WHEN LOCAL GEOMETRY MAKES ONLINE LEARNING EASY

Anonymous authors
Paper under double-blind review

ABSTRACT

Convex relaxations make combinatorial online problems tractable, but the relaxed decision space can still have exponentially many vertices. We exploit a finer structure: the space is partitioned into polyhedral cells, where a cell is a region on which every loss is affine and is therefore determined by its values at that region's vertices. Existing global methods ignore these smaller local action sets, while existing cell-based arguments can conflate two different tasks: learning within a known cell and identifying the relevant cell from past observations. We separate these tasks and first study the setting in which the relevant cell and each change of cell are supplied before play. Our key insight is that the best fixed decision within an affine cell is attained at a vertex. Learning during an epoch with an unchanged cell therefore reduces exactly to prediction with a small set of experts. Restarting multiplicative weights at cell changes yields epochwise regret of order $M\sqrt{ET \log V_{\max}}$ over T rounds and E epochs, where losses are bounded by M and each cell has at most $V_{\max}$ vertices. Thus, for fixed loss and geometry parameters, average regret vanishes whenever $E \log V_{\max} = o(T)$. In submodular–concave games, each Lovász cell has only $n+1$ chain vertices; combining the reduction with projected ascent gives order $M\sqrt{ET \log(n+1)} + L_y D_{\mathcal{Y}} \sqrt{ET}$.

1 INTRODUCTION

Convex relaxation makes combinatorial optimization amenable to online learning by replacing a discrete action set with a convex domain (Hazan & Kale, 2012; Hazan, 2016). Yet the relaxed domain may still have exponentially many vertices, so global geometry can obscure simpler local structure. We study a fixed polyhedral partition whose regions, or *cells*, have losses affine on each round; cumulative loss over a cell is therefore minimized at a vertex (Schrijver, 2003).

Submodular optimization provides the motivating example. The Lovász extension is convex on the cube and affine on each simplex induced by a coordinate ordering (Lovász, 1983; Bach, 2013). The cube has 2^n set-indicator vertices, but each simplex has only $n+1$ chain vertices. If the relevant simplex is known, the learner need only compare against these local actions. We ask: *When can regret depend on local-cell complexity rather than on the geometry of the full relaxation?*

We study an online 2-player setting (Cardoso et al., 2025; xin Meng & wei Liu, 2023). The learner acts before observing each objective; the horizon is partitioned into epochs, each associated with one cell. Within an epoch, performance is measured against the best fixed decision in that cell and best fixed opposing action, both chosen in hindsight. This lies between one comparator for the horizon and one that changes every round (Cesa-Bianchi & Lugosi, 2006; Herbster & Warmuth, 1998; Zhao et al., 2024): comparators change only at declared cell boundaries, and regret sums the two players' external regrets over epochs rather than comparing against a new saddle point each round.

The key distinction is between learning within a cell and identifying it. We ask what regret is achievable with supplied cells, if cells can be selected causally from past objectives, and when local learning improves on global convex optimization. A current-equilibrium cell generally depends on the unseen objective, so revealing it before play is an oracle assumption. Restarting projected gradient descent at the same boundaries yields square-root dependence on horizon and epoch count (Zinkevich, 2003; Hazan, 2016); any local advantage must come from a smaller complexity factor.

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.

Under *supplied-cell access*, the learner receives each cell, its vertices, and epoch boundary before play, then observes active-vertex losses after each round. Since the best fixed decision in a cell is a vertex, the problem reduces to expert learning over local vertices (Littlestone & Warmuth, 1994; Cesa-Bianchi & Lugosi, 2006). Restarted multiplicative weights with projected ascent (Freund & Schapire, 1997; Hazan, 2016; Juditsky & Nemirovski, 2011) yields regret scaling as the square root of horizon times epoch count, with logarithmic dependence on the largest cell size. For Lovász relaxations this replaces 2^n global actions by $n+1$ local chain vertices. Epoch lengths need not be known in advance, and for fixed dimension average regret vanishes when the number of epochs is sublinear. Compared with global projected gradient descent, local learning is preferable when its bounded-loss–local-complexity factor is smaller than the global gradient-bound–diameter factor.

We distinguish two access models. Under *oracle access*, the learner receives the cell containing a canonically selected equilibrium of the current objective; regret is controlled by changes in this label. The label can range over all $n!$ coordinate orderings and cannot generally be predicted from past objectives. A causal alternative selects each new cell from a best response to an observed objective. The same analysis controls regret on these causal epochs, but not against the oracle sequence without an additional stability argument. Thus, learning within supplied or causal cells is distinct from reproducing the oracle's cells.

Relation to online learning frameworks. Our analysis uses restart-and-sum reductions for expert advice and online convex optimization (Freund & Schapire, 1997; Littlestone & Warmuth, 1994; Hazan, 2016). Tracking experts and specialists allow changing comparators or availability within a fixed expert framework (Herbster & Warmuth, 1998; Freund et al., 1997); supplied cells instead expose a small geometry-induced expert set. Strongly adaptive methods control all intervals without supplied boundaries but typically retain a global-domain complexity factor (Daniely et al., 2015; Zhang et al., 2018). Appendix F gives details.

2 Background

Submodular–concave games. Let $[n] = \{1, \dots, n\}$. A submodular–concave game consists of a ground set $[n]$, a compact convex set $\mathcal{Y} \subseteq \mathbb{R}^m$, and a payoff function $f\colon 2^{[n]} \times \mathcal{Y} \to \mathbb{R}$ such that $f(\cdot, y)$ is submodular for every $y \in \mathcal{Y}$ and $f(S, \cdot)$ is concave for every $S \subseteq [n]$.

Lovász extension. For $x \in [0,1]^n$ sorted as $x_{\pi(1)} \geq \cdots \geq x_{\pi(n)}$, the Lovász extension is

$$f_L(x, y) = \sum_{i=0}^{n} (x_{\pi(i)} - x_{\pi(i+1)}) f(C_i, y), \tag{1}$$

where $C_0 = \varnothing$, $C_i = \{\pi(1), \dots, \pi(i)\}$, $x_{\pi(0)} = 1$, and $x_{\pi(n+1)} = 0$. For each online objective f_t, we write f_L^t for its Lovász extension. If $f(\cdot, y)$ is submodular, then $f_L(\cdot, y)$ is convex and $\min_{x \in [0,1]^n} f_L(x, y) = \min_{S \subseteq [n]} f(S, y)$ for every $y \in \mathcal{Y}$ (Lovász, 1983). The mixed minimax value equals $V^\star := \min_{x \in [0,1]^n} \max_{y \in \mathcal{Y}} f_L(x, y)$, and an optimal mixed strategy exists with support on a chain of at most $n+1$ nested subsets (Theorem A.3 in Section A.2). For each permutation π of $[n]$, define the *Lovász cell* $\mathcal{P}_\pi = \{x \in [0,1]^n : x_{\pi(1)} \geq \cdots \geq x_{\pi(n)}\}$. These closed simplices cover the cube and overlap on boundary faces. The extension is affine on each cell, but need not be linear because $f(\varnothing, y)$ may be nonzero.

3 Problem Setup

Online protocol and feedback. We study online learning over T rounds and write $[T] = \{1, \dots, T\}$. At round $t \in [T]$, the learner chooses (x_t, y_t) before observing f_t. After play, it receives full access to f_t, including set values and the bounded first-order oracles specified below; revealed objectives remain available for later queries. Losses may depend adaptively on the past, and the bounds hold pathwise for every realized sequence satisfying the assumptions. The oracle result explicitly grants additional current-cell information before play.

Epochwise comparator sequence and regret. The affine-cell reduction requires a comparator that is fixed while a cell is active. Let $\mathcal{Q} = ((I_j, P_j))_{j=1}^{E}$ be a declared sequence of nonempty contiguous

epochs that cover $[T]$, together with closed Lovász cells. The *epochwise sum of external regrets* is

$$\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}) := \sum_{j=1}^{E} \left[\max_{y\in\mathcal{Y}} \sum_{t\in I_j} f_L^t(x_t, y) - \min_{x\in P_j} \sum_{t\in I_j} f_L^t(x, y_t) \right]. \tag{2}$$

Problem studied. For a declared cell sequence $\mathcal{Q}$, we seek an online algorithm whose epochwise regret is sublinear in T and depends on the local cell complexity rather than the global action-space complexity. Our main result addresses the supplied-cell setting, in which every epoch boundary and active cell are available before play. We then consider two distinct instantiations: an equilibrium-cell oracle, which supplies information about the current objective, and a causal rule based only on previously revealed objectives. The causal rule competes against the cells it selects; without an additional stability condition, it does not compete against the oracle's equilibrium-cell sequence.

Assumption 3.1 (Payoffs, geometry, and access)**.** For every t, $f_t(\cdot, y)$ is submodular and $f_t(S, \cdot)$ is continuous and concave on a nonempty compact convex set $\mathcal{Y} \subset \mathbb{R}^m$. Uniformly, $|f_t(S, y)| \le M$, where $M > 0$ is known. At queried points, the first-order oracle supplies a subgradient $g_t^x \in \partial_x f_L^t(x_t, y_t)$ and a supergradient $g_t^y \in \partial_y f_L^t(x_t, y_t)$ satisfying $\|g_t^x\|_2 \le L_x$ and $\|g_t^y\|_2 \le L_y$ for known positive bounds. All norms and projections are Euclidean. Define $D_{\mathcal{Y}} := \sup_{u,v\in\mathcal{Y}} \|u - v\|_2$ and assume exact projection onto $\mathcal{Y}$ is available.

Appendix A.3 discusses their scaling of M, L_x, L_y, and $D_{\mathcal{Y}}$.

4 Cell-Local Epochwise Upper Bounds

The central obstacle is that a global action space can be large even when each stable region has few vertices. We first prove that supplied cells reduce each epoch to local expert learning. We then distinguish two ways to obtain the cells: a before-play equilibrium-label oracle and a causal rule whose benchmark is defined by its own observed cells.

4.1 Supplied-cell guarantee

Theorem 4.1 (Supplied-cell epochwise regret)**.** *Under Assumption 3.1, let $\mathcal{Q}$ consist of E chain-cell epochs, with each cell and boundary supplied before its first play. Cold MW and projected ascent, restarted at each boundary and using the doubling schedule* (5)*, satisfy* $\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}) \le 5M\sqrt{ET\log(n+1)} + 4L_y D_{\mathcal{Y}}\sqrt{ET}$*. The epoch lengths need not be known in advance, and the cells need not contain equilibria.*

Proof. Lemmas C.1 and C.2 bound the two external regrets in each epoch. Their played-payoff terms cancel. Summing over epochs and applying Cauchy–Schwarz proves the claim. □

Thus the local vertex count replaces global geometry once the cell sequence is supplied. To connect this guarantee to the game sequence, we next define cells through canonical equilibrium labels.

4.2 Canonical equilibrium labels

Definition 4.2 (Permutation-switch count)**.** Let $X_t^\star = \arg\min_{x\in[0,1]^n} \max_{y\in\mathcal{Y}} f_L^t(x, y)$. Compactness and continuity ensure that $X_t^\star$ is nonempty and compact. Successively minimize coordinates $x_1, \ldots, x_n$ over the remaining set to select $x_t^\star$. Sort its coordinates in descending order, breaking ties by increasing coordinate index, to obtain $\pi_{x_t^\star}$. Define

$$\mathrm{SC}_T = \left|\{t \in [T-1] : \pi_{x_{t+1}^\star} \ne \pi_{x_t^\star}\}\right|. \tag{3}$$

This quantity is a property of the objective sequence under the stated deterministic convention. The label selects a closed comparator simplex without removing its boundary faces.

The deterministic convention makes the label well defined. Proposition G.3 quantifies its convention sensitivity. Coordinate ordering of a minimax solution need not agree with singleton-marginal ordering, and actions can change within one cell (Appendix E). Canonical labels define a cell sequence, but the current label depends on the current game. We therefore state the corresponding result under explicit before-play oracle access.

Algorithm 1 Geometric WS-CAMW with causal cell queries and doubling

Require: Known $M, L_y, D_{\mathcal{Y}}$; $\alpha \in (0,1)$; $y^0 \in \mathcal{Y}$; past-objective selection oracle
$q \leftarrow p \leftarrow \mathrm{Unif}(\{0,\dots,n\})$; $\pi_{\mathrm{old}} \leftarrow \mathrm{id}$; $v \leftarrow y^0$; $H \leftarrow 1$; $h \leftarrow 0$
for $t = 1, \dots, T$ **do**
  $\pi_t \leftarrow \mathrm{id}$ if $t = 1$; otherwise $\pi_t \leftarrow \pi_{t-1}^{\mathrm{BR}}(v)$, with v the preceding update, before resets
  if $\pi_t \neq \pi_{\mathrm{old}}$ **then**
    Transfer current p from π_{old} to π_t by (9), obtaining q
    $p \leftarrow q$; $v \leftarrow y^0$; $H \leftarrow 1$; $h \leftarrow 0$
  else if $h = H$ **then**
    $H \leftarrow 2H$; $h \leftarrow 0$; $p \leftarrow q$; $v \leftarrow y^0$
  end if
  Set $y_t \leftarrow v$, $x_{t,\pi_t(j)} \leftarrow \sum_{i=j}^{n} p_i$ for $j \in [n]$; play (x_t, y_t) (p_0 is on $C_0 = \varnothing$)
  Observe f_t and g_t^y; set $p_i \leftarrow p_i \exp(-\eta_H f_t(C_i^{\pi_t}, y_t))$, then normalize p
  $v \leftarrow \Pi_{\mathcal{Y}}(y_t + \beta_H g_t^y)$; $h \leftarrow h + 1$; $\pi_{\mathrm{old}} \leftarrow \pi_t$
end for

4.3 Oracle cell access

Before round t, an oracle supplies $\pi_{x_t^\star}$ but no other information about f_t. Let $\mathcal{Q}^\star$ denote the maximal constant-label epochs and their closed cells. Oracle CAMW runs MW on the active chain and projected ascent on $\mathcal{Y}$, restarting both at each boundary under the doubling schedule below.

Corollary 4.3. *Under Assumption 3.1 and the stated before-play cell oracle,*

$$\frac{\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}^\star)}{T} \le O\left(M\sqrt{\frac{(1+\mathrm{SC}_T)\log(n+1)}{T}} + L_y D_{\mathcal{Y}} \sqrt{\frac{1+\mathrm{SC}_T}{T}} \right). \tag{4}$$

This is Theorem 4.1 with $E = 1 + \mathrm{SC}_T$. Proposition G.4 shows that the label alphabet can have entropy $\log_2(n!)$ and that independent labels can force $T(1 - 1/n!)$ expected causal prediction errors. The corollary is therefore an access-conditioned instantiation of the supplied-cell theorem.

Corollary 4.4. *If $\mathrm{SC}_T = 0$, the average epochwise regret is $O((M\sqrt{\log(n+1)} + L_y D_{\mathcal{Y}})/\sqrt{T})$.*

4.4 Causal observed cells and anytime updates

After f_s is revealed, define $x_s^{\mathrm{BR}}(y)$ by successive coordinate minimization over $\arg\min_{x \in [0,1]^n} f_L^s(x, y)$; let $\pi_s^{\mathrm{BR}}(y)$ sort this point with index-based tie breaking. This requires an exact best-response and selection oracle for a past objective; sorting singleton marginal values is not sufficient. The theorem also applies to any other causal permutation rule, with resulting cells as its benchmark. Let v denote the tentative maximizer state used in the cell query and y_t be the action after any reset. At an observed switch, geometric transfer uses the latest post-update weights p, retains their contribution on shared sets, and assigns every new-chain expert the floor in (9).

Doubling schedule. Within each cell epoch, use block capacities $H = 1, 2, 4, \dots$, truncating the last block at the next cell change. Let q be the epoch's initial distribution, uniform for cold CAMW and transferred for geometric WS-CAMW. Restart MW from this same q at every internal block boundary, and restart ascent from a fixed $y^0 \in \mathcal{Y}$. Set $A = \log(n+1)$ for cold CAMW and $A = \log((n+1)/\alpha)$ for geometric WS-CAMW. On a block of capacity H, use

$$\eta_H = \left(\frac{2A}{M^2 H}\right)^{1/2}, \qquad \beta_H = \frac{D_{\mathcal{Y}}}{L_y\sqrt{H}}. \tag{5}$$

Internal block resets do not change the comparator epoch or switch count. These schedules require no future epoch length. For the fixed default $\alpha = 1/2$, the floor is $1/(2(n+1))$ and $A = \log(2(n+1))$. If a shared vertex has old mass at least $\rho > 0$, Lemma B.2 further bounds its initialization KL by $\log(2/\rho)$. The update itself remains tuned to the worst-case value of A.

Theorem 4.5 (Observed-cell upper bound)**.** *Under Assumption 3.1 and the specified past-objective oracle, let $\widehat{\mathcal{Q}}$ be the maximal constant-π_t epochs used by Algorithm 1. Define $\widehat{\mathrm{SC}}_T := |\{t \in [T-1] :$*

$\pi_{t+1} \neq \pi_t\}|$ *and* $\widehat{E} := 1 + \widehat{\mathrm{SC}}_T$. *Then*

$$\frac{\mathrm{Reg}_T^{\mathrm{epoch}}(\widehat{\mathcal{Q}})}{T} \leq O\left(M\sqrt{\frac{\widehat{E}\log((n+1)/\alpha)}{T}} + L_y D_{\mathcal{Y}}\sqrt{\frac{\widehat{E}}{T}}\right). \tag{6}$$

Equivalently, the algorithm certifies total regret at most $5M\sqrt{\widehat{E}T\log((n+1)/\alpha)} + 4L_y D_{\mathcal{Y}}\sqrt{\widehat{E}T}$ on its maximal constant-cell epochs. Comparing $\widehat{\mathcal{Q}}$ with oracle sequence $\mathcal{Q}^\star$ requires a benchmark-transfer bound, and Appendix E.2 shows that maximizer tracking alone is insufficient.

5 Lower Bounds and the Polyhedral Principle

The supplied-cell upper bound has two sources of dependence: the horizon term already present in the model class and the epoch-by-expert term induced by local learning. We separate them because the strongest available lower bounds apply at different levels of generality.

Proposition 5.1 (In-class lower bound with zero switches)**.** *For every $n \geq 1$ and $T \geq 1$, there exists an oblivious distribution over globally modular losses with singleton $\mathcal{Y}$, $|f_t| \leq M$, and $\mathrm{SC}_T = 0$ in every realization such that every causal learner, even with the equilibrium-label oracle, satisfies $\mathbb{E}\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}^\star) \geq \frac{M}{2\sqrt{3}}\sqrt{T}$.*

The construction uses $f_t(S) = M\varepsilon_t|S|/n$ with independent Rademacher signs. The optimum is either 0 or $\mathbf{1}$, and both receive the same label, so the oracle reveals no sign. Appendix G proves the bound. Hence the $\sqrt{T}$ dependence is necessary even when the equilibrium-cell label never changes. We next characterize the additional epoch and expert dependence in the local-expert surrogate.

5.1 The Epoch Dependence Is Sharp for Local Expert Learning

Consider the finite-expert problem induced by a supplied cell, allowing arbitrary bounded losses on its local vertices. For loss vectors $\ell_t \in [-M, M]^K$ and learner distributions $p_t \in \Delta_K$, define

$$R_T^{\mathrm{experts}} = \sum_{j=1}^{E}\left[\sum_{t\in I_j}\langle p_t, \ell_t\rangle - \min_{i\in[K]}\sum_{t\in I_j}\ell_{t,i}\right].$$

Theorem 5.2 (Epochwise local-expert lower bound)**.** *Fix $E \in \{1, \ldots, T\}$ nonempty epochs of as-equal-as-possible lengths and $K \geq 2$ local experts per epoch. For every possibly randomized algorithm, there exists an oblivious distribution over loss sequences $\ell_t \in [-M, M]^K$ such that, with expectation over the losses and the learner's randomization,*

$$\mathbb{E}[R_T^{\mathrm{experts}}] \geq \Omega\Big(M\min\Big\{T, \sqrt{ET\log K}\Big\}\Big). \tag{7}$$

The uncapped square-root expression applies when $\log K \leq T/E$. This result concerns the abstract local-expert benchmark; arbitrary expert losses need not admit a globally submodular realization with a prescribed equilibrium-cell sequence.

Appendix D gives the construction and proof. With Proposition 5.1, this result identifies which parts of the upper bound are established in class and which are established only for the surrogate.

Corollary 5.3. *Corollary 4.3 gives the epochwise submodular–concave upper bound. Theorem 5.2 shows that its $\sqrt{ET\log K}$ statistical form is unavoidable in the broader local-expert surrogate, while Proposition 5.1 establishes in-class necessity of the $\sqrt{T}$ factor.*

5.2 General polyhedral model

The proof of the supplied-cell theorem uses only finite vertices and within-cell affinity. This observation yields the following extension beyond Lovász simplices. Let $\mathcal{X} \subset \mathbb{R}^d$ be nonempty, compact, and convex, and let finitely many nonempty closed compact polytopes $\{\mathcal{R}_a\}_{a=1}^J$ cover $\mathcal{X}$. Their

relative interiors are disjoint, while shared faces are retained. Let V_a be the nonempty vertex set of $\mathcal{R}_a$ and define $V_{\max} := \max_a |V_a|$. Each loss $\ell_t \colon \mathcal{X} \to \mathbb{R}$ is single-valued, affine on every cell, compatible on intersections, and bounded by $|\ell_t(x)| \le M$. Global convexity is not required for the first result. The active label a_t, its vertex list, and each new epoch boundary are supplied before round t; full vertex losses are revealed after play. Epochs are the maximal constant-label intervals, so $E = 1 + \mathrm{RS}_T$, where $\mathrm{RS}_T := |\{t \in [T-1] : a_{t+1} \neq a_t\}|$. Write $a(j)$ for the label on epoch I_j and define $\mathcal{Q} := ((I_j, \mathcal{R}_{a(j)}))_{j=1}^{E}$ and

$$R_T^{\mathrm{poly}}(\mathcal{Q}) = \sum_{j=1}^{E} \left[\sum_{t \in I_j} \ell_t(x_t) - \min_{x \in \mathcal{R}_{a(j)}} \sum_{t \in I_j} \ell_t(x) \right].$$

If labels are instead defined from roundwise optima, a deterministic selector and containing-cell convention must be fixed, and the current-loss-dependent label must be supplied as an oracle.

Theorem 5.4 (Local MW on fixed polytopal cells)**.** *In the preceding model, play convex combinations of the active cell's vertices using cold MW and the doubling schedule* (5)*, with $A = \log |V_{a(j)}|$ and no maximizer. Then $R_T^{\mathrm{poly}}(\mathcal{Q}) \le O\big(M\sqrt{ET \log V_{\max}}\big)$.*

Remark 5.5. The guarantee assumes enumeration and evaluation of active vertices; after access is granted, the update requires no projection (Appendix G.5). A one-simplex instance realizes the $E = 1$ expert lower bound, while matching dependence for arbitrary affine cell sequences remains open.

5.3 Comparison with global projected OGD

Theorem 5.6 (Epochwise upper-bound comparison)**.** *Assume additionally that every ℓ_t is globally convex on $\mathcal{X}$, that a subgradient of norm at most L_x is available at each played point, and that exact Euclidean projection onto $\mathcal{X}$ is available. Let $D_{\mathcal{X}} := \mathrm{diam}_2(\mathcal{X})$. Global OGD, restarted on the same supplied epochs with doubling blocks and step size $D_{\mathcal{X}}/(L_x\sqrt{H})$, satisfies $R_T^{\mathrm{poly}}(\mathcal{Q}) \le O(L_x D_{\mathcal{X}} \sqrt{ET})$. Selecting MW or OGD before play using the known constants gives*

$$\frac{R_T^{\mathrm{poly}}(\mathcal{Q})}{T} \le O\left(\sqrt{\frac{E}{T}} \min\{M\sqrt{\log V_{\max}}, L_x D_{\mathcal{X}}\} \right). \tag{8}$$

With additional feedback at both candidate actions, a two-expert master could add $O(M\sqrt{T})$ to the better base bound. The comparison is determined by the prefactors $M\sqrt{\log V_{\max}}$ and $L_x D_{\mathcal{X}}$, not by E or T. For two-player games, both methods also incur the common ascent term $O(L_y D_{\mathcal{Y}} \sqrt{ET})$.

Shortest paths and flows. For path losses $\ell_t(x) = c_t^\top x$, the path polytope is one valid cell, so $\mathrm{RS}_T = 0$ and $V_{\max} = N$, the number of paths. A bounded flow polytope has the same single-cell interpretation. Changes of the best path or optimal basis are not changes of these decision-space affine cells; a cost-space normal-fan model would require a separate analysis. The historical path study in Appendix F.4 uses prescribed optimal-path changes only as a diagnostic index.

6 Evaluation

6.1 Experimental Setup

Comparison protocol. All methods use the same declared epoch–cell sequence and paired loss streams generated with common random numbers. Cell-aware methods receive declarations before play; no-cell baselines receive only the losses. A common evaluator computes signed, unclipped regret from Eqn. (2). In the main comparisons, each cell contains the epoch's best decision, keeping the restricted comparator meaningful even for methods that can play outside the cell. A random-cell diagnostic shows why this matters: cube-roaming OGD attains regret -51.5 at $T = 4096$ when compared against an unrelated cell. Additional ablations perturb the declarations or replace oracle with causal labels.

Exact comparators. We choose losses that are affine in y over a box and affine in x within each cell. For the unit-box maximizer with $f_L^t(x, y) = \langle y, \phi_t(x) \rangle$, the epoch maximum is $\|\sum_{t \in I_j} \phi_t(x_t)\|_1$.

Table 1: Summary of predictions and measured values. Exponents are OLS slopes in log–log with 95% bootstrap intervals; acceptance was declared before running, except for the row marked †, which was added post hoc. All quantities are computed from (2).

Result	Predicted	Measured	
Thm. 4.1, epochs	$\partial \log \mathrm{Reg}/\partial \log E = 0.500$	$0.495\ [0.489, 0.504]$	✓
Thm. 4.1, horizon, E fixed	0.500	$0.497\ [0.488, 0.503]$	✓
Thm. 4.1, $E = T^{1/4}$	$(1+\theta)/2 = 0.625$	$0.624\ [0.620, 0.631]$	✓
Thm. 4.1, $E = T^{1/2}$	$(1+\theta)/2 = 0.750$	$0.750\ [0.746, 0.752]$	✓
Thm. 5.4, local vertices	$\propto \sqrt{\log V_{\max}}$	CV 0.128 vs. 0.471 for $\sqrt{V_{\max}}$	✓
Thm. 4.1, dimension	$\sqrt{\log(n{+}1)}$: slope 0.118	$0.124\ [0.094, 0.159]$	✓
Thm. 5.6, OGD side	$L_x D_{\mathcal{X}} = \sqrt{n/k}$: slope 0.500	$0.447\ [0.433, 0.464]$	✓
Thm. 5.2, tightness	$\mathrm{Reg} \asymp M\sqrt{ET \log K}$	ratio $\in [1.08, 1.17]$, $E \in [1, 512]$	✓
Prop. 5.1, floor	$\geq \frac{M}{2\sqrt{3}}\sqrt{T}$ for every learner	$0/11$ horizons below, all 5 methods	✓
Prop. G.4, labels	mismatch $\to 1 - 1/n!$	$\lvert\cdot\rvert \leq 1.1 \cdot 10^{-3}$ (n=3), $3.9 \cdot 10^{-7}$ (n=8)	✓
Prop. G.3, convention	$\mathrm{SC}_T : 0 \to T-1$	0 vs. 999 at $T = 1000$	✓
Lem. B.2, transfer	gain increasing in overlap	0.8–1.7%, *no* overlap trend	×
Lem. B.2, persistent†	as above	$2.8\% \to 70.4\%$	✓

The cell-restricted minimum, $\min_{x \in P_j} \sum_{t \in I_j} f_L^t(x, y_t)$, is attained at a vertex and can be computed by enumerating at most $V_{\max}$ vertices (Lemma C.1); singleton-$\mathcal{Y}$ instances require only this minimization. This eliminates iterative-solver error, leaving only floating-point summation error.

Instances. We use two families satisfying Assumption 3.1. *(i) Local-expert / polyhedral* instances (Section 5.2) use an ambient pool of N vertices, cells $P_j = \mathrm{conv}\{e_i : i \in A_j\}$ with $|A_j| = K = V_{\max}$, and losses $\ell_t(x) = \langle z_t, x \rangle$, $z_t \in \{-M, M\}^N$. Varying K at fixed N separates local from ambient complexity. For $M = 1$, one vertex per epoch has mean loss $-\Delta$, where $\Delta = \min\{1, \gamma\sqrt{\log K / T_j}\}$; varying γ spans noise-dominated and strong-signal regimes. *(ii) In-model submodular–concave* instances use the modular family of Proposition G.4, $f_t(S, y) = \frac{1}{n}\sum_i y_i(\mathbb{1}\{i \in S\} - a_{t,i})$. Since its minimax minimizer is $x_t^\star = a_t$, sorting a_t yields the canonical label and exact switch count SC_T, enabling direct comparison of oracle and past-objective labels. A singleton-$\mathcal{Y}$ cube variant removes the ascent term and isolates the minimizer comparison in (8).

Algorithms. Cold MW and geometric WS-MW operate on the active cell using the doubling schedule (5), avoiding advance knowledge of epoch lengths. Restarted global projected OGD uses step $D_{\mathcal{X}}/(L_x\sqrt{H})$ with the same supplied boundaries and doubling blocks. Cell-aware OGD and OLMDA provide gradient-based comparisons, while SAOL and ambient-pool fixed share test tracking without cell access; fixed share receives the switch count but not the cells. Ablations separate the effects of cell restriction, restarts, and weight transfer using ambient-pool MW, boundary restarts with anytime MW or AdaHedge, and warm versus cold starts under matched boundaries. Main learning rates use theoretical constants, with held-out tuning used to assess sensitivity.

Measurements and statistics. We sweep horizon, epoch count, local vertex count, dimension, signal strength, and cell overlap. Each configuration uses 20 seeds, except the zero-switch lower-bound experiment, which uses 200. We report 95% percentile-bootstrap intervals and paired Wilcoxon tests for algorithm contrasts. Scaling exponents are OLS slopes in log–log coordinates. Table 1 distinguishes pre-registered predictions from the post hoc persistent-optimum experiment.

6.2 Experimental Results

Observed regret remains below the supplied-cell bound. Across 122 (configuration, method) pairs comprising 2,440 seed-runs and spanning $T \in [2^{12}, 2^{20}]$, $E \in [1, 1024]$, $V_{\max} \in [2, 1024]$ and $n \in [8, 1024]$, no run exceeded the bound in Theorem 4.1; the worst ratio was 0.430 while family means ranged from 0.04 to 0.31 (Figure 1d). The scaling tests below examine whether this margin reflects conservative constants rather than different parameter dependence.

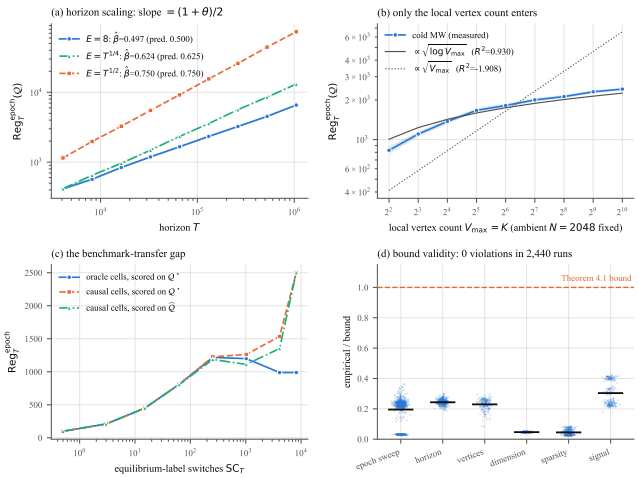

Figure 1: Theorem-aligned measurements of (2). (a) Horizon exponent tracks $(1+\theta)/2$ across three coupled regimes $E = T^{\theta}$, a two-parameter prediction. (b) With ambient vertex pool fixed at $N = 2048$, regret follows $\sqrt{\log V_{\max}}$ and not $\sqrt{V_{\max}}$: only the local count enters. (c) Benchmark-transfer gap: an oracle-cell and a causal-cell learner are scored on the same comparator sequences; the gap is nil while labels are predictable and diverges once they are not. (d) Empirical regret against the Theorem 4.1 bound over all 2,440 runs; bars are means, the worst single run reaches 0.430.

Measured scaling follows predicted epoch, horizon, and local-complexity dependence. Fitted epoch and fixed-epoch horizon exponents are 0.495 and 0.497, with 95% intervals containing the predicted 0.500. With $E = T^{\theta}$, fitted exponents are 0.624 ($\theta = 1/4$) and 0.750 ($\theta = 1/2$), matching predictions 0.625 and 0.750 (Fig. 1a). For the local-expert surrogate (Thm. 5.2), $\mathrm{Reg}/(M\sqrt{ET\log K})$ remains in $[1.08, 1.17]$ over $E \in [1, 512]$, consistent with the rate shared by upper and lower bounds. Finally, with $N = 2048$ and $V_{\max} \in [2, 1024]$, the coefficient of variation is 0.128 for $\mathrm{Reg}/\sqrt{\log V_{\max}}$ vs. 0.471 for $\mathrm{Reg}/\sqrt{V_{\max}}$, supporting dependence on *local* vertex count (Fig. 1b).

Oracle-to-causal transfer incurs a gap as label changes become unpredictable. The exact-comparator protocol also separates learning within cells from selecting their labels. On the family of Prop. G.4, the causal selector's label-mismatch rate differs from $1 - 1/n!$ by only $1.1 \cdot 10^{-3}$ at $n = 3$ and $3.9 \cdot 10^{-7}$ at $n = 8$ (Fig. 3b). The corresponding regret comparison reveals a growing benchmark-transfer gap: at $\mathrm{SC}_T = 0$ the oracle-cell and causal-cell learners are indistinguishable on $\mathcal{Q}^{\star}$ (gap -0.0), at $\mathrm{SC}_T = 63$ the gap is 1.4, at $\mathrm{SC}_T = 8180$ it is 1513.4, a factor of 2.53 (Fig. 1c). The causal learner's regret on its *own* cells $\widehat{\mathcal{Q}}$ stays within (6). These measurements distinguish the guarantee on learner's own cells from performance against the oracle's comparator sequence.

Zero switches retain the predicted $\sqrt{T}$ regret dependence. On the construction of Proposition 5.1, all five tested methods- including two constant strategies and follow-the-leader- have regret above $\frac{M}{2\sqrt{3}}\sqrt{T}$ at all 11 horizons, with fitted exponents in $[0.479, 0.493]$. Thus, for these tested strategies, eliminating label switches does not eliminate the baseline $\sqrt{T}$ dependence (Figure 3a).

Small declaration delays have small measured cost. Delayed declarations provide a setting between exact oracle access and fully causal label selection. Delays of up to 64 rounds (1.6% of an epoch) increase regret by $< 2\%$. Even with a delay of half an epoch, anytime-rate MW attains 3086 versus fixed share's 3125. These results quantify tolerance to delayed boundaries on tested instances.

Boundary-only restarts improve high-signal regret without knowing epoch lengths. In the high-signal regime, fixed share without cell access attains regret 3085 versus 3916 for cold MW ($\gamma = 8$, Fig. 2a). A matched ablation shows that internal doubling restarts discard useful within-epoch information: with the same supplied boundaries, anytime MW and AdaHedge require no epoch lengths and reduce regret to 502 and 526, respectively, both below MW tuned to the true epoch length (1512). AdaHedge is also 5.9× better than fixed share, and under held-out tuning remains below

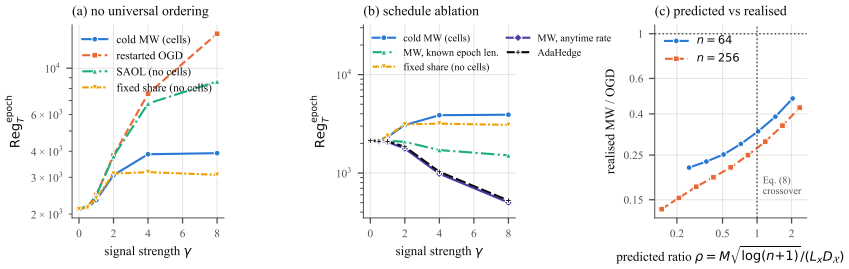

Figure 2: (a) As per-epoch signal γ grows, cold MW on supplied cells overtakes restarted OGD but is crossed by a fixed-share tracker that receives no cells. (b) Matched schedule ablation: boundary-only restarts with anytime MW or AdaHedge require no epoch lengths and achieve lower regret in signal-bearing regimes than doubling-schedule method and MW tuned to known epoch length. (c) Prefactor ratio ρ of (8) predicts direction of MW/OGD comparison but not location of crossing.

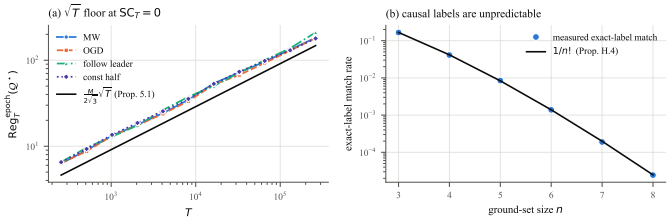

Figure 3: Lower-bound scope. (a) Zero-switch construction of Prop. 5.1: every method stays above $\frac{M}{2\sqrt{3}}\sqrt{T}$ at all 11 horizons; the $\sqrt{T}$ dependence is not an artifact of any one learner. (b) Exact-label match rate of a causal selector on the family of Prop. G.4 follows $1/n!$ over five orders of magnitude.

every tuned no-cell method by $2.6\times$ at $\gamma = 8$ (Fig. 2b). These results identify restart scheduling as an important practical factor in exploiting strong signals.

Prefactor ratio tracks relative regret, but does not locate empirical crossover. Writing $\rho = M\sqrt{\log(n+1)}/(L_x D_{\mathcal{X}})$ for the prefactor ratio in (8), the realized ratio of MW to OGD regret grows as $\rho^{0.41\,[0.34,0.47]}$: the direction of (8) is confirmed, but the realized crossing sits near $\rho \approx 17.7$ rather than $\rho = 1$ (Fig. 2c). In tested Lovász-cube configurations, $\rho \le \sqrt{\log(n+1)} \le 2.4$, so the crossing is not reached and the local method has lower regret throughout that sweep, while a different in-model instance with $L_x D_{\mathcal{X}} = 1$ reverses the ordering. This is a comparison of two upper bounds under a common partition, and the experiments confirm it should not be read as an ordering of the algorithms.

Large transfer gains need persistent useful weights, not overlap alone. With identical supplied boundaries, the pre-registered redrawn-optimum experiment gives only 0.8–1.7% warm-start gains (paired $p < 10^{-3}$), with no increasing trend as overlap ranges from 0 to 1. Thus, overlap alone does not support the pre-registered transfer prediction. In a post hoc persistent-optimum experiment, one-block-per-epoch gains rise from 2.8% (zero overlap) to 70.4% (full overlap), though intermediate values are nonmonotone; with doubling blocks, gains remain 1.3–3.0%. This suggests that substantial transfer requires both persistent useful weights and a schedule that preserves them across epochs.

7 Conclusion

We showed that supplied affine cells reduced epochwise learning to restarted expert advice over local vertices, yielding $O(M\sqrt{ET\log V_{\max}})$ for general cells and $O(M\sqrt{ET\log(n+1)} + L_y D_{\mathcal{Y}}\sqrt{ET})$ for submodular–concave games. We instantiated this reduction with before-play equilibrium labels and with causal observed-cell epochs. We characterized the benchmark, established an in-class $\Omega(M\sqrt{T})$ lower bound at zero switches, and isolated the broader local-expert lower bound. Future work will develop a stable causal cell-identification rule with a provable benchmark-transfer bound.

AI USE STATEMENT

In this work, we used generative AI tools for implementing and running experiments, finding related work, generating synthetic data sets, cleaning and reformatting dataset, and improving the clarity of the manuscript. We have not used generative AI tools for research ideation, assisting with translation, supporting qualitative and thematic data analysis, or interpreting results. Developing theoretical models or conceptual frameworks, formulating mathematical claims, providing critical ingredients for proving mathematical claims, proposing or refining hypotheses are not applicable to this work. Additionally, we used generative AI tools for editing paper to improve readability, identifying relevant literature, formatting references, suggesting a structure for a research paper, proposing a title or keywords for a research paper. We have reviewed all AI-assisted work including code, figures, paper edit suggestions, search results. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work studies regret guarantees for online optimization and evaluates them on synthetic instances; it does not involve human participants, personal data, or a deployed decision system. The methods could support applications such as routing and resource allocation, but similar optimization tools could also be used for targeting, surveillance, or inequitable allocation. Any deployment should therefore assess its objective, data practices, and effects on affected groups in the specific application.

REPRODUCIBILITY STATEMENT

The assumptions, algorithms, and proofs are given in the paper and appendices. The code includes the evaluation harness and commands for regenerating the reported results and figures. Each run uses a separate evaluator to compute Eq. (2) from the played trajectory and declared cells; the synthetic comparators have closed-form solutions, avoiding numerical solver error

REFERENCES

Jean-Yves Audibert, Sébastien Bubeck, and Gábor Lugosi. Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1):31–45, 2014. doi: 10.1287/moor.2013.0598.

Francis Bach. Learning with submodular functions: A convex optimization perspective. *Foundations and Trends in Machine Learning*, 6(2–3):145–373, 2013.

Adrian Rivera Cardoso, He Wang, and Huan Xu. The online saddle point problem and online convex optimization with knapsacks. *Mathematics of Operations Research*, 2025.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Nicolò Cesa-Bianchi and Gábor Lugosi. Combinatorial bandits. *Journal of Computer and System Sciences*, 78(5):1404–1422, 2012. doi: 10.1016/j.jcss.2012.01.001.

Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2018.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23, pp. 6.1–6.20. PMLR, 2012. URL `https://proceedings.mlr.press/v23/chiang12.html`.

Amit Daniely, Alon Gonen, and Shai Shalev-Shwartz. Strongly adaptive online learning. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 1405–1411, 2015.

Amir Ali Farzin, Yuen-Man Pun, Philipp Braun, Tyler Summers, and Iman Shames. Solving the offline and online min-max problem of non-smooth submodular-concave functions: A zeroth-order approach. *arXiv preprint arXiv:2601.21243*, 2026. URL https://arxiv.org/abs/2601.21243.

Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997. doi: 10.1006/jcss.1997.1504.

Yoav Freund, Robert E. Schapire, Yoram Singer, and Manfred K. Warmuth. Using and combining predictors that specialize. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pp. 334–343, 1997.

Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3–4):157–325, 2016. doi: 10.1561/2400000013. URL https://www.nowpublishers.com/article/Details/OPT-013.

Elad Hazan and Satyen Kale. Online submodular minimization. *Journal of Machine Learning Research*, 13:2903–2922, 2012.

Mark Herbster and Manfred K. Warmuth. Tracking the best expert. *Machine Learning*, 32(2):151–178, 1998. doi: 10.1023/A:1007424614876.

Anatoli Juditsky and Arkadi Nemirovski. First-order methods for nonsmooth convex large-scale optimization, I: General purpose methods. In Suvrit Sra, Sebastian Nowozin, and Stephen J. Wright (eds.), *Optimization for Machine Learning*, pp. 121–148. MIT Press, 2011. doi: 10.7551/mitpress/8996.003.0007.

David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146, 2003.

Jure Leskovec and Andrej Krevl. SNAP datasets: Stanford large network dataset collection. https://snap.stanford.edu/data, 2014.

Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994. doi: 10.1006/inco.1994.1009.

László Lovász. Submodular functions and convexity. In *Mathematical Programming: The State of the Art*, pp. 235–257. Springer, 1983. doi: 10.1007/978-3-642-68874-4_10.

Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *Conference on Learning Theory (COLT)*, 2013.

Alexander Schrijver. *Combinatorial Optimization: Polyhedra and Efficiency*. Springer, 2003.

Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958. doi: 10.2140/pjm.1958.8.171.

Bryan Wilder. Equilibrium computation and robust optimization in zero sum games with submodular structure. In *AAAI Conference on Artificial Intelligence*, 2018.

Qing xin Meng and Jian wei Liu. Online saddle point problem and online convex-concave optimization. *arXiv preprint arXiv:2312.06957*, 2023.

Lijun Zhang, Tianbao Yang, Rong Jin, and Zhi-Hua Zhou. Dynamic regret of strongly adaptive methods. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80, pp. 5882–5891. PMLR, 2018. URL https://proceedings.mlr.press/v80/zhang18o.html.

Peng Zhao, Yu-Jie Zhang, Lijun Zhang, and Zhi-Hua Zhou. Adaptivity and non-stationarity: Problem-dependent dynamic regret for online convex optimization. *Journal of Machine Learning Research*, 25, 2024.

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

APPENDIX

A BACKGROUND AND STRUCTURAL FOUNDATIONS

A.1 LANDSCAPE OF ONLINE POLYHEDRAL LEARNING

Table 2 compares upper bounds with their comparator and access conditions; constants are not suppressed.

Table 2: Total-regret upper bounds; no structural minimax tightness is asserted.

Method	Upper bound	Comparator/access
Projected descent–ascent	$(L_x D_{\mathcal{X}} + L_y D_{\mathcal{Y}})\sqrt{T}$, $D_{\mathcal{X}} = \sqrt{n}$	Global static; full information
Cold cell MW + ascent	$O((M\sqrt{\log(n+1)} + L_y D_{\mathcal{Y}})\sqrt{ET})$	Supplied chain epochs
Geometric WS-CAMW	$O((M\sqrt{\log((n+1)/\alpha)} + L_y D_{\mathcal{Y}})\sqrt{\widehat{E}T})$	Own causal cells only

A.2 BACKGROUND ON STRUCTURAL PROPERTIES OF THE LOVÁSZ EXTENSION

We present the following structural results from Lovász (1983); Schrijver (2003) for completeness.

Lemma A.1 (Lovász extension: convexity and tight relaxation)**.** *Let $f\colon 2^{[n]} \to \mathbb{R}$ be submodular. Then the Lovász extension $f_L\colon [0,1]^n \to \mathbb{R}$ defined in* (1) *satisfies: (1) f_L is convex and (2) $\min_{x\in[0,1]^n} f_L(x) = \min_{S\subseteq[n]} f(S)$.*

Lemma A.2 (Mixed minimax value via threshold coupling)**.** *Consider a submodular-concave game $(f, [n], \mathcal{Y})$ with continuous payoffs and compact convex $\mathcal{Y}$ as in Assumption 3.1. Define the* mixed value $V^{\mathrm{mix}} := \min_{\mu\in\Delta(2^{[n]})} \max_{y\in\mathcal{Y}} \mathbb{E}_{S\sim\mu}[f(S,y)]$ *and the* relaxed value $V^\star := \min_{x\in[0,1]^n} \max_{y\in\mathcal{Y}} f_L(x,y)$. *Then $V^\star = V^{\mathrm{mix}}$.*

Lemma A.3 (Chain support structure)**.** *Under Assumption 3.1, there exist $x^\star \in \arg\min_{x\in[0,1]^n} \max_{y\in\mathcal{Y}} f_L(x,y)$, a chain $\emptyset = C_0 \subset C_1 \subset \cdots \subset C_K = [n]$ with $K \le n$, and weights $w \in \Delta_{K+1}$ such that $f_L(x^\star, y) = \sum_{j=0}^K w_j f(C_j, y)$ for every $y \in \mathcal{Y}$. The distribution assigning mass w_j to C_j is an optimal mixed strategy.*

Corollary A.4 (Mixed equilibrium existence)**.** *Under Assumption 3.1, every game admits a saddle point $(x^\star, y^\star) \in [0,1]^n \times \mathcal{Y}$ with $f_L(x^\star, y) \le V^\star \le f_L(x, y^\star)$ for all $x \in [0,1]^n$, $y \in \mathcal{Y}$. Moreover, the threshold distribution representing $x^\star$ is supported on a chain of at most $n+1$ subsets, and $V^\star = V^{\mathrm{mix}}$.*

Justification of the structural lemmas. For fixed y, the Lovász extension of the normalized submodular function $f(\cdot, y) - f(\varnothing, y)$ is the support function of its base polytope (Lovász, 1983; Schrijver, 2003); adding back $f(\varnothing, y)$ proves convexity. Equation (1) expresses every value as a convex combination of set values, so $f_L(x,y) \ge \min_S f(S,y)$, and equality of the two minima follows by including all set indicators in the cube.

For any distribution μ on sets with marginals $x = \mathbb{E}_\mu \mathbb{1}_S$, Jensen's inequality gives $f_L(x,y) \le \mathbb{E}_\mu f(S,y)$ for all y, hence $V^\star \le V^{\mathrm{mix}}$. Conversely, draw U uniformly on $[0,1]$ and take $S_U = \{i : x_i \ge U\}$. The sets form one chain, and their probabilities are exactly the nonnegative differences in (1), summing to one. Therefore $\mathbb{E} f(S_U, y) = f_L(x,y)$ simultaneously for all y. Applied to a minimax $x^\star$, this gives $V^{\mathrm{mix}} \le V^\star$ and the chain-support existence statement. Continuity and compactness give attainment, while convexity in x and concavity in y permit Sion's minimax theorem (Sion, 1958), yielding a saddle point. Not every optimal distribution need have chain support. □

A.3 ASSUMPTION SCALING ACROSS DOMAINS

A comparison must keep M, L_x, L_y, the Euclidean cube diameter $\sqrt{n}$, and $D_{\mathcal{Y}}$ visible. The minimizer prefactors are $M\sqrt{\log(n+1)}$ and $L_x\sqrt{n}$. For example, the upper scalings $M = O(n)$ and $L_x =$

$O(\sqrt{n})$ yield $O(n\sqrt{\log(n+1)})$ and $O(n)$, respectively. Under the additional normalization $M = 1$ and $L_x\sqrt{n} = \Theta(\sqrt{n})$, their ratio is $\Theta(\sqrt{n/\log(n+1)})$. Thus the preferred upper bound depends on explicit problem scales and the shared ascent term; no universal ordering follows.

B Algorithms and Implementation

B.1 Euclidean OLMDA baseline

With subgradient g_t^x and supergradient g_t^y from Assumption 3.1, play first, then update

$$x_{t+1} = \Pi_{[0,1]^n}(x_t - \eta_x g_t^x), \qquad y_{t+1} = \Pi_{\mathcal{Y}}(y_t + \eta_y g_t^y).$$

This is the Euclidean specialization of mirror descent–ascent. A Lovász x-subgradient can be computed by sorting x and evaluating its $n+1$ chain sets; this is not a best-response computation.

Theorem B.1 (Global-static external regret)**.** *Define*

$$R_T^{\text{static}} = \max_{y\in\mathcal{Y}} \sum_t f_L^t(x_t, y) - \min_{x\in[0,1]^n} \sum_t f_L^t(x, y_t).$$

For known T, the preceding updates with $\eta_x = \sqrt{n}/(L_x\sqrt{T})$ and $\eta_y = D_{\mathcal{Y}}/(L_y\sqrt{T})$ satisfy $R_T^{\text{static}} \le (L_x\sqrt{n} + L_y D_{\mathcal{Y}})\sqrt{T}$.

Proof. Projection nonexpansiveness gives, for any fixed $u \in [0,1]^n$,

$$\sum_t \langle g_t^x, x_t - u\rangle \le \frac{n}{2\eta_x} + \frac{\eta_x T L_x^2}{2}.$$

The corresponding ascent inequality is $\sum_t \langle g_t^y, v - y_t\rangle \le D_{\mathcal{Y}}^2/(2\eta_y) + \eta_y T L_y^2/2$ for fixed $v \in \mathcal{Y}$. Convexity and concavity bound the two external regrets by these quantities. Optimize over u, v, cancel realized payoffs, and substitute the rates. Singleton $\mathcal{Y}$ contributes zero. □

The global-static benchmark differs from (2): it uses one fixed action per player over the entire horizon. Appendix H summarizes the scope of the guarantees.

B.2 Geometric Chain-Overlap Transfer

Geometric WS-CAMW (Algorithm 1) replaces uniform initialization at a cell change with a transfer retaining prior mass contributions on shared chain elements; shrinkage and normalization do not preserve exact weights. We state and prove the initialization guarantee used in Theorem 4.5.

Lemma B.2 (Geometric transfer initialization)**.** *Let w be the MW weight distribution at the end of epoch j over chain $\mathcal{C}^{\pi_j}$, and let $\mathcal{S} = \{k : C_k^{\pi_j} = C_k^{\pi_{j+1}}\}$ be the shared chain elements with total weight $p = \sum_{k\in\mathcal{S}} w_k$. Under geometric transfer with parameter $\alpha \in (0,1)$, the transferred distribution*

$$\tilde{w}_k = \begin{cases} (1-\alpha)\, w_k + \alpha/(n+1) & \text{if } k \in \mathcal{S}, \\ \alpha/(n+1) & \text{if } k \notin \mathcal{S}, \end{cases} \qquad w'_k = \tilde{w}_k \Big/ \sum_\ell \tilde{w}_\ell, \tag{9}$$

satisfies:

(a) $w'_k \ge \alpha/(n+1)$ *for every k, so $D_{\mathrm{KL}}(e_{i^\star}\|w') \le \log((n+1)/\alpha)$ for any comparator $i^\star$.*

(b) *If $i^\star \in \mathcal{S}$ with prior weight $w_{i^\star}$, then $w'_{i^\star} \ge (1-\alpha)\, w_{i^\star} + \alpha/(n+1)$, and the initialization KL is*

$$D_{\mathrm{KL}}(e_{i^\star}\|w') \le \log\big(1/\big((1-\alpha)\, w_{i^\star} + \alpha/(n+1)\big)\big).$$

When the shared expert has large prior mass, this can be $O(1)$ rather than $O(\log n)$.

Proof. **Part (a).** The normalization constant is $Z = \sum_k \tilde{w}_k = (1-\alpha)\, p + \alpha \le 1$, since $p \le 1$. Thus $w'_k = \tilde{w}_k / Z \ge \tilde{w}_k \ge \alpha/(n+1)$ for all k. The KL bound follows: $D_{\mathrm{KL}}(e_{i^\star} \| w') = \log(1/w'_{i^\star}) \le \log((n+1)/\alpha)$.

Part (b). For $i^\star \in \mathcal{S}$: $\tilde{w}_{i^\star} = (1-\alpha)\, w_{i^\star} + \alpha/(n+1)$. Since $Z \le 1$, $w'_{i^\star} = \tilde{w}_{i^\star}/Z \ge \tilde{w}_{i^\star}$, giving the stated bound. The guarantee therefore depends on retained probability mass rather than on the number of shared chain elements. □

Remark B.3 (Scope of transfer guarantees). At a fixed learning rate, the initialization term is $\log(1/q_i)/\eta$. Lemma B.2 controls this term, while Algorithm 1 tunes η using the worst-case floor. Comparator-adaptive rates and the historical Jaccard transfer require separate analyses.

C Proofs: Upper Bound

C.1 Fixed-comparator block and epoch bounds

Lemma C.1 (Within-epoch MW bound)**.** *Fix a closed chain cell with $K \ge 2$ vertices, losses $\ell_{t,i} = f_t(C_i, y_t) \in [-M, M]$, and epoch prior q satisfying $\log(1/q_i) \le A$ for all i. The doubling MW updates with $\eta_H = \sqrt{2A/(M^2 H)}$ and prior q in every block give*

$$\sum_{t \in I_j} f_L^t(x_t, y_t) - \min_{x \in P_j} \sum_{t \in I_j} f_L^t(x, y_t) \le 5M\sqrt{AT_j}. \tag{10}$$

For a singleton chain the minimizer regret is zero.

Proof. Affineness on the closed simplex gives $f_L^t(x_t, y_t) = \sum_i p_{t,i} \ell_{t,i}$. The exponential-weights potential and Hoeffding's lemma for losses in $[-M, M]$ give, in any block of length $h \le H$, regret to every vertex i at most

$$\frac{\log(1/q_i)}{\eta_H} + \frac{\eta_H M^2 h}{2} \le M\sqrt{2AH}.$$

This inequality holds pathwise even for adaptively generated losses. Apply it to the same vertex throughout the epoch. For capacities $1, 2, \ldots, 2^b$, including a truncated last block, $2^b \le T_j$ and $\sum_{k=0}^{b} \sqrt{2^k} \le (2+\sqrt{2})\sqrt{T_j}$. Thus the constant can be bounded by 5. Finally the cumulative loss is affine on the compact cell, so its minimum is attained at one of the chain vertices. No comparison to a per-round game value is used (Freund & Schapire, 1997). □

Lemma C.2 (Within-epoch projected-ascent bound)**.** *With the same blocks and projected ascent restarted at y^0 in each block,*

$$\max_{y \in \mathcal{Y}} \sum_{t \in I_j} f_L^t(x_t, y) - \sum_{t \in I_j} f_L^t(x_t, y_t) \le 4 L_y D_{\mathcal{Y}} \sqrt{T_j}. \tag{11}$$

Proof. For any fixed y, concavity bounds each payoff difference by $\langle g_t^y, y - y_t \rangle$. Projection gives block regret at most $D_{\mathcal{Y}}^2/(2\beta_H) + \beta_H L_y^2 h/2 \le L_y D_{\mathcal{Y}} \sqrt{H}$. Sum the geometric series of square roots as above and then maximize over the same y for the entire epoch. If $\mathcal{Y}$ is a singleton, the left side is zero. □

Lemma C.3 (Epoch aggregation via Cauchy–Schwarz)**.** *Suppose E epochs of lengths $T_1, \ldots, T_E$ with $\sum_j T_j = T$ each contribute regret $a\sqrt{T_j}$ for some constant $a > 0$. Then*

$$\sum_{j=1}^{E} a\sqrt{T_j} \le a\sqrt{ET}. \tag{12}$$

Proof. By Cauchy–Schwarz, $\sum_{j=1}^{E} \sqrt{T_j} \le \sqrt{E \sum_j T_j} = \sqrt{ET}$. □

C.2 Proof of Corollary 4.3

We prove the bound for oracle CAMW with the current equilibrium label supplied before play. This gives an oracle upper bound for the epochwise fixed-comparator benchmark as a function of the permutation-switch complexity SC_T.

Proof strategy. We decompose the horizon into $E = \mathrm{SC}_T + 1$ epochs. Within each epoch, MW competes with the best fixed chain vertex under the realized maximizer actions, while restarted mirror ascent competes with the best fixed maximizer action. Adding these two no-regret inequalities gives exactly the epochwise sum of external regrets in (2); Cauchy–Schwarz then aggregates the epoch costs.

Proof. **Epoch decomposition.** Let $1 = t_1 < t_2 < \cdots < t_{S+1} < t_{S+2} = T + 1$ be the epoch boundaries induced by the $S = \mathrm{SC}_T$ selected-equilibrium permutation changes. Then $I_j = \{t_j, \dots, t_{j+1} - 1\}$, $j = 1, \dots, S + 1$, is the oracle epoch partition, using the deterministic convention in Definition 4.2.

Within-epoch structure. Fix an epoch j of length T_j. During this epoch, the equilibrium permutation $\pi^{(j)}$ is fixed. The Lovász extension is affine on this closed coordinate-order simplex, whose associated chain is

$$C_0^{(j)} \subset C_1^{(j)} \subset \cdots \subset C_n^{(j)}.$$

Thus the minimizer's problem reduces to online linear optimization over the $n+1$ vertices $\{C_i^{(j)}\}_{i=0}^n$.

Within-epoch external regrets. With $c_x = 5$ and $c_y = 4$, by Lemmas C.1 and C.2, and because the minimum of an affine function over P_j is attained at a chain vertex,

$$\begin{aligned}\max_{y \in \mathcal{Y}} \sum_{t \in I_j} f_L^t(x_t, y) - \min_{x \in P_j} \sum_{t \in I_j} f_L^t(x, y_t) \\ \le c_x M \sqrt{T_j \log(n+1)} + c_y L_y D_{\mathcal{Y}} \sqrt{T_j}. \end{aligned} \tag{13}$$

The two middle terms $\sum_{t \in I_j} f_L^t(x_t, y_t)$ cancel when the minimizer and maximizer inequalities are added. This is the precise reason for using the epochwise fixed-comparator benchmark.

Aggregation across epochs. Summing (13) over the $E = S + 1$ epochs and applying Lemma C.3 separately to both terms gives

$$\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}^\star) \le O\Big(M \sqrt{(1 + \mathrm{SC}_T) T \log(n+1)} + L_y D_{\mathcal{Y}} \sqrt{(1 + \mathrm{SC}_T) T} \Big). \tag{14}$$

Dividing by T proves (4). □

C.3 Proof of Theorem 4.5

Proof. Use $c_x = 5, c_y = 4$ as above. The observed switches induce $\widehat{E} = 1 + \widehat{\mathrm{SC}}_T$ epochs. Within epoch j, the observed permutation is fixed, so the minimizer operates on a fixed chain $\mathcal{C}^{(j)}$. The transfer rule guarantees every new-epoch expert initial mass at least $\alpha/(n+1)$; hence the anytime MW analysis gives

$$\sum_{t \in I_j} f_L^t(x_t, y_t) - \min_{x \in P_j} \sum_{t \in I_j} f_L^t(x, y_t) \le c_x M \sqrt{T_j \log((n+1)/\alpha)}. \tag{15}$$

Because mirror ascent is also restarted at each observed boundary, Lemma C.2 gives

$$\max_{y \in \mathcal{Y}} \sum_{t \in I_j} f_L^t(x_t, y) - \sum_{t \in I_j} f_L^t(x_t, y_t) \le c_y L_y D_{\mathcal{Y}} \sqrt{T_j}. \tag{16}$$

Adding the two inequalities yields the epoch-j summand in (2). Summing over epochs and applying Cauchy–Schwarz proves (6). In particular, the maximizer term also scales with $\sqrt{\widehat{E}}$ because the benchmark permits a different fixed maximizer comparator in every epoch.

The transferred prior is reused at internal doubling boundaries, so the same worst-case initialization bound applies in every block. Only the initialization-KL refinement of Lemma B.2(b) is claimed; the schedule is not tuned to an unknown comparator-specific KL. □

C.4 Chain restriction and advance-known union pools

Theorem C.4 (Chain-restricted epochwise upper bound)**.** *Under Assumption 3.1, for a supplied sequence $\mathcal{Q}$ of E convex hulls of chains, each with at most K vertices, cold restarted MW and ascent give*

$$\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}) \le O(M\sqrt{ET\log K} + L_y D_{\mathcal{Y}}\sqrt{ET}).$$

Proof. Apply Lemmas C.1 and C.2, cancel realized payoffs, and sum using Cauchy–Schwarz. Singleton cells contribute zero on the minimizer side. □

Union alternative and its access model. Suppose the union of the E chains is supplied in advance, with $U \le EK$ distinct set-indicator vertices, and the same epoch boundaries are known before play. Restart MW over the entire union at each boundary, allowing weight on inactive-cell vertices and querying all U losses. The union is not a chain. Nevertheless convexity gives

$$f_L^t\left(\sum_i p_{t,i} v_i, y_t\right) \le \sum_i p_{t,i} f_L^t(v_i, y_t).$$

MW's loss therefore upper-bounds the played loss. Competing with the best fixed active-chain vertex yields minimizer regret $O(M\sqrt{T_j \log U})$; ascent cancels the actual played payoff as before. Thus the same benchmark has upper bound $O(M\sqrt{ET\log U} + L_y D_{\mathcal{Y}}\sqrt{ET})$, although these plays need not lie in the active cell. Thus advance knowledge of the union recovers logarithmic dependence on its size at the cost of querying all union vertices and allowing plays outside the active cell. The local algorithm needs only the current chain.

D Proofs: Lower-Bound Scope

D.1 A Modular Construction Within the Model Class

For fixed $A, B \subseteq [n]$, bounded coefficients $\varepsilon_{t,i}$, and a submodular h, the example

$$f_t(S, y) = y \sum_{i \in S \cap A} \varepsilon_{t,i} + h(S \cap B), \qquad \mathcal{Y} = [-1, 1], \tag{17}$$

is globally submodular in S (a modular term plus a restricted submodular function) and affine in y. This verifies admissibility only; it neither proves a regret lower bound nor prescribes equilibrium switches. When $h = 0$, $x = 0$ and $y = 0$ already neutralize the bilinear term.

D.2 Proof of Theorem 5.2

Proof. Let E be the number of epochs and let their as-equal-as-possible lengths be $T_1, \dots, T_E$, so each $T_j \in \{\lfloor T/E \rfloor, \lceil T/E \rceil\}$. Let R_j denote the jth summand of R_T^{experts}, and let $\mathcal{F}_{j-1}$ be the history through epoch $j-1$. In epoch j, draw a fresh standard hard distribution for prediction with K experts and losses in $[-M, M]$. The classical full-information experts lower bound and Yao's principle give

$$\mathbb{E}[R_j \mid \mathcal{F}_{j-1}] \ge cM \min\{T_j, \sqrt{T_j \log K}\} \tag{18}$$

for a universal constant $c > 0$ (Cesa-Bianchi & Lugosi, 2006). Fresh independence across epochs makes the bound valid even for algorithms that transfer state between epochs. Summing the conditional bounds yields

$$\mathbb{E}[R_T^{\mathrm{experts}}] \ge cM \sum_{j=1}^{E} \min\{T_j, \sqrt{T_j \log K}\}. \tag{19}$$

If $\log K \le T/E$, every epoch lies in the square-root regime up to universal rounding constants, and the balanced lengths give $\sum_j \sqrt{T_j} = \Theta(\sqrt{ET})$. Otherwise the capped experts lower bound itself supplies a linear-order lower bound in every balanced epoch, up to universal constants. Combining the two cases proves the stated $\Omega(M \min\{T, \sqrt{ET \log K}\})$ bound.

This argument assigns arbitrary losses directly to the K local experts. It does not extend those values to a globally submodular set function, nor does it prescribe an equilibrium-cell switch count. Consequently, it supports only the surrogate lower-bound statement in Theorem 5.2. □

D.3 A Vertex-Loss Surrogate Lower Bound

Proposition D.1 (Vertex-loss surrogate lower bound)**.** *Suppose $V_{\max} \geq 2$ and a regional feedback model permits arbitrary losses in $[-M, M]$ on the $V_{\max}$ active vertices in each of E balanced epochs, without requiring an affine extension or agreement across shared facets. For every possibly randomized learner, there exists an oblivious distribution over such vertex-loss sequences for which*

$$\mathbb{E}[R_T^{\text{experts}}] \geq \Omega\Big(M \min\{T, \sqrt{ET \log V_{\max}}\}\Big). \tag{20}$$

This surrogate result is not a lower bound for the affine-per-region loss class in Theorem 5.4.

Proof. Apply Theorem 5.2 with $K = V_{\max}$. The qualification is essential: an affine function on a d-dimensional region has at most $d + 1$ degrees of freedom, so arbitrary independent values cannot generally be assigned to all vertices when $V_{\max} > d + 1$. Moreover, independently specified affine pieces need not be continuous on shared facets. A vertexwise Rademacher construction therefore does not establish a general polyhedral OCO lower bound. □

D.4 Proof of Theorem 5.4

Proof. Let $s = \mathrm{RS}_T$ and let $1 = t_1 < \cdots < t_{s+1} < t_{s+2} = T + 1$ be the boundary points of the $s + 1$ epochs. In epoch I_j, the comparator is restricted to the fixed active region $\mathcal{R}_{a(j)}$. Because each loss is affine on this region, the best fixed comparator in hindsight is attained at a vertex of $\mathcal{R}_{a(j)}$. Anytime MW over at most $V_{\max}$ vertices therefore gives

$$\sum_{t \in I_j} \ell_t(x_t) - \min_{x \in \mathcal{R}_{a(j)}} \sum_{t \in I_j} \ell_t(x) \leq cM\sqrt{T_j \log V_{\max}}. \tag{21}$$

Summing over epochs and applying Cauchy–Schwarz yields

$$R_T^{\text{poly}}(\mathcal{Q}) \leq cM\sqrt{(1 + \mathrm{RS}_T)T \log V_{\max}}, \tag{22}$$

which proves Theorem 5.4. Proposition D.1 applies only to the less constrained vertex-loss surrogate and is not invoked as a matching lower bound here. □

D.5 Proof of the epochwise upper-bound comparison

Proof. The MW branch is Theorem 5.4. Under the additional global convexity and projection assumptions, projected OGD in a block has regret at most $L_x D_{\mathcal{X}} \sqrt{H}$ against every fixed point of $\mathcal{X}$, hence also the epoch's best point in its declared cell. Summing doubling blocks and then epochs gives $O(L_x D_{\mathcal{X}} \sqrt{ET})$. Choose the algorithm with the smaller known displayed prefactor before play. Both bounds refer to the same externally supplied comparator sequence; neither provides a structural lower bound or a crossover in E. □

E Separations Between Cell Stability and Dynamic Comparators

E.1 Stable cell, changing optimal action

Proposition E.1 (Cell stability does not control dynamic regret)**.** *There are globally modular losses on $n = 2$ with singleton $\mathcal{Y}$ and $|f_t| \leq 2$ such that $\mathrm{SC}_T = 0$ but the selected optimal action changes $T - 1$ times. Moreover, an oblivious random sequence with $\mathrm{SC}_T = 0$ forces expected regret at least $T/2$ to the roundwise optimum for every causal learner without current-loss side information.*

Proof. Let $f_t(S) = -\mathbb{1}\{1 \in S\} + \varepsilon_t \mathbb{1}\{2 \in S\}$, with $\varepsilon_t \in \{-1, +1\}$. The unique minimizer is $(1, 0)$ for $\varepsilon_t = +1$ and $(1, 1)$ for $\varepsilon_t = -1$. Both have index-tie-broken descending permutation $(1, 2)$. Alternating signs prove the first claim. For the lower bound, choose independent unbiased signs. Any causal $x_t \in [0, 1]^2$ has expected loss $\mathbb{E}[-x_{t,1} + \varepsilon_t x_{t,2}] = -\mathbb{E}x_{t,1} \geq -1$, whereas the expected roundwise minimum is $-3/2$. Sum over t. Every loss is modular on all subsets; no extension from arbitrary chain values is needed. □

For the same deterministic selected-action sequence, SC_T is at most its action-switch count. The construction therefore separates equilibrium-cell switches from roundwise action switches; it does not compare SC_T with path length or gradient variation.

E.2 Why maximizer tracking is insufficient

Take $c > 0$, $\mathcal{Y} = [-1, 1]$, and the modular–affine game

$$f(S, y) = c\mathbb{1}\{1 \in S\} + y(\mathbb{1}\{2 \in S\} - 1/2), \qquad f_L(x, y) = cx_1 + y(x_2 - 1/2).$$

Its minimax objective is $cx_1 + |x_2 - 1/2|$, so $x^\star = (0, 1/2)$ is unique and has permutation $(2, 1)$. The dual optimum is $y^\star = 0$. At $y^\star$, successive-coordinate minimization of the best-response set selects $(0, 0)$ with permutation $(1, 2)$. The singleton marginal scores are c and 0, strictly separated. Hence perfect maximizer tracking and distinct marginal scores do not identify the equilibrium cell.

Even if synchronous label sequences differ on only m rounds, each discrepancy can affect two adjacent transitions, so their switch counts differ by at most $2m$, not necessarily m. A lagged detector has further initialization and endpoint issues. Count agreement alone still does not align comparator cells or partitions. A bridge from maximizer tracking to equilibrium-cell regret therefore requires a stable selection rule, control of lagged cell errors, and a benchmark-discrepancy bound.

E.3 What remains of structural comparisons

The ratio of displayed minimizer prefactors is $L_x D_{\mathcal{X}} / (M\sqrt{\log V_{\max}})$ for $V_{\max} > 1$. Upper big-O assumptions alone do not give a two-sided ratio: in particular $M = O(1)$ supplies no lower bound on its denominator. No exponential improvement or actual regret ordering is inferred.

Proposition E.2 (Declared constant-label regimes)**.** *If the canonical equilibrium label is constant within each of k declared contiguous regimes, then $\mathrm{SC}_T \le k - 1$ and the oracle bound is $O(M\sqrt{kT\log(n+1)} + L_y D_{\mathcal{Y}}\sqrt{kT})$.*

Proof. Only the $k - 1$ regime boundaries can change the label; substitute into Corollary 4.3. □

This assumes label stability rather than deriving it from small payoff drift.

F Related Work and Positioning

Related work falls into four categories: *global online convex optimization*, *online saddle-point learning*, *tracking and interval adaptivity*, and *combinatorial oracle models*.

Online convex and submodular optimization. Projected OGD and mirror descent express regret through gradient norms and domain geometry (Zinkevich, 2003; Hazan, 2016; Juditsky & Nemirovski, 2011). The Lovász extension connects submodular minimization to convex optimization (Lovász, 1983; Bach, 2013; Hazan & Kale, 2012). Under supplied cell and vertex access, our result replaces global-domain geometry by a dependence on $\log(n + 1)$ for Lovász cells. Smooth relaxations for submodular maximization (Chen et al., 2018; Wilder, 2018) address different objectives and feasible sets.

External regret and online saddle points. Online saddle-point formulations use several comparator and evaluation criteria (Cardoso et al., 2025; xin Meng & wei Liu, 2023; Farzin et al., 2026). Our criterion adds the two external regrets along the realized opponent actions within each epoch. Because objectives vary with time, the resulting sum is signed and differs from fixed-game exploitability. A direct comparison with the zeroth-order method of Farzin et al. (2026) must therefore account for both the criterion and the access model: their function queries versus our exact set values, supergradients, projections, and cell-selection operations.

Tracking, specialists, and strong adaptivity. Fixed-Share tracks changing comparators in a fixed expert universe (Herbster & Warmuth, 1998); specialists model changing availability (Freund et al., 1997); strongly adaptive methods provide simultaneous interval guarantees (Daniely et al., 2015; Zhang et al., 2018). Supplying a cell instead reveals a small local expert set. Proposition G.5 states how simultaneous interval bounds imply Eq. (2) for every external partition while retaining their interval overhead, global-domain prefactor, and feedback assumptions. For Fixed-Share on all subsets, the universe has size 2^n and hence $\log N = n$; a smaller union requires advance availability and additional value queries (Appendix C.4). Geometric transfer provides a prior based on shared chain elements and controls its initialization KL.

Variation and predictability. Path-length, gradual-variation, and predictable-sequence methods exploit other forms of temporal structure (Chiang et al., 2012; Rakhlin & Sridharan, 2013; Zhang et al., 2018). An equilibrium-cell label alone does not control motion within a cell, as the modular construction in Appendix E demonstrates. These measures can therefore complement cell stability when additional sensitivity assumptions connect payoff variation to the selected equilibrium.

Polyhedral access and computation. Combinatorial online learning (Cesa-Bianchi & Lugosi, 2012; Audibert et al., 2014) and projection-free optimization often use oracle models that do not enumerate a local vertex set. Our MW guarantee assumes that the active vertices and their losses are available. In shortest-path problems, paths are vertices of one globally linear decision-space model, whereas the cost-space normal fan is a different geometric object. Appendix G.5 makes the corresponding operation costs explicit.

The main results are theoretical. We retain the archived measurements below to document prior behavior under their original metrics, then state precisely why they do not constitute a test of the supplied-cell theorem.

F.1 Historical Experimental Diagnostics

The archived experiments measure roundwise chain-based diagnostics under a protocol different from Algorithm 1. We report them as descriptive evidence, with their metrics, source bundles, and implementation conditions stated explicitly. Section 6 gives the protocol for evaluating Eq. (2).

F.1.1 Metrics and controlled synthetic instances

Let $\mathcal{C}_t^{\mathrm{score}}(y_t)$ be the chain obtained by sorting singleton marginal scores of $f_t(\cdot, y_t)$ in increasing order, and let

$$m_t^{\mathrm{chain}} = \min_{C \in \mathcal{C}_t^{\mathrm{score}}(y_t)} f_t(C, y_t).$$

The saved one-sided diagnostic and approximate two-sided diagnostic are

$$\bar{d}_T^x = \frac{1}{T}\sum_t [f_L^t(x_t, y_t) - m_t^{\mathrm{chain}}], \tag{23}$$

$$\bar{d}_T^{xy} = \frac{1}{T}\sum_t [f_L^t(x_t, \widetilde{y}_t) - m_t^{\mathrm{chain}}].$$

Here $\widetilde{y}_t$ is the output of a single midpoint-initialized L-BFGS-B maximization with at most 15 iterations, without an optimality certificate. The chain minimum is not generally the full-domain minimum. Both optimizations are roundwise, not fixed within epochs. The stored field `final_avg_minimizer_static_regret` means $\bar{d}_T^x$, despite its name; `final_avg_regret` aliases the approximate two-sided field, or $\bar{d}_T^x$ when `skip_duality_gap=True`. Neither is the revised benchmark or a certified exploitability measure.

The controlled generator uses

$$f_t(S, y) = \frac{\sum_{i \in S}(a_{t,i} + b_{t,i}^\top y) - \sum_{i<j,\, i,j \in S} w_{ij}}{\widehat{M}_t}, \qquad w_{ij} \geq 0.$$

This is genuinely submodular in S and affine in y. Its normalization $\widehat{M}_t$ is a target-chain range at $y = 0$, not a certified uniform bound over all sets and y. Designed changes concern singleton-score order, not solved equilibrium cells. We denote the designed count by S_{design}, the runtime accepted-label count by S_{obs}, and the offline score-order count by S_{score}; none is identified with intrinsic SC_T. In some continuous baselines the stored observed statistic instead counts iterate sorts and is not comparable across methods.

Stability sweep. At $n = 20$, $T = 10{,}000$, ten seeds per level, cold CAMW's $\bar{d}_T^{xy}$ increases from 0.06553 at $S_{\text{design}} = 0$ to 0.84130 at 9999. Ordinary least-squares fits of log diagnostic against $\log(1 + S_{\text{design}})$ give slope 0.2706 over all nine levels and 0.3061 over $S_{\text{design}} \in \{1, 5, 20, 50, 200\}$. These values characterize growth on the designed sequences; they are not estimates of an equilibrium-switch exponent.

Dimension sweep. For $T = 20{,}000$, $S_{\text{design}} = 0$, and five seeds at each of seven sizes $n = 10$ through 1000, the coefficient of variation across sizes of CAMW's $\bar{d}_T^x/\sqrt{\log n/T}$ is 0.1193; that of OGD's $\bar{d}_T^x/\sqrt{n/T}$ is 0.1494. At $n = 1000$, the means are 0.01603 (CAMW), 0.02797 (OGD), and 0.42929 (OLMDA): OGD/CAMW is 1.75, not over an order of magnitude. OGD is better at $n \leq 200$, while CAMW is better at 500 and 1000. The OGD/OLMDA distinction includes different step sizes in the same Euclidean geometry. These finite-instance observations do not establish a dimension-dependent theorem comparison.

F.1.2 Transfer, noise, and domain diagnostics

The integrity-checked B5 bundle (ten seeds per cell) gives the following means of $\bar{d}_T^x$. These are descriptive summaries, not significance claims; equal tuning budgets and matched detector persistence were not established. Geometric transfer and OGD have nearly equal diagnostics in the first instance.

Table 3: B5 saved one-sided diagnostic, $T = 10{,}000$. "Jaccard" and "Geometric" are distinct historical WS-CAMW implementations.

Instance	CAMW	Jaccard	Geometric	OGD	OLMDA	SAOL-inspired
$n = 50$, $S_{\text{design}} = 10$	.05475	.05170	.03779	.03789	.09249	.30381
$n = 200$, $S_{\text{design}} = 0$	.02129	.02129	.02129	.01771	.25001	.50824

OGD has the lowest reported diagnostic in the second. The SAOL-inspired engineering baseline differs from the strongly adaptive method analyzed by Daniely et al. (2015), so its rows describe only the archived implementation.

In the integrity-checked B6 noise bundle, cold CAMW's mean $\bar{d}_T^x$ rises from 0.35769 to 0.61238 as noise standard deviation rises from 0.01 to 0.1. Both warm variants range from 0.05099 to 0.05416, while OGD ranges from 0.03826 to 0.06519. In every noisy run, the warm variants have $S_{\text{obs}} = 0$ and $S_{\text{score}} = 9999$: their persistence-two detectors accept no changes, so the observed advantage isolates restart suppression rather than weight transfer. Cold CAMW has both counts equal to 9999. Normalization cancels the generator's margin parameter, so the margin sweep does not measure boundary separation. The noisy wrapper also perturbs numerical gradients.

Shortest-path experiments measure excess cost over the per-round best path. The historical "OGD-FW" baseline is a shortest-path averaging heuristic rather than projected OGD. It has the lower diagnostic at every tested size with zero prescribed changes; with 200 changes, MW has the lower diagnostic on the two smallest grids (Appendix F.4). Simulated SNAP sequences use real graph topology with designed drift and stabilized score-order seeds; OLMDA has the lowest reported diagnostic, and Jaccard WS-CAMW improves on cold CAMW (Appendix F.6). These application-inspired payoffs fall outside the concavity assumptions of the theory.

What a theorem-aligned evaluation requires. A new study must use a common externally declared epoch/cell sequence for all algorithms, record played trajectories and transition rounds, minimize accumulated vertex losses once per epoch, and maximize the accumulated played-x payoff over one fixed y per epoch with certified or explicitly bounded solver error. It must align causal feedback,

both-player restarts, rates, loss normalization, detector persistence, and tuning budgets. The archived scalar summaries cannot supply these quantities retrospectively. Signed values of Eq. (2) must be reported without silently clipping them or taking their logarithms.

F.2 Historical protocol and parameters

The archived runner exposes the current objective to `play`; these runs therefore use a different feedback protocol from Algorithm 1.

- Cold CAMW and both warm variants compute their initial chain using the current first objective. Later plays use the existing chain; after play, `update` computes a current singleton-score ordering, accepts or delays a change, transfers/resets weights, and trains on the resulting chain's losses. The observed permutation is recorded after this update and need not be the chain that generated x_t on a transition round. A zero logged trigger delay is therefore not zero pre-play adoption delay.
- Cold CAMW has persistence one, while Jaccard WS-CAMW and Geometric WS-CAMW default to persistence two and floor parameter 0.1. Geometric transfer matches (9); Jaccard transfer uses row-normalized overlap coefficients and a subsequent renormalization, with no geometric-floor guarantee proved here. The no-restart variant freezes the first chain rather than continuing to change chains without resetting weights.
- MW initializes at $\sqrt{\log(n+1)/T}$ and resets its rate using the remaining global horizon, not elapsed epoch time. It divides losses by each round's observed chain range, which can change the weighting of rounds relative to the raw-loss benchmark. The maximizer state and rate are not restarted on cell changes. Default ascent uses box diameter divided by $\sqrt{T}$.
- Euclidean OGD uses minimizer step $\sqrt{n/T}$; the default Euclidean OLMDA implementation uses $1/\sqrt{T}$. Neither restarts on cell changes. Fixed-Share MW and the SAOL-inspired baseline query the current objective's chain in every play, receiving stronger feedback than the causal algorithm.
- The implemented ZO-EG-style baseline uses two projected extragradient stages, Gaussian finite differences, default smoothing 0.1, and ten samples. Its normalization and projected perturbations differ from the cited method (Farzin et al., 2026). We therefore treat it as a ZO-EG-style diagnostic rather than an implementation of that guarantee.

Algorithm query counts exclude evaluator calls, whereas round timers include some offline score-order bookkeeping. Prefix-set construction and overlap matrices also add costs beyond sorting. We therefore remove generic end-to-end complexity and wall-clock superiority claims. No revised-algorithm runtime is reported.

F.3 Artifact provenance and replication limits

All paths in this subsection are relative to `code/Online-Min-Max/`. The code and saved artifacts were inspected without modification. The archived dimension plot, omitted here because its axis calls the diagnostic regret, is `experiments/results/figures/fig3_dimension_scaling.pdf`; its plotting source is `experiments/generate_final_figures.py`. That script reads `experiments/h100_results_normalized_v2/dimension_scaling/` and `experiments/h100_results_normalized_v2/stability_sweep/`. Identically named files under the top-level `results/` are not pooled with these versions. Dimension rows use five seeds, stability rows ten, with the horizons reported in Section F.1.

B5 is `b5_saol_h100_output_20260428T211706Z.zip`, member `outputs/b5_saol_bundle.pkl`; B6 is `b6_noisy_h100_output_20260429T061516Z.zip`, member `outputs/b6_noisy_bundle.pkl`. Both ZIP hashes match their sidecars, and both pickle hashes match their internal manifests. B5 contains 12 configuration cells and B6 30, each with seeds 0–9. Both use the one-sided diagnostic (`skip_duality_gap=True`); B5 uses $(n, T, y_{\mathrm{dim}}) = (50, 10000, 5)$ with ten designed switches and $(200, 10000, 5)$ with zero. B6 uses $(50, 10000, 5)$ with ten designed switches, noise levels $0.01, 0.05, 0.1$ at margin 0.5, or clean margins $0.05, 0.1, 0.25$.

Only B5 and B6 enter the reported tables because their integrity metadata validate the archived contents; A2 and B1 do not meet this criterion. The single-seed Slashdot CAMW run ($n = 4349$, $T = 500$, ten Monte Carlo samples) serves as a smoke test rather than a comparison. A theorem-aligned study will require complete benchmark trajectories, corrected implementations, symmetric tuning, additional graph replications, and validated packaging. The available metadata do not support a complete compute-budget estimate.

F.4 SHORTEST-PATH DIAGNOSTIC

The source is `experiments/h100_results_normalized_v2/shortest_path/`, the GPU dataset used by the plotting script, at $T = 50{,}000$ and ten seeds.

Its per-round diagnostic is

$$\bar{d}_T^{\text{path}} = \frac{1}{T}\sum_t \left[c_t^\top x_t - \min_{p\in\text{paths}} c_t^\top \mathbb{1}_p\right].$$

MW enumerates all paths; normalization uses a fixed path-length bound, and its learning rate receives the oracle epoch length $T/(S_{\text{path}} + 1)$. Here S_{path} counts prescribed optimal-path changes, not decision-space affine-cell changes. The baseline named OGD-FW updates $x \leftarrow (1-\gamma)x + \gamma \mathbb{1}_{p_t^{\text{best}}}$ with $\gamma = 2/(r+2)$ and a global update counter r. It does not compute a projected-gradient target or an approximation to its projection. We call it the FW heuristic, without the theorem's OGD guarantee. At zero prescribed changes, FW has the lower diagnostic at every tested size. At 200 changes, MW

Table 4: Historical dynamic path-cost diagnostic means, ten seeds. No minimax or switch-threshold interpretation.

Grid k	MW, $S_{\text{path}} = 0$	FW, $S_{\text{path}} = 0$	MW, $S_{\text{path}} = 200$	FW, $S_{\text{path}} = 200$
5	.14752	.00011	1.63951	2.32506
6	.21036	.00015	2.23721	2.62964
8	.35733	.00022	3.73134	2.75641
10	.52896	.00028	5.50351	2.87880
12	.72218	.00035	7.44533	3.15121

has the lower diagnostic for $k = 5, 6$, while FW has the lower diagnostic for the remaining sizes. These descriptive results use a different criterion from the revised theorem.

F.5 APPLICATION-INSPIRED PAYOFFS

The retained theoretical model is unconstrained Lovász minimization on the cube. Ordinary budgeted monotone coverage or influence maximization is a different problem because negating its payoff generally produces a supermodular function. Accordingly, the historical implementation is treated as minimizing the cost-regularized payoffs below, not as maximizing coverage.

The sensor implementation uses a grid fraction

$$f_t(S, y) = \frac{|\{g : \exists i \in S,\ \|g - \text{loc}_i\| \le r_{t,i}(y)\}|/|G| - \sum_{i\in S} c_i}{s},$$

where $s > 0$ is a fixed calibration scale, $y \in [-1, 1]^m$, and $r_{t,i}(y) = \bar{r}_{t,i}\,\text{clip}(1 + b_i^\top y, 0.3, 2.0)$. At fixed y this is coverage minus modular cost and is submodular. As a function of y it is a discontinuous threshold payoff, not concave. The named water-network instance uses randomly generated locations, not loaded water-distribution measurements.

The feature implementation uses cost-regularized facility location,

$$f_t(S, y) = \frac{\sum_j \max_{i\in S}[a_{t,ij} + b_{ij}^\top y]_+ - \sum_{i\in S} c_i}{s},$$

with $y \in [-1, 1]^m$, fixed positive calibration scale s, and zero coverage for $S = \varnothing$. It is submodular for fixed y but generally convex, not concave, in y. The wine/breast-cancer labels refer to dimensions

used to synthesize block similarities; the builders do not load the measurements. This is not mutual information or adversarial regression error. General mutual information need not be submodular: independent fair bits X_1, X_2 and $Z = X_1 \text{ XOR } X_2$ have $I(X_i; Z) = 0$ but $I(X_1, X_2; Z) = 1$. Ordinary squared prediction error is convex in regression weights. These implementations therefore serve as out-of-model diagnostics, and we omit their validation plots.

F.6 Simulated graph sequences

SNAP (Leskovec & Krevl, 2014) provides real graph topology, not the temporal adversary. The influence implementation evaluates Monte Carlo independent-cascade spread minus $\lambda|S|$, with probabilities $p_e(y) = \text{clip}((\alpha_t + 0.1y)p_e^0, 0, 1)$ and $y \in [-1, 1]$. It uses five scheduled drift regimes. Fixed-probability spread is submodular (Kempe et al., 2003), but it is not generally concave in the common probability scale (a two-edge directed chain seeded at its start has expected spread $1+p+p^2$). Sampled live-edge evaluations are discontinuous in y. Score-order queries use regime-stable seeds while payoff queries use round seeds, so the two sampled oracles also differ. Small score-switch counts do not certify small equilibrium SC_T or naturally occurring temporal stability.

Extended rows come from `experiments/h100_results_new/h100_results/snap_extended/`, filtered to Wiki-Vote $n = 787$ and Epinions $n = 2744$, deduplicated by (T, seed), three seeds per horizon.

They use the legacy approximate two-sided diagnostic, not the B5/B6 metric. At $T = 1000$ and 20000, respectively, the Wiki-Vote means are $234.680, 72.667$ (cold CAMW), $111.124, 26.019$ (Jaccard WS-CAMW), and $15.960, 5.147$ (OLMDA). The Epinions means are $4153.363, 1307.147$, $2500.452, 582.753$, and $111.028, 105.840$ in the same order. OLMDA remains better throughout the tested horizons. The Jaccard improvement over cold is about 2.11–2.79 times on Wiki-Vote and 1.66–2.24 times on Epinions. No geometric-transfer rows occur in this plotted dataset.

The separate historical large-influence figure uses Wiki-Vote $n = 1979$, a different graph extraction from the $n = 787$ extended study, so the artifacts are analyzed separately. With three seeds and uncertified maximization on nonconcave payoffs, these runs remain descriptive. We omit the old SNAP figures and their horizon-crossover interpretation because the displayed theoretical ratio cancels T and cannot explain a T/n transition.

G Benchmark Structure, Access, and Comparisons

G.1 Refinement and the sign of the benchmark

Proposition G.1 (Refinement and comparator inclusion)**.** *Fix the played trajectories (x_t, y_t). Splitting an epoch into contiguous subintervals while retaining its comparator cell cannot decrease Eq.* (2)*. For fixed intervals, enlarging each comparator set cannot decrease it either. Arbitrary changes of cells have no such monotonicity.*

Proof. Write $A(I) = \max_y \sum_{t\in I} f_L^t(x_t, y)$ and $B(I, P) = \min_{x\in P} \sum_{t\in I} f_L^t(x, y_t)$. For a partition $I = \bigcup_k I_k$,

$$A(I) \le \sum_k A(I_k), \qquad B(I, P) \ge \sum_k B(I_k, P).$$

Subtracting proves refinement; $P \subseteq P'$ implies $B(I, P') \le B(I, P)$ and proves inclusion. For nonmonotonicity let $n = 2$, $\mathcal{Y}$ be a singleton, $T = 2$, and $x_t = 0$. Under $f_L^t(x) = -x_1 + x_2$, one epoch with P_{12} has regret 2, whereas two singleton epochs with P_{21} have regret 0. Reversing the loss signs reverses these two values. Both losses are modular and bounded by one. This compares evaluation of the same trajectories, not reruns with different restart schedules. □

Proposition G.2 (A sufficient condition for nonnegativity)**.** *Within an epoch with cell P, suppose all $x_t \in P$ and the games on $P \times \mathcal{Y}$ have a common saddle $(x^\dagger, y^\dagger)$. Then that epoch's contribution to Eq.* (2) *is nonnegative. In particular this holds for a stationary game on the cell. Existence of a separate saddle in each round, or in the cumulative game, is insufficient.*

Proof. Put $v_t = f_L^t(x^\dagger, y^\dagger)$. The common-saddle inequalities give

$$\max_y \sum_{t\in I} f_L^t(x_t, y) \ge \sum_{t\in I} f_L^t(x_t, y^\dagger) \ge \sum_{t\in I} v_t \ge \sum_{t\in I} f_L^t(x^\dagger, y_t) \ge \min_{x\in P} \sum_{t\in I} f_L^t(x, y_t).$$

A stationary continuous convex–concave game on compact $P \times \mathcal{Y}$ has such a saddle. The signed example in Section 3 has roundwise saddles and an identically zero cumulative game, yet regret -1, proving the final assertion. □

G.2 Canonical labels: convention and information

Proposition G.3 (Maximal convention sensitivity)**.** *Changing only the coordinate-tie convention can change* SC_T *from* 0 *to* $T-1$*, even for bounded modular losses on two elements.*

Proof. Take singleton $\mathcal{Y}$ and alternate $f_{2r-1}(S) = 0$ and $f_{2r}(S) = -\mathbb{1}\{1 \in S\}$. Successive coordinate minimization selects $(0, 0)$ on odd rounds and $(1, 0)$ on even rounds. Increasing-index ties give $(1, 2)$ throughout. Decreasing-index ties give $(2, 1)$ and $(1, 2)$ alternately. Selection of the point is unchanged, and $|f_t| \le 1$. □

Thus the convention is part of the complexity measure, not an innocuous implementation detail. This sensitivity does not affect a theorem for a fixed declared convention.

Proposition G.4 (Oracle alphabet capacity and prediction)**.** *For* $n \ge 2$*, an equilibrium-cell label has at most* $\log_2(n!)$ *bits of entropy. This bound is attained by bounded submodular–concave games. There is an independent-round distribution of such games for which any permutation-valued predictor using only past objectives and independent internal randomness has expected exact-label mismatches* $T(1 - 1/n!)$.

Proof. There are $n!$ possible labels, proving the upper bound. For $a \in (0, 1)^n$, let $\mathcal{Y} = [-1, 1]^n$ and define

$$f_a(S, y) = \frac{1}{n} \sum_{i=1}^{n} y_i(\mathbb{1}\{i \in S\} - a_i).$$

Its Lovász extension f_L^a satisfies

$$\max_{y\in\mathcal{Y}} f_L^a(x, y) = \frac{1}{n}\|x - a\|_1.$$

The function is modular in S, affine in y, and bounded in absolute value by one; its unique minimax minimizer is a. Its first-order norms are bounded by $1/\sqrt{n}$, so it meets the bounded-oracle assumptions as well. Permuting a fixed list of n distinct coordinates uniformly gives a uniform equilibrium label. Independently redrawing this permutation each round makes the label independent of the predictor's past and randomization. Every predicted permutation then has success probability $1/n!$, and summing mismatch probabilities proves the claim. □

This is an alphabet and exact-prediction statement, not an oracle/no-oracle regret separation. It rules out universal sublinear-mismatch simulation under the present assumptions, but does not attach a regret price to the oracle's bits.

G.3 Proof of the zero-switch in-class lower bound

Proof of Proposition 5.1. Let $\mathcal{Y}$ be a singleton and let independent unbiased signs ε_t define

$$f_t(S) = M\varepsilon_t \frac{|S|}{n}, \qquad f_L^t(x) = \frac{M\varepsilon_t}{n} \sum_{i=1}^{n} x_i.$$

These losses are modular, bounded by M, and have subgradient norm $M/\sqrt{n}$. Their unique minimizers are 0 and $\mathbf{1}$, respectively, both with increasing-index label id. The label oracle therefore reveals no current sign and $\mathrm{SC}_T = 0$ in every realization. A causal learner, even with that oracle, has

expected cumulative loss zero. With $Z_T = \sum_t \varepsilon_t$, the best fixed point in P_{id} has loss $M \min(0, Z_T)$. Consequently

$$\mathbb{E}\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}^\star) = \frac{M}{2}\mathbb{E}|Z_T|.$$

Hölder's inequality and the second and fourth moments give

$$T = \mathbb{E}Z_T^2 \le (\mathbb{E}|Z_T|)^{2/3}(\mathbb{E}Z_T^4)^{1/3}, \qquad \mathbb{E}Z_T^4 = 3T^2 - 2T \le 3T^2.$$

Thus $\mathbb{E}|Z_T| \ge \sqrt{T/3}$. This is an oblivious distributional lower bound valid for every randomized causal learner, and hence implies a worst-case expected lower bound. □

The same calculation gives $(M/(2\sqrt{3}))\sum_j \sqrt{|I_j|} = \Omega(M\sqrt{ET})$ for balanced, externally declared subdivisions of this same cell. These subdivisions still have $\mathrm{SC}_T = 0$ and are not the maximal oracle epochs; this does not prove a lower bound with prescribed positive equilibrium-switch count.

A representable simplex case. The affine-per-cell class also contains the ordinary one-cell experts lower bound: take $\mathcal{X} = \Delta_K$, one cell with K vertices, and $\ell_t(x) = \langle x, z_t\rangle$ for $z_t \in [-M, M]^K$. Every expert-loss vector is now a bounded globally affine loss. Theorem 5.2 with $E = 1$ gives $\Omega(M \min\{T, \sqrt{T \log K}\})$. The affine dimension is $K - 1$, so there is no conflict with the degrees-of-freedom obstruction in Proposition D.1. Neither example supplies a matching structural lower bound for arbitrary cell geometry and equilibrium switches.

G.4 What interval guarantees imply

Supplied boundaries are sufficient for the restart proof, but they are not necessary for an algorithm that controls every interval. The next proposition formalizes this alternative route.

Proposition G.5 (Conditional interval-to-epoch reduction)**.** *Suppose two causal procedures satisfy, simultaneously for every interval I, on the actual opponent-induced losses,*

$$\sum_{t\in I} f_L^t(x_t, y_t) - \min_{x\in[0,1]^n} \sum_{t\in I} f_L^t(x, y_t) \le C_x(T)\sqrt{|I|},$$
$$\max_{y\in\mathcal{Y}} \sum_{t\in I} f_L^t(x_t, y) - \sum_{t\in I} f_L^t(x_t, y_t) \le C_y(T)\sqrt{|I|}.$$

Then, for every E-epoch comparator sequence $\mathcal{Q}$ of cells in the cube, without supplying those cells or boundaries to the procedures,

$$\mathrm{Reg}_T^{\mathrm{epoch}}(\mathcal{Q}) \le (C_x(T) + C_y(T))\sqrt{ET}.$$

Proof. For each I_j, the minimum over the cube is at most the minimum over P_j. Replace it by the latter in the first inequality, add the second, cancel played payoffs, and sum using Cauchy–Schwarz. □

This is the precise benchmark implication of simultaneous interval regret, not a claim that the cited strongly adaptive algorithms have the local MW constant. Their hypotheses must cover the feedback and realized opponent-induced losses. Global convex OCO may incur geometry-dependent constants and interval overheads. The general piecewise-affine model in Section 5.2 need not be globally convex, so this convex-OCO route is not automatic there.

G.5 Oracle-operation accounting

A small active cell gives a statistical reduction only after the corresponding operations are available. The table counts those operations rather than wall-clock time. Let C_f denote one set-value evaluation, C_y one played-point supergradient call, P_y projection onto $\mathcal{Y}$, and C_{sel} exact past-objective best-response/coordinate selection and labeling. These costs depend on the representation and are not assumed constant.

Method	Per-round requirements, excluding evaluator queries
Supplied-chain MW + ascent	$n+1$ set values; one C_y and one P_y; $O(n)$ arithmetic for MW and coordinate sums. Chain access is supplied.
Causal geometric WS-CAMW	The same, plus one C_{sel} on each round after the first; comparison and overlap work when labels change.
Cube projected descent–ascent	One x-subgradient, one C_y, cube clipping ($O(n)$), and one P_y. A Lovász x-subgradient can itself require sorting ($O(n \log n)$) and $n+1$ set values.
General-cell MW	Enumeration/storage of $\lvert V_a \rvert$ vertices, $\lvert V_a \rvert$ loss queries, and barycenter formation (up to $O(d\lvert V_a \rvert)$ for dense vertices); no projection after these are supplied.

One supergradient call at a mixed point may itself aggregate chain supergradients. Explicitly materializing all prefix-set indicators can use $O(n^2)$ storage; a nested-list representation uses $O(n)$ indices but requires a compatible value oracle. Given old and new permutations, equal prefixes for transfer can be detected in $O(n)$ time: scan new prefixes, maintaining the largest old position; size-k prefixes agree exactly when that maximum is k. Transfer arithmetic is then $O(n)$. None of this bounds C_{sel} or proves end-to-end speedup over OGD.

G.6 Switch and epoch notation

Symbol	Meaning and information model
$E, \mathcal{Q}$	Number and sequence of externally declared comparator epochs/cells; boundaries supplied for restarted bounds.
SC_T	Changes of canonical current equilibrium labels; before-play oracle for the corresponding learner; convention-relative.
$\widehat{\mathrm{SC}}_T, \widehat{\mathcal{Q}}$	Changes and maximal epochs of the causal learner's played labels; endogenous comparator.
RS_T	Changes of supplied general-polytope labels, not changes of the roundwise minimizer.
S_{design}	Synthetic generator's prescribed singleton-score-order changes.
S_{obs}	Historical accepted runtime-label counter (some continuous baselines instead count iterate sorts).
S_{score}	Historical offline singleton-score-order changes.
S_{path}	Prescribed optimal-path changes in the historical dynamic-cost diagnostic.

Only $E = 1 + \mathrm{SC}_T$ for maximal oracle epochs and $\widehat{E} = 1 + \widehat{\mathrm{SC}}_T$ for maximal observed epochs are identified in the game theorems. The historical counters are not proxies for these quantities.

H Scope Summary

The main guarantees concern fixed comparators within supplied cell epochs. The in-class lower bound establishes the $\sqrt{T}$ horizon dependence at zero equilibrium-label switches, while the unrestricted local-expert surrogate establishes the full epoch and expert dependence. The comparison with global OGD is a comparison of upper-bound prefactors under the same supplied partition. The archived experiments retain their recorded diagnostic metrics and do not validate the theorem-aligned benchmark. Appendices C–G provide the corresponding benchmark, access, representability, and implementation details.