ell}$ on the surrogate linear rewards $\ell_t(u) = \langle \nabla \Psi_t(y_t), u\rangle$. Let $\beta = 1 - \frac{e^{-k}k^k}{k!}$ and $\alpha = \beta(1-1/e)$. Then, for any $p \in (0,1)$, the Relax-and-Round algorithm (Algorithm 1) using the CRS rounding from Appendix A satisfies, with probability at least $1-p$,*

$$R_{\alpha,X}^{\Phi_\ell} \le \beta R_{1,Y}^{\widetilde{\Phi}_\ell} + O\left(M\sqrt{T\left(d_\ell + \log\frac{1}{p}\right)}\right).$$

Throughout this subsection, fix $\ell \in \{1,2\}$. Let $\boldsymbol{y}_t \in \mathcal{Y}$ be the fractional action and $\boldsymbol{x}_t \in \mathcal{X}$ its rounded counterpart. We use the same pre-round conditioning as in Appendix C.3: $\boldsymbol{y}_t$ and f_t are fixed before the fresh rounding randomness at round t. All conditional expectations below are with respect to the corresponding filtration $(\mathcal{H}_t)_{t=0}^T$.

We first state the concentration inequality used in the proof.

Lemma C.5 (Azuma–Hoeffding inequality for supermartingales)**.** *Let $(X_t)_{t=1}^T$ be adapted to a filtration $(\mathcal{H}_t)_{t=0}^T$, and suppose that*

$$\mathbb{E}[X_t \mid \mathcal{H}_{t-1}] \le 0, \qquad |X_t| \le c \qquad \forall t \in [T],$$

almost surely, for a deterministic constant $c \ge 0$. Then, for every $q \in (0,1)$, with probability at least $1-q$,

$$\sum_{t=1}^T X_t \le c\sqrt{2T\log\frac{1}{q}}.$$

Counting integral transformations. Every transformation in Φ_ℓ has a unique coefficient representation $\phi_\Theta(\boldsymbol{x}) = \Theta z_\ell(\boldsymbol{x})$. The coefficient bounds established in Appendices C.1 and C.2 give

$$0 \le b_i \le 1, \qquad 0 \le A_{ij} \le 1, \qquad |Q_{i,\{p,q\}}| \le 1,$$

where the quadratic coefficients are present only when $\ell = 2$.

For an integral transformation, these coefficients are integers. Indeed, writing e_j for the j-th standard basis vector,

$$b_i = \phi_i(\mathbf{0}), \qquad A_{ij} = \phi_i(e_j) - \phi_i(\mathbf{0}),$$

and, when $\ell = 2$,

$$Q_{i,\{p,q\}} = \phi_i(e_p + e_q) - \phi_i(e_p) - \phi_i(e_q) + \phi_i(\mathbf{0}).$$

All evaluations on the right-hand sides are Boolean. Consequently,

$$b_i, A_{ij} \in \{0, 1\}, \qquad Q_{i,\{p,q\}} \in \{-1, 0, 1\}.$$

Since there are $d_\ell = np_\ell$ coefficients, this implies $|\Phi_\ell| \le 3^{d_\ell}$. This counting bound is used only in the analysis; the algorithm optimizes over the fractional coefficient domain $\mathcal{P}_\ell$.

Lemma C.6 (Comparator concentration)**.** *Fix $q \in (0, 1)$. Then, with probability at least $1 - q$,*

$$\max_{\phi \in \Phi_\ell} \sum_{t=1}^{T} f_t(\phi(\boldsymbol{x}_t)) \le \max_{\phi \in \widetilde{\Phi}_\ell} \sum_{t=1}^{T} F_t(\phi(\boldsymbol{y}_t)) + M\sqrt{2T\left(d_\ell \log 3 + \log \frac{1}{q}\right)}.$$

Proof. Fix $\phi \in \Phi_\ell$ and define

$$D_{t,\phi} = f_t(\phi(\boldsymbol{x}_t)) - F_t(\phi(\boldsymbol{y}_t)).$$

By Lemma C.3,

$$\mathbb{E}[D_{t,\phi} \mid \mathcal{H}_{t-1}] \le 0.$$

Both terms defining $D_{t,\phi}$ lie in $[0, M]$, so $|D_{t,\phi}| \le M$. Applying Lemma C.5 with failure probability $q/|\Phi_\ell|$ and taking a union bound over Φ_ℓ shows that, with probability at least $1 - q$,

$$\sum_{t=1}^{T} D_{t,\phi} \le M\sqrt{2T\left(\log |\Phi_\ell| + \log \frac{1}{q}\right)} \qquad \forall \phi \in \Phi_\ell.$$

Using $|\Phi_\ell| \le 3^{d_\ell}$, taking the maximum over $\phi \in \Phi_\ell$, and then using $\Phi_\ell \subseteq \widetilde{\Phi}_\ell$ proves the claim. □

We next bound the learner's realized reward using the conditional rounding guarantee.

Lemma C.7 (Learner-side concentration)**.** *Let $q \in (0, 1)$. Assume that*

$$\mathbb{E}[f_t(\boldsymbol{x}_t) \mid \mathcal{H}_{t-1}] \ge \beta F_t(\boldsymbol{y}_t) \qquad \forall t \in [T].$$

Then, with probability at least $1 - q$,

$$\sum_{t=1}^{T} f_t(\boldsymbol{x}_t) \ge \beta \sum_{t=1}^{T} F_t(\boldsymbol{y}_t) - M\sqrt{2T \log \frac{1}{q}}.$$

Proof. Define

$$D_t = \beta F_t(\boldsymbol{y}_t) - f_t(\boldsymbol{x}_t).$$

The conditional rounding guarantee gives $\mathbb{E}[D_t \mid \mathcal{H}_{t-1}] \le 0$. Since $0 \le F_t(\boldsymbol{y}_t) \le M$, $0 \le f_t(\boldsymbol{x}_t) \le M$, and $0 \le \beta \le 1$, we also have $|D_t| \le M$. The claim follows from Lemma C.5. □

Proof of Theorem 6.3. For every $\phi \in \widetilde{\Phi}_\ell$, we have $\phi(\boldsymbol{y}_t) \in \mathcal{Y}$. Applying the auxiliary-function inequality with comparator $\phi(\boldsymbol{y}_t)$ gives

$$\left(1 - \frac{1}{e}\right) F_t(\phi(\boldsymbol{y}_t)) - F_t(\boldsymbol{y}_t) \le \langle \nabla \Psi_t(\boldsymbol{y}_t), \phi(\boldsymbol{y}_t) - \boldsymbol{y}_t \rangle .$$

Summing over t, maximizing over $\phi \in \widetilde{\Phi}_\ell$, and using the assumed surrogate-regret guarantee yields

$$\left(1 - \frac{1}{e}\right) \max_{\phi \in \widetilde{\Phi}_\ell} \sum_{t=1}^{T} F_t(\phi(\boldsymbol{y}_t)) - \sum_{t=1}^{T} F_t(\boldsymbol{y}_t) \le \mathcal{R}_{1,\mathcal{Y}}^{\widetilde{\Phi}_\ell}.$$

Multiplying by β gives

$$\alpha \max_{\phi \in \widetilde{\Phi}_\ell} \sum_{t=1}^{T} F_t(\phi(\boldsymbol{y}_t)) \le \beta \sum_{t=1}^{T} F_t(\boldsymbol{y}_t) + \beta\, \mathcal{R}_{1,\mathcal{Y}}^{\widetilde{\Phi}_\ell}.$$

Apply Lemmas C.6 and C.7, each with failure probability $p/2$, and define

$$\Delta_{\mathrm{cmp}} = M\sqrt{2T\left(d_\ell \log 3 + \log \frac{2}{p}\right)}, \qquad \Delta_{\mathrm{crs}} = M\sqrt{2T \log \frac{2}{p}}.$$

By a union bound, with probability at least $1 - p$, both concentration bounds hold. On this event,

$$\begin{aligned}
\mathcal{R}_{\alpha,\mathcal{X}}^{\Phi_\ell} &= \alpha \max_{\phi \in \Phi_\ell} \sum_{t=1}^{T} f_t(\phi(\boldsymbol{x}_t)) - \sum_{t=1}^{T} f_t(\boldsymbol{x}_t) \\
&\le \alpha \max_{\phi \in \widetilde{\Phi}_\ell} \sum_{t=1}^{T} F_t(\phi(\boldsymbol{y}_t)) + \alpha\Delta_{\mathrm{cmp}} - \sum_{t=1}^{T} f_t(\boldsymbol{x}_t) \\
&\le \beta\, \mathcal{R}_{1,\mathcal{Y}}^{\widetilde{\Phi}_\ell} + \alpha\Delta_{\mathrm{cmp}} + \Delta_{\mathrm{crs}}.
\end{aligned}$$

Since $\alpha \le 1$, the definitions of Δ_{cmp} and Δ_{crs} give

$$R_{\alpha,X}^{\Phi_\ell} \le \beta R_{1,Y}^{\widetilde{\Phi}_\ell} + O\left(M\sqrt{T\left(d_\ell + \log \frac{1}{p}\right)}\right),$$

which proves the claimed bound. Finally, since Φ_ℓ contains every constant map to a feasible integral action, the same guarantee also controls static regret. □

D Approximate Fixed Points

D.1 Fixed points of $\widetilde{\Phi}_\ell$

Lemma D.1 (Restatement of Lemma 6.1)**.** *Let $\ell \in \{1, 2\}$ and let $\phi \in \widetilde{\Phi}_\ell$, with associated parameter $\rho \in [0, 1]$. Then*

$$\|\phi(\boldsymbol{x}) - \phi(\boldsymbol{y})\|_1 \le \rho\|\boldsymbol{x} - \boldsymbol{y}\|_1 \le \|\boldsymbol{x} - \boldsymbol{y}\|_1,\ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{Y}.$$

Proof. Since ϕ is a multilinear polynomial map, it is continuously differentiable. Let $J(\boldsymbol{x}) = D\phi(\boldsymbol{x})$ denote its Jacobian. Coordinatewise monotonicity implies

$$J_{ij}(\boldsymbol{x}) \ge 0, \qquad \forall i, j \in [n], \quad \forall \boldsymbol{x} \in [0, 1]^n,$$

where the boundary cases follow by continuity.

Writing $\mathbf{b} = \phi(\mathbf{0})$, the affine total-mass identity gives

$$\sum_{i=1}^{n} \phi_i(\boldsymbol{x}) = \sum_{i=1}^{n} b_i + \rho \sum_{i=1}^{n} x_i.$$

Differentiating with respect to x_j yields

$$\sum_{i=1}^{n} J_{ij}(\boldsymbol{x}) = \rho, \qquad \forall j \in [n], \quad \forall \boldsymbol{x} \in [0,1]^n.$$

Thus the induced ℓ_1 operator norm satisfies

$$\|J(\boldsymbol{x})\|_{1\to 1} = \max_{j\in[n]} \sum_{i=1}^{n} |J_{ij}(\boldsymbol{x})| = \rho.$$

For any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{Y}$, convexity of $\mathcal{Y}$ ensures that the segment joining them lies in $\mathcal{Y}$. By the fundamental theorem of calculus,

$$\phi(\boldsymbol{y}) - \phi(\boldsymbol{x}) = \int_0^1 J\big(\boldsymbol{x} + t(\boldsymbol{y}-\boldsymbol{x})\big)(\boldsymbol{y}-\boldsymbol{x})\, dt.$$

Consequently,

$$\begin{aligned}\|\phi(\boldsymbol{y}) - \phi(\boldsymbol{x})\|_1 &\le \int_0^1 \big\|J\big(\boldsymbol{x} + t(\boldsymbol{y}-\boldsymbol{x})\big)\big\|_{1\to 1} \|\boldsymbol{y}-\boldsymbol{x}\|_1\, dt \\ &= \rho\|\boldsymbol{y}-\boldsymbol{x}\|_1 \\ &\le \|\boldsymbol{y}-\boldsymbol{x}\|_1,\end{aligned}$$

as claimed. □

D.2 Anchored Fixed Points

Lemma D.2 (Restatement of Lemma 6.2)**.** *Let $\psi : \mathcal{Y} \to \mathcal{Y}$ be non-expansive in ℓ_1, and suppose that $\mathcal{Y}$ has ℓ_1-diameter at most D. Set $\epsilon = 1/\sqrt{T}$ and $\mu = 2D\sqrt{T}/(2D\sqrt{T}+1)$. Starting from $\boldsymbol{x}^0 = \boldsymbol{x}_0 \in \mathcal{Y}$, consider the iteration*

$$\boldsymbol{x}^{s+1} = T_\mu(\boldsymbol{x}^s) = (1-\mu)\boldsymbol{x}_0 + \mu\psi(\boldsymbol{x}^s).$$

After $s = \mathcal{O}(D\sqrt{T}\log(2 + D\sqrt{T}))$ iterations, the iterate satisfies $\|\psi(\boldsymbol{x}^s) - \boldsymbol{x}^s\|_1 \le \frac{1}{\sqrt{T}}$.

Proof of Lemma 6.2. For all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{Y}$,

$$\|T_\mu(\boldsymbol{x}) - T_\mu(\boldsymbol{y})\|_1 = \mu\|\psi(\boldsymbol{x}) - \psi(\boldsymbol{y})\|_1 \le \mu\|\boldsymbol{x}-\boldsymbol{y}\|_1,$$

so T_μ is a contraction with factor μ and therefore admits a unique fixed point $\boldsymbol{x}_\mu \in \mathcal{Y}$. For any $\boldsymbol{x} \in \mathcal{Y}$ we have

$$T_\mu(\boldsymbol{x}) - \boldsymbol{x} = (1-\mu)(\boldsymbol{x}_0 - \boldsymbol{x}) + \mu(\psi(\boldsymbol{x}) - \boldsymbol{x}),$$

which implies

$$\mu(\psi(\boldsymbol{x}) - \boldsymbol{x}) = (T_\mu(\boldsymbol{x}) - \boldsymbol{x}) + (1-\mu)(\boldsymbol{x} - \boldsymbol{x}_0).$$

Taking ℓ_1 norms and using $\|\boldsymbol{x} - \boldsymbol{x}_0\|_1 \le D$ yields

$$\|\psi(\boldsymbol{x}) - \boldsymbol{x}\|_1 \le \frac{1}{\mu}\|T_\mu(\boldsymbol{x}) - \boldsymbol{x}\|_1 + \frac{1-\mu}{\mu}D. \tag{12}$$

With the above choice of μ and $\epsilon = 1/\sqrt{T}$ we have $\frac{1-\mu}{\mu}D \le \epsilon/2$, so it suffices to find an iterate satisfying

$$\|T_\mu(\boldsymbol{x}) - \boldsymbol{x}\|_1 \le \frac{\mu\epsilon}{2} = \frac{\mu}{2\sqrt{T}}. \tag{13}$$

Running power iteration $\boldsymbol{x}^{t+1} = T_\mu(\boldsymbol{x}^t)$, contraction of T_μ implies

$$\|\boldsymbol{x}^t - \boldsymbol{x}_\mu\|_1 \le \mu^t\|\boldsymbol{x}^0 - \boldsymbol{x}_\mu\|_1 \le \mu^t D.$$

Moreover,

$$\|T_\mu(\boldsymbol{x}^t) - \boldsymbol{x}^t\|_1 = \|\boldsymbol{x}^{t+1} - \boldsymbol{x}^t\|_1 \le \|\boldsymbol{x}^{t+1} - \boldsymbol{x}_\mu\|_1 + \|\boldsymbol{x}^t - \boldsymbol{x}_\mu\|_1 \le (\mu^{t+1} + \mu^t)D \le 2\mu^t D.$$

Hence the stopping condition (13) is satisfied whenever $2\mu^t D \le \mu/(2\sqrt{T})$, equivalently $\mu^{t-1} \le 1/(4D\sqrt{T})$, which holds for

$$t \ge 1 + \frac{\log(4D\sqrt{T})}{-\log \mu}.$$

Substituting $\mu = \frac{2D\sqrt{T}}{2D\sqrt{T}+1}$ and using $-\log\mu \ge 1-\mu = 1/(2D\sqrt{T}+1)$ yields

$$t = \mathcal{O}\Big(D\sqrt{T}\,\log(D\sqrt{T})\Big).$$

Combining this with (12) shows that $\|\psi(\boldsymbol{x}^t) - \boldsymbol{x}^t\|_1 \le 1/\sqrt{T}$. □

E Separation of Static, Affine, and Quadratic Regret

We give two examples separating the three regret notions. Both use $k = 1$ and linear rewards $f_t(\boldsymbol{x}) = \boldsymbol{w}_t^\top \boldsymbol{x}$. The action sequences are prescribed; we do not claim that a particular online algorithm generates them.

E.1 Static versus Affine Regret

Let $n = 3$ and repeat the following three-round cycle:

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = (\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_3), \qquad (\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3).$$

The action sequence earns 1 per cycle, as does the best fixed feasible action. Hence static regret is zero.

The permutation $\phi(\boldsymbol{x}) = (x_2, x_1, x_3)$ belongs to Φ_1. It transforms the actions into $(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_3)$ and earns 2 per cycle. Thus, whenever T is divisible by three,

$$R_{\text{static}}(T) = 0, \qquad R_{\Phi_1}(T) \ge T/3.$$

The same lower bound holds for quadratic regret because $\Phi_1 \subseteq \Phi_2$.

E.2 Affine versus Quadratic Regret

Let $n = 4$ and repeat the actions $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$. The corresponding reward vectors are the rows of

$$W = \begin{pmatrix} 1/2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The action sequence earns 3 per cycle, while the best fixed feasible action earns 2.

We show that every affine comparator earns at most 3, even in the fractional class $\widetilde{\Phi}_1$. Write $\phi(\boldsymbol{x}) = \mathbf{b} + A\boldsymbol{x}$ and let

$$u = b_2 + b_4, \qquad v = b_1 + b_3, \qquad d = A_{11} + A_{33}.$$

The defining constraints imply $\mathbf{b} \ge 0$, $A \ge 0$, $A^\top \mathbf{1} = \rho \mathbf{1}$, $\mathbf{b} + A\mathbf{1} \le \mathbf{1}$, and $u + v + \rho \le 1$. In particular,

$$d \le 2\rho, \qquad A_{21} + A_{22} + A_{43} + A_{44} \le 2 - u.$$

Therefore the comparator's total reward in one cycle satisfies

$$\begin{aligned} \sum_{t=1}^{4} \boldsymbol{w}_t^\top \phi(\mathbf{e}_t) &= \tfrac{1}{2}v + 2u + \tfrac{1}{2}d + A_{21} + A_{22} + A_{43} + A_{44} \\ &\le 2 + u + \tfrac{1}{2}v + \rho \le 3. \end{aligned}$$

The identity attains 3, so affine regret is zero.

Now consider

$$\psi(\boldsymbol{x}) = \big(x_1 x_2,\ x_1 + x_2 - x_1 x_2,\ x_3 x_4,\ x_3 + x_4 - x_3 x_4\big).$$

This map is monotone, maps the cube into itself, preserves total mass, and maps binary vectors to binary vectors. It therefore belongs to Φ_2. It transforms the four actions into $(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_4, \mathbf{e}_4)$, earning 4 per cycle. No feasible comparator can earn more, since each round's reward is at most one.

Consequently, whenever T is divisible by four,

$$R_{\text{static}}(T) = -T/4, \qquad R_{\Phi_1}(T) = 0, \qquad R_{\Phi_2}(T) = T/4.$$

The same regret values hold when comparators range over $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$, evaluating the linear rewards on their fractional outputs.

F Lower Bound: OGA Can Have Linear Affine Regret

We show that static OGA can achieve negative expected static regret while incurring linear expected affine regret, even after applying the CRS. The construction matches the cyclical instance in Section 7. All regrets in this section are ordinary regrets, with approximation factor one.

Construction. Let $n = 6$, $k = 2$, and partition the items into pairs $B_1 = \{1, 2\}$, $B_2 = \{3, 4\}$, and $B_3 = \{5, 6\}$. Each item covers one distinct universe element. When group B_i is active, its elements have weight $1/2$ and all others have weight zero, giving

$$f_t(S) = \frac{|S \cap B_i|}{2}, \qquad F_t(\boldsymbol{y}) = \frac{1}{2} \sum_{j \in B_i} y_j.$$

Run OGA on $\mathcal{Y} = \{\boldsymbol{y} \in [0,1]^6 : \mathbf{1}^\top \boldsymbol{y} \le 2\}$:

$$\boldsymbol{y}_{t+1} = \Pi_{\mathcal{Y}}\big(\boldsymbol{y}_t + \eta \nabla F_t(\boldsymbol{y}_t)\big), \qquad \boldsymbol{y}_1 = \tfrac{1}{3}\mathbf{1}, \qquad \eta = \frac{D}{G\sqrt{T}} = 2\sqrt{\frac{2}{T}},$$

where $D = 2$ and $G = 1/\sqrt{2}$. For $T \ge 128$, set

$$\delta = \eta/2 = \sqrt{2/T}, \qquad m = \left\lfloor \frac{3}{2\delta} \right\rfloor.$$

Groups B_1, B_2, B_3 are active successively for m rounds each, repeating until round T. Let S_t be the result of rounding $\boldsymbol{y}_t$ using the CRS of Appendix A, independently across rounds.

Proposition F.1. *For this construction,*

$$\mathbb{E}[R_{\text{static}}(T)] \le -0.036\,T + O(\sqrt{T}), \qquad \mathbb{E}[R_{\Phi_1}(T)] \ge 0.070\,T - O(\sqrt{T}).$$

Consequently, OGA incurs linear expected affine and quadratic regret despite negative expected static regret for sufficiently large T.

Proof. Consider the fixed permutation

$$\phi(\boldsymbol{x}) = (x_5, x_6, x_1, x_2, x_3, x_4),$$

which moves each group to the next group in the cycle. This map belongs to Φ_1.

Projected dynamics. Symmetry keeps the two coordinates within each group equal. Their common values (u_1, u_2, u_3) sum to one throughout: the initialization saturates the budget, and projection after a positive coordinate increment preserves this equality. The squared projection distance is twice the corresponding three-coordinate distance, so the update reduces to Euclidean projection onto $\{u \ge 0 : \sum_i u_i = 1\}$.

Let $h = m\delta \in [11/8, 3/2]$. Starting from $(1/3, 1/3, 1/3)$, the first block reaches the active vertex after accumulated increment one and remains there until the block ends. In each subsequent block, order the groups as active, previously active, and remaining. The initial state has the form $(0, a, 1-a)$. After an accumulated increment $s = r\delta$, its exact state is

$$v_a(s) = \begin{cases} (2s/3,\ a - s/3,\ 1 - a - s/3), & 0 \le s \le 3(1-a), \\ ((s+1-a)/2,\ (1+a-s)/2,\ 0), & 3(1-a) \le s \le h. \end{cases}$$

The first expression applies before the remaining group reaches zero; the second applies afterwards. These formulas also hold for a discrete step crossing that boundary.

For $a \in [h/2, 1]$, we have $3(1-a) \le h$ and $h < 1 + a$. Thus the remaining group reaches zero by the block's end, while the previously active group stays positive. At the next block, the previously active coordinate is therefore

$$a_{\text{next}} = \frac{1 + h - a}{2}.$$

Starting from $a = 1$, the interval $[h/2, 1]$ is invariant, and the recurrence contracts towards $a_* = (1+h)/3$ by a factor of $1/2$. At $h = 3/2$, the stationary block trajectory is

$$v(s) = \begin{cases} (2s/3,\ 5/6 - s/3,\ 1/6 - s/3), & 0 \le s \le 1/2, \\ (s/2 + 1/12,\ 11/12 - s/2,\ 0), & 1/2 \le s \le 3/2. \end{cases}$$

Let $\boldsymbol{y}(s)$ be the six-dimensional vector obtained by repeating each coordinate of $v(s)$ twice.

Expected rounded rewards. For a fractional point $\boldsymbol{y}$, define $q_g(\boldsymbol{y}) = \mathbb{E}[|S \cap B_g|/2]$, where S is its CRS output. Let

$$p_{\boldsymbol{y}}(R) = \prod_{i \in R} y_i \prod_{i \notin R} (1 - y_i)$$

be the probability of independently sampling R. Conditional on R, an item $i \in R$ is retained with probability

$$c_i(R, \boldsymbol{y}) = \begin{cases} 1, & |R| \le 2, \\ \dfrac{2}{|R|} + \dfrac{\sum_{j \in R} y_j - |R| y_i}{|R|(|R| - 1)}, & |R| > 2. \end{cases}$$

The second expression follows by summing the CRS probabilities over the two-element subsets containing i. Hence

$$q_g(\boldsymbol{y}) = \frac{1}{2} \sum_{R \subseteq [6]} p_{\boldsymbol{y}}(R) \sum_{i \in R \cap B_g} c_i(R, \boldsymbol{y}).$$

Each q_g is therefore a polynomial of degree at most seven and is Lipschitz on the cube.

Along the stationary trajectory, the learner receives the active group's reward, whereas ϕ receives the previously active group's reward. Substituting $\boldsymbol{y}(s)$ into the preceding finite sum and integrating over the two pieces gives

$$L = \frac{2}{3} \int_0^{3/2} q_{\text{active}}(\boldsymbol{y}(s))\, ds \approx 0.370, \qquad P = \frac{2}{3} \int_0^{3/2} q_{\text{previous}}(\boldsymbol{y}(s))\, ds \approx 0.441.$$

Evaluating these polynomial integrals in exact rational arithmetic certifies

$$L - \frac{1}{3} > 0.036, \qquad P - L > 0.070.$$

Finite-horizon bounds. The contraction of the block dynamics and Lipschitz continuity of q_g imply that initialization contributes $O(m)$ total error relative to the stationary trajectory. For each stationary block, the discrete reward sum differs from δ^{-1} times the corresponding integral over $[0, h]$ by $O(1)$, giving $O(T/m)$ total error. Since $|h - 3/2| < \delta$ and the stationary block averages are Lipschitz in h, the floor in m contributes $O(T\delta)$. An incomplete final block contributes another $O(m)$. Because $m = \Theta(\sqrt{T})$ and $\delta = \Theta(T^{-1/2})$,

$$\sum_{t=1}^{T} \mathbb{E}[f_t(S_t)] = LT + O(\sqrt{T}), \qquad \sum_{t=1}^{T} \mathbb{E}[f_t(\phi(S_t))] = PT + O(\sqrt{T}).$$

If d_g is the number of rounds during which group B_g is active, the best fixed feasible set earns $\max_g d_g = T/3 + O(m)$. Consequently,

$$\mathbb{E}[R_{\text{static}}(T)] = \left(\frac{1}{3} - L\right) T + O(\sqrt{T}) \le -0.036\, T + O(\sqrt{T}).$$

Since affine regret maximizes over a class containing the fixed map ϕ,

$$\mathbb{E}[R_{\Phi_1}(T)] \geq \sum_{t=1}^{T} \mathbb{E}\big[f_t(\phi(S_t)) - f_t(S_t)\big] \geq (P-L)T - O(\sqrt{T}) \geq 0.070\,T - O(\sqrt{T}).$$

Finally, $\Phi_1 \subseteq \Phi_2$ gives the same lower bound for quadratic regret. □

G Experiments

G.1 Experimental Setup

Goal and scope. The purpose of the experiment is to illustrate the distinction between static regret and affine regret, and to evaluate learners operating directly over the affine and quadratic transformation classes. We evaluate all learners on integral actions obtained using the same contention resolution scheme (CRS) (see Appendix A). The comparison uses ordinary regret, with approximation factor one. The objective is a weighted-coverage function; thus the experiment lies within monotone submodular maximization but does not involve overlapping coverage.

We make the code used for the experiments publicly available in an anonymized GitHub repository.

Weighted-coverage instance. We construct a weighted-coverage instance with a ground set of $n = 6$ items and cardinality budget $k = 2$. The items are partitioned into three groups $B_1 = \{1, 2\}$, $B_2 = \{3, 4\}$, and $B_3 = \{5, 6\}$. The universe contains six elements, and item i covers only element i.

At round t, one group B_{a_t} is active. Its two universe elements each receive weight $1/2$, while all other elements receive weight zero. For an integral set $S \subseteq [n]$, the reward is therefore

$$f_t(S) = \frac{1}{2}|S \cap B_{a_t}|.$$

In particular, $0 \leq f_t(S) \leq 1$ for every feasible set. The multilinear extension and its gradient are

$$F_t(\boldsymbol{y}) = \frac{1}{2}\sum_{i \in B_{a_t}} y_i, \qquad [\nabla F_t(\boldsymbol{y})]_i = \frac{1}{2}\mathbf{1}\{i \in B_{a_t}\}.$$

The fractional action domain is $\mathcal{Y} = \{\boldsymbol{y} \in [0,1]^6 : \mathbf{1}^\top \boldsymbol{y} \leq 2\}$.

Reward schedules. We consider two reward schedules. In the static instance, $a_t = 1$ at every round, so the best fixed feasible action is B_1. In the cyclical instance, the active group rotates through the three groups in consecutive blocks:

$$a_t = 1 + \left(\left\lfloor \frac{t-1}{m} \right\rfloor \bmod 3\right).$$

Both schedules are fixed before the experiment and are shared by all learners. Actions are chosen before observing the current reward, after which the learner receives full-information feedback.

Algorithms and initialization. We compare three algorithms:

1. **Static OGA:** Online Gradient Ascent on $\mathcal{Y}$.
2. **Affine OGA:** OGA over $\mathcal{P}_1$, the coefficient domain of $\widetilde{\Phi}_1$, with actions obtained as approximate fixed points of the current affine transformation.
3. **Quadratic OGA:** OGA over $\mathcal{P}_2$, the coefficient domain of $\widetilde{\Phi}_2$, with actions obtained analogously.

All methods start from the uniform fractional point $\boldsymbol{y}_1 = (1/3, \ldots, 1/3)$. For both transformation learners, the initial transformation is the constant map $\phi_1(\boldsymbol{y}) = \boldsymbol{y}_1$.

Writing $\boldsymbol{g}_t = \nabla F_t(\boldsymbol{y}_t)$, the updates are

$$\boldsymbol{y}_{t+1} = \Pi_{\mathcal{Y}}(\boldsymbol{y}_t + \eta_0 \boldsymbol{g}_t)$$

for static OGA, and

$$\Theta_{t+1} = \Pi_{\mathcal{P}_\ell}\left(\Theta_t + \eta_\ell \boldsymbol{g}_t \mathbf{z}_\ell(\boldsymbol{y}_t)^\top\right), \qquad \ell \in \{1, 2\},$$

for the transformation learners.

Learning rates and block length. Each learner uses $\eta = D/(G\sqrt{T})$, where D is the diameter of its optimization domain and G bounds its update-gradient norm. For the transformation learners, distances and gradient norms use the Frobenius norm in coefficient space.

The raw action gradient has norm $1/\sqrt{2}$. Since $\sup_{\boldsymbol{y}\in\mathcal{Y}} \|\mathbf{z}_1(\boldsymbol{y})\|_2 = \sqrt{3}$ and $\sup_{\boldsymbol{y}\in\mathcal{Y}} \|\mathbf{z}_2(\boldsymbol{y})\|_2 = 2$, the corresponding coefficient-gradient bounds are $\sqrt{3/2}$ and $\sqrt{2}$. The diameters, gradient bounds, and resulting rates at $T = 10000$ are:

Learner	D	G	η
Static OGA	2	$1/\sqrt{2}$	0.0282843
Affine OGA	$\sqrt{12}$	$\sqrt{3/2}$	0.0282843
Quadratic OGA	6	$\sqrt{2}$	0.0424264

These diameters and gradient bounds are exact for this instance. The equality of the static and affine learning rates follows from their respective ratios D/G.

For these objectives, $\nabla\Psi_t = (1 - 1/e)\nabla F_t$. Using the auxiliary gradient and scaling its norm bound by the same factor leaves the update $\eta\boldsymbol{g}_t$ unchanged. We therefore implement the equivalent raw-gradient updates.

We set

$$\delta = \frac{\eta_0}{2} = \sqrt{\frac{2}{T}}, \qquad m = \left\lfloor \frac{3}{2\delta} \right\rfloor.$$

We use $T = 10000$, giving a block length of $m = 106$: the active pair changes every 106 rounds. The experiment stops after exactly T rounds, possibly before the final block is complete.

Rounding. Every learner rounds its fractional actions using the CRS described in Appendix A. We evaluate each fractional trajectory over 64 independent rounding runs, using common random numbers across learners. The fractional trajectories are deterministic and do not depend on the rounding draws.

Metrics. For each realized trajectory $(S_t)_{t=1}^T$, we report ordinary static regret

$$R_{\mathrm{static}}(T) = \max_{\substack{S\subseteq[n]\\|S|\le k}} \sum_{t=1}^{T} f_t(S) - \sum_{t=1}^{T} f_t(S_t),$$

and transformation regret

$$R_{\Phi_\ell}(T) = \max_{\phi\in\Phi_\ell} \sum_{t=1}^{T} f_t\big(\phi(\mathbf{1}_{S_t})\big) - \sum_{t=1}^{T} f_t(S_t), \qquad \ell \in \{1, 2\},$$

where we identify a feasible binary vector with its corresponding set.

Comparator optimization covers the complete integral classes, including all constant maps to feasible actions. For $n = 6$ and $k = 2$, these classes contain 742 affine and 54,742 quadratic transformations. Since items within each pair always have equal weights, swapping output coordinates within a pair does not change the reward. We avoid evaluating these equivalent transformations separately, while computing the maximum over the full comparator class exactly.

For each rounding run, we compute regret by maximizing the cumulative comparator reward over the full class, then subtracting the learner's cumulative reward. We divide by the number of rounds and average over the 64 independent runs. For transformation regret, this estimates $\mathbb{E}[R_{\Phi_\ell}(T)]/T$. Shaded regions show approximate pointwise 95% confidence intervals for this mean, computed as the sample mean plus or minus $1.96\, s/\sqrt{64}$, where s is the sample standard deviation of regret per round across runs. Negative static regret indicates that the learner outperforms every fixed feasible action; affine and quadratic regret are nonnegative because their comparator classes contain the identity transformation.

Table 1: Average regret per round on the cyclical instance with $n = 6$, $k = 2$, and $T = 10000$, over 64 independent CRS rounding runs.

Learner	Static regret	Affine regret	Quadratic regret
Static OGA	-0.030169	0.071057	0.071057
Affine OGA	-0.164763	0	0
Quadratic OGA	-0.396109	0	0

Implementation. Coefficient projections use sparse, warm-started quadratic programs solved with OSQP. The projection objective penalizes only changes in the transformation coefficients; ρ and the auxiliary variables used to describe $\mathcal{P}_2$ have zero objective weight. The solver's absolute and relative tolerances are 10^{-9}. Fixed-point computation uses anchoring and checked numerical acceleration, and every returned action satisfies $\|\phi_{\Theta_t}(\boldsymbol{y}_t) - \boldsymbol{y}_t\|_1 \leq 10^{-7}$. The implementation checks projection feasibility and the update-gradient bounds throughout each run.

G.2 Results

On the cyclical instance, static OGA achieves negative average static regret while incurring positive average affine and quadratic regret. After $T = 10000$ rounds, both transformation learners have negative static regret and zero affine and quadratic regret in every sampled run. Table 1 reports the cumulative regret divided by T, averaged over the 64 independent rounding runs.

Figure 2 shows the evolution of these metrics and the results on the static instance. On the static instance, all three learners achieve low average static regret per round: 0.003042, 0.002685, and 0.001655 for static, affine, and quadratic OGA, respectively.

These results illustrate that strong performance against fixed actions can coexist with poor performance against affine deviations, even after CRS rounding. Both transformation learners perform well under all three metrics on this instance, capturing affine deviations while remaining competitive with every fixed feasible action. The comparison concerns ordinary regret and does not establish a separation between affine and quadratic regret. The observed performance of the transformation learners is specific to this experiment and does not imply a general ordinary-regret guarantee.

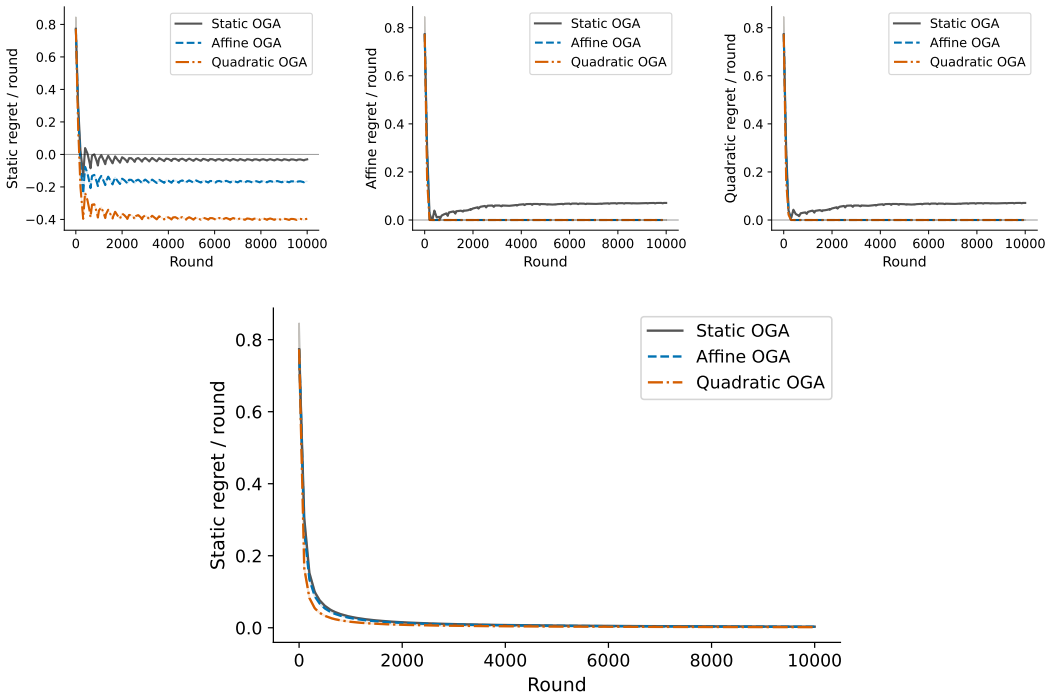


Figure 2: Average regret per round under CRS rounding with $n = 6$, $k = 2$, and $T = 10000$. Top: static (left), affine (middle), and quadratic (right) regret on the cyclical instance. Bottom: static regret on the static instance. Curves show means over 64 independent rounding runs, with shaded regions indicating 95% confidence intervals. All learners use their respective theoretical learning rates.