CONTEXTUAL SWAP REGRET VIA LOCALIZED MARTINGALE ABSORPTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Online forecasts can be assessed through errors that remain correctable in the issued record. Contextual swap regret measures the reduction in cumulative squared loss from contextual replacements and supports calibration guarantees. Distribution-weighted guarantees do not directly control randomized forecasters' realized records. Over T prediction rounds, at fixed dimension and comparator radius, existing methods attain $\widetilde{O}(T^{1/3})$ distribution-weighted regret, whereas the prior realized bound is $\widetilde{O}(T^{3/5})$ with high probability. For affine replacements, we improve the latter to $\widetilde{O}(T^{1/3})$, matching the distribution-weighted horizon exponent. Sampling fluctuations shrink with replacement distance. Our analysis exploits this relationship to let square loss absorb the sampling cost, preserving a finer prediction grid. Lower bounds establish the optimal horizon exponent at fixed confidence. We also characterize expected minimax regret across dimensions and horizons up to logarithmic factors for every fixed $R \geq 1$, where R bounds replacement parameter norms. Experiments assess the grid choice on synthetic streams and census income prediction data. Within a shared ridge implementation, the finer grid reduces median realized regret to $1/15.5$ of the coarse grid value at $T = 10^6$ on synthetic data, and improves realized regret and affine calibration error on the census benchmarks.

1 INTRODUCTION

After an online prediction run, we can ask how much replacing the issued predictions using context could reduce cumulative squared loss. Contextual swap regret formalizes this audit (Garg et al., 2024; Luo et al., 2025). Each prediction value receives its own contextual replacement rule, and regret measures the largest total improvement within a given function class. This quantity measures correctable contextual errors; controlling it supports calibration and multicalibration guarantees (Garg et al., 2024; Luo et al., 2025). We study the optimal high-probability growth of realized contextual swap regret with the number of rounds T, at fixed context dimension.

A randomized forecaster chooses a prediction distribution each round and samples its output. Finite grid methods achieve $\widetilde{O}(T^{1/3})$ regret weighted by sampling probabilities (Luo et al., 2025, Section 2.3). Transferring this guarantee to the realized record requires accounting for sampling and hindsight selection together. Sampling groups rounds by prediction value; the audit then selects each group's best replacement after observing the record. The conversion must control all such improvements simultaneously. The analysis of Luo et al. (2025) gives a realized regret bound of $\widetilde{O}(T^{3/5})$. Does sampling necessarily worsen the horizon exponent?

We prove that sampling need not worsen the horizon exponent. A replacement close to the issued prediction changes the loss only slightly, limiting its sampling fluctuations. The prior conversion uses a uniform worst-case bound, without exploiting this dependence. Accumulating these bounds over prediction values restricts the analysis to a coarser grid.

Square loss lets us exploit this relationship further. A replacement's loss improvement equals the gain from correcting prediction error minus a cost proportional to the squared replacement distance. The same squared distance controls the variance of sampling fluctuations. The proof moves each replacement toward the prediction while keeping its comparator admissible. This absorbs the variance cost into the existing regret bound, preserving the finer grid. For unclipped affine comparators with

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.

bounded parameter norm and an oblivious sequence, we obtain a high-probability realized regret bound of $\widetilde{O}(T^{1/3})$ at fixed context dimension.

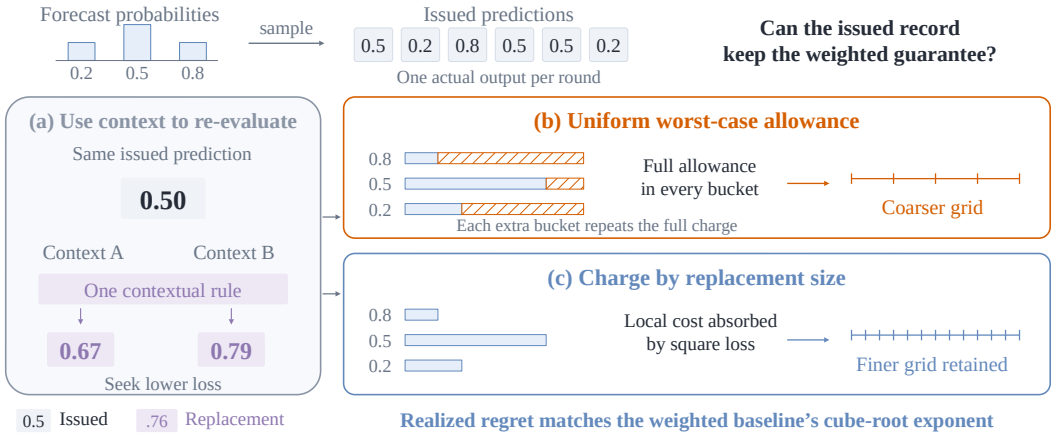

Figure 1: **Why the charging rule changes the affordable grid.** Uniform worst-case analysis repeats one worst-case sampling allowance at every prediction bucket, so excess charges accumulate as the grid is refined. Localized charging follows the size of each post-hoc replacement and is paid from the existing regret budget, allowing the realized analysis to retain the finer grid used before sampling.

Our contributions are as follows.

- **Realized regret bound and analysis.** Localized sampling deviations and an admissible comparator pullback yield a high-probability bound of $\widetilde{O}(T^{1/3}d^{2/3})$ for realized contextual swap regret. This analysis retains a finer grid and improves the horizon exponent at fixed dimension from $3/5$ in Luo et al. (2025) to $1/3$.
- **Matching lower bounds and expected minimax rate.** A fixed-confidence lower bound of $\Omega(T^{1/3})$ establishes the optimal horizon exponent. A dimension lower bound valid at radius one, combined with a dimension-independent upper bound, gives expected minimax regret $\widetilde{\Theta}_R(\min\{T^{1/3}d^{2/3}, T^{3/5}\})$ for every fixed $R \geq 1$ (Corollary 8).
- **Calibration guarantees for affine tests.** A deterministic reduction converts realized regret into swap multicalibration guarantees for affine tests, giving normalized ℓ_1 error $\widetilde{O}(T^{-1/3})$ at fixed dimension (Corollary 5). This recovers the existing calibration rate from improved realized regret (Luo et al., 2025).

2 Background and Motivation

Calibration, multicalibration, and contextual swap regret. Calibration compares forecast probabilities with empirical outcome frequencies (Dawid, 1982; Foster & Vohra, 1998). Online regression yields calibrated forecasts (Foster & Kakade, 2006), while forecast hedging gives a common analysis of several calibration notions (Foster & Hart, 2021). Multicalibration imposes conditional calibration across groups (Hébert-Johnson et al., 2018); multiaccuracy controls residual correlations with contextual tests (Kim et al., 2019). Related formulations study groupwise loss guarantees (Rothblum & Yona, 2021) and outcome indistinguishability (Dwork et al., 2021). Omniprediction connects multicalibration to guarantees across loss functions (Gopalan et al., 2022). Garg et al. (2024) formalize contextual swap regret through a separate contextual replacement for each prediction value and connect swap multicalibration to online regression. We study square-loss swap regret on the realized forecast record within this framework.

Regret guarantees from online regression. Empirical play approaches correlated equilibrium under calibrated best responses (Foster & Vohra, 1997) or regret matching (Hart & Mas-Colell, 2000).

Internal regret also arises in portfolio selection (Stoltz & Lugosi, 2005). Blum & Mansour (2007) reduce swap regret to external regret. Fishelson et al. (2025) extend this approach to continuous action spaces using discretization and randomized rounding. Luo et al. (2025) combine a finite prediction grid with online Newton steps on probability-weighted square losses. Pseudo regret weights gains by prediction probabilities; realized regret uses sampled predictions. For bounded linear classes at fixed dimension and radius, Luo et al. (2025, Section 2.3 and Appendix C.2) obtain $\widetilde{O}(T^{1/3})$ pseudo regret and $\widetilde{O}(T^{3/5})$ realized regret with high probability, using different grids. We close this horizon gap for oblivious sequences and unclipped bounded affine comparators, as summarized in Table 1.

Calibration rates and alternative methods. Gupta et al. (2022) give online multivalidity guarantees for means, moments, and prediction intervals conditional on groups. Lee et al. (2022) study multicalibration through online minimax multiobjective optimization. For bounded linear tests, Luo et al. (2025, Corollary 1) already achieve normalized ℓ_1 swap multicalibration error $\widetilde{O}(T^{-1/3})$ at fixed dimension. Ghuge et al. (2025) directly optimize ℓ_1 multicalibration, attaining normalized error $\widetilde{O}(T^{-1/3})$ with enumerable groups and $\widetilde{O}(T^{-1/4})$ with an oracle-efficient algorithm using an offline optimization oracle. Farina & Perdomo (2026) combine online learning over context and prediction tests with expected variational inequality solvers; their error depends on the test learner's regret. These guarantees differ in error norm, test class, computational access, and whether each prediction value can select its own test. Our affine calibration corollary recovers the existing normalized ℓ_1 time exponent; our main improvement concerns realized contextual swap regret.

Table 1: **Contextual swap regret at fixed dimension, radius, and confidence.** $\widetilde{O}$ hides logarithmic factors. Our row uses the setting of Section 3; its lower bound matches the horizon exponent. Full parameter dependence and assumptions appear in the formal results.

Source	Objective	Bound in T	Status
Luo et al. (2025, Sec. 2.3)	Pseudo regret	$\widetilde{O}(T^{1/3})$	Known pseudo regret bound
Luo et al. (2025, App. C.2)	Realized regret	$\widetilde{O}(T^{3/5})$, h.p.	Prior realized regret bound
Theorem 4 (ours)	Realized regret	$\widetilde{O}(T^{1/3})$, h.p.	Matches pseudo exponent; matching lower bound

3 Problem Setting

Online protocol. Throughout, T is the horizon, namely the number of online interactions, and d is the context dimension. We study contextual prediction against an oblivious adversary, following the randomized prediction protocol of Cesa-Bianchi & Lugosi (2006, Chapter 4). Before the learner's randomization, the adversary fixes a sequence

$$(x_1, y_1), \ldots, (x_T, y_T), \qquad x_t \in \mathbb{R}^d, \quad \|x_t\|_2 \le 1, \quad y_t \in [0, 1].$$

On round t, the learner observes x_t, chooses a distribution P_t over predictions in $[0, 1]$, draws $z_t \sim P_t$, and incurs squared loss $(z_t - y_t)^2$. We write $B_{t,p} = \mathbf{1}\{z_t = p\}$ when predictions are supported on a finite grid.

Obliviousness supplies the conditional independence used in our martingale argument. The sequence can be arbitrary and nonstochastic. Conditional on the past and the current context, the bucket draw is fresh learner randomization and the outcome is fixed.

Comparators and prediction grid. Fix a radius $R > 0$ and let $\tilde{R} = \max(R, 1)$. The *comparator class* is the unclipped affine ball over the augmented input $\tilde{x} = (x, 1) \in \mathbb{R}^{d+1}$,

$$\mathcal{F} = \left\{\tilde{x} \mapsto \langle \tilde{\theta}, \tilde{x} \rangle : \|\tilde{\theta}\|_2 \le \tilde{R}\right\}.$$

This class is convex, symmetric, and contains every constant $x \mapsto p$ for $p \in [0, 1]$, with parameter $(0, \ldots, 0, p)$. We evaluate raw affine values, which may lie outside $[0, 1]$. By convexity and the inclusion of constants, the pullback $(1 - \alpha)p + \alpha f$ stays in the ball for $\alpha \in [0, 1]$. Lemma 3 and Corollary 5 use this closure property. To lighten notation, we write $f(x)$ for the value of an affine comparator on the augmented input $(x, 1)$.

For an integer $N \geq 1$, let

$$Z_N = \left\{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\right\}.$$

The algorithm uses this grid. The benchmark permits $N+1$ independently chosen comparators, one per prediction value, with zero contribution from unvisited values. It equals the atom benchmark defined in Section 5, so different grids are evaluated by the same rule on their realized prediction atoms.

Realized and distribution-weighted regret. A *contextual swap map* (Garg et al., 2024; Luo et al., 2025) assigns to each bucket $p \in Z_N$ a comparator $f_p \in \mathcal{F}$. The swap transformation replaces the learner's prediction $z_t = p$ by $f_p(x_t)$. The *contextual swap regret* is

$$\mathrm{SReg}_T(\mathcal{F}) = \sup_{(f_p)_{p\in Z_N} \in \mathcal{F}^{Z_N}} \sum_{t=1}^{T} \sum_{p\in Z_N} B_{t,p}\left[(p - y_t)^2 - (f_p(x_t) - y_t)^2\right].$$

The *pseudo contextual swap regret* replaces the realized bucket draw by its conditional probability,

$$\mathrm{PSReg}_T(\mathcal{F}) = \sup_{(f_p)_{p\in Z_N} \in \mathcal{F}^{Z_N}} \sum_{t=1}^{T} \sum_{p\in Z_N} P_t(p)\left[(p - y_t)^2 - (f_p(x_t) - y_t)^2\right].$$

For every fixed swap map, replacing $B_{t,p}$ by $P_t(p)$ gives the conditional expectation of its sampled gain on round t. This correspondence does not survive optimization over swap maps. The pseudo quantity optimizes the probability-weighted ledger, while the realized quantity optimizes after the sampled buckets are known. Thus PSReg_T is generally not $\mathbb{E}[\mathrm{SReg}_T]$. The upper bound first controls PSReg_T and then converts it to SReg_T; this is where the localization argument enters.

4 Algorithm and Upper Bound

Section 2 introduced the distribution-weighted baseline and the remaining gap for the realized sequence. We now derive that baseline by balancing discretization and per-bucket learning, then show that the localized sampling offset can be absorbed without changing the balance; Figure 2 previews this decomposition into three costs.

The learner is the standard Blum and Mansour construction, with one weighted online Newton step (ONS) learner per bucket and a master using a stationary distribution. Localization permits the finer resolution $N \sim (T/d)^{1/3}$. Fix a sufficiently large constant C_R depending only on R; throughout Algorithm 1, two-point rounding denotes the distribution on neighboring grid values whose mean is $w_{t,i}$.

Algorithm 1 Bucketed swap learner at the localized grid

Input. horizon T, dimension d, confidence δ, radius R
$\Theta \leftarrow \{\theta \in \mathbb{R}^{d+1} : \|\theta\|_2 \leq 4\tilde{R}\}$
$L \leftarrow (d+1)\log(C_R T) + \log(T/\delta)$; $N \leftarrow \lceil (T/L)^{1/3} \rceil$; grid $Z_N = \{0, \frac{1}{N}, \ldots, 1\}$
for each bucket $i = 0, \ldots, N$ **do** initialize a weighted ONS learner $\mathcal{A}_i$ on Θ
end for
for $t = 1, \ldots, T$ **do**
  observe x_t; set $\tilde{x}_t = (x_t, 1)$
  for each bucket i **do** obtain $\theta_{t,i}$ from $\mathcal{A}_i$; $w_{t,i} \leftarrow \mathrm{clip}_{[0,1]}\langle \theta_{t,i}, \tilde{x}_t \rangle$; $q_{t,i} \leftarrow$ two-point rounding of $w_{t,i}$ onto Z_N
  end for
  assemble column-stochastic Q_t from $\{q_{t,i}\}$; solve $Q_t P_t = P_t$ for the stationary P_t
  sample $z_t \sim P_t$, predict z_t, incur $(z_t - y_t)^2$
  observe y_t; **for** each bucket i, update $\mathcal{A}_i$ with the loss function $\phi_{t,i}(\theta) = P_t(p_i)(\langle \theta, \tilde{x}_t \rangle - y_t)^2$
end for

The rank-one matrix updates cost $O(Nd^2)$ per round, in addition to the ONS projections and the stationary solve for Q_t. Each column has at most two nonzero entries; Appendix D reports measured solver costs.

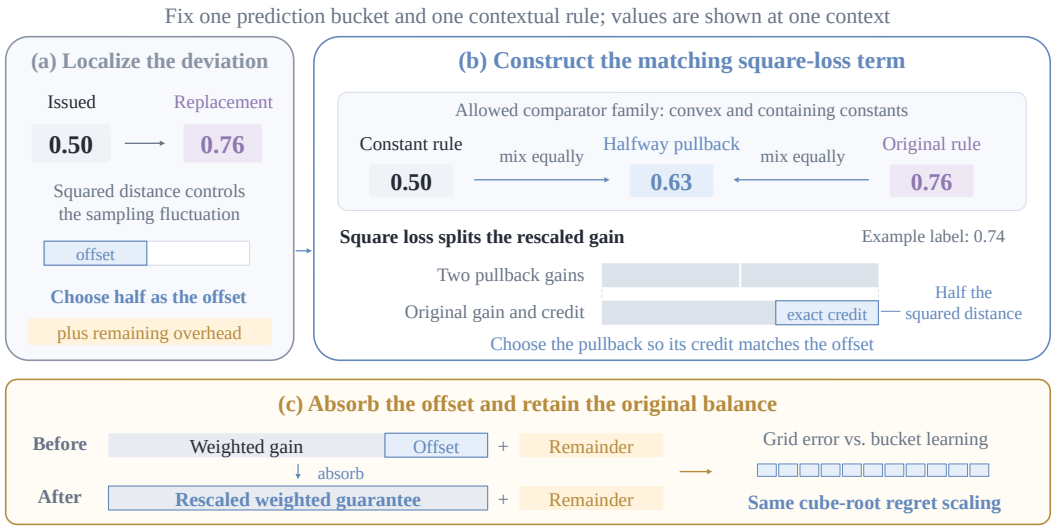

Figure 2: **How the sampling charge is absorbed.** First measure each bucket's deviation at the scale of its proposed replacement. Next pull that replacement toward the issued prediction while staying inside the comparator family. The advantage deliberately given up by this pullback pays the localized sampling charge, so the original balance between grid approximation and per-bucket learning is preserved.

The pseudo-regret guarantee follows the contextual instantiation of Luo et al. (2025, Section 2.3), using the reduction of Blum & Mansour (2007) and ONS (Hazan et al., 2007).

Lemma 1 (Pseudo-regret baseline)**.** *For every oblivious sequence $(x_t, y_t)_{t=1}^T$ with $\|x_t\|_2 \le 1$ and $y_t \in [0,1]$, every integer $N \ge 1$, and every $T \ge 2$, there is a learner such that*

$$\mathrm{PSReg}_T(\mathcal{F}) \le C_R \left(\frac{T}{N^2} + Nd \log(C_R T) \right).$$

Here C_R depends only on the comparator radius R.

The third cost is the sampling deviation, and the next lemma controls it locally. Its key point, compared with prior analyses, is that the martingale variance is localized separately at each prediction bucket p. For

$$\Delta_{t,p,f} = (p - y_t)^2 - (f(x_t) - y_t)^2, \qquad v_{t,p,f} = p - f(x_t),$$

the square-loss structure gives $\Delta_{t,p,f}^2 \le C_R v_{t,p,f}^2$. The conversion in Luo et al. (2025, Lemma 8) also applies Freedman's inequality to each bucket and comparator, but uses a uniform bound on $\Delta_{t,p,f}^2$. Localizing that bound makes the offset absorbable. Appendix A.4 compares the resulting grid balances with dimension and confidence terms retained.

Lemma 2 (Localized martingale bound)**.** *For every oblivious sequence $(x_t, y_t)_{t=1}^T$ with $\|x_t\|_2 \le 1$ and $y_t \in [0,1]$, every integer $N \ge 1$, every $\delta \in (0,1)$, and every $\gamma \in (0,1)$, with probability at least $1-\delta$ over the learner's draws, the following event holds simultaneously for all $p \in Z_N$ and all $f \in \mathcal{F}$,*

$$\sum_{t=1}^{T} (B_{t,p} - P_t(p)) \Delta_{t,p,f} \le \gamma \sum_{t=1}^{T} P_t(p) v_{t,p,f}^2 + \frac{C_R}{\gamma} \left(d \log(C_R T) + \log \frac{N+1}{\delta} \right).$$

Localization leaves a positive offset proportional to the squared distance between the bucket prediction p and the comparator value $f(x_t)$, and convexity of $\mathcal{F}$ absorbs it. For any $p \in Z_N$, $f \in \mathcal{F}$, and $\alpha \in [0,1]$, define $f_\alpha = (1-\alpha)p + \alpha f$; by convexity and the presence of constants, $f_\alpha \in \mathcal{F}$. Shifting the comparison from f to the pullback f_α exposes, under square loss, exactly the same squared distance as the offset, so the variance price folds back into the swap benchmark as part of the regret charge. Choosing $\alpha = 1 - \gamma$ gives the following.

Lemma 3 (Offset absorption)**.** *For every oblivious sequence $(x_t, y_t)_{t=1}^T$ with $\|x_t\|_2 \le 1$ and $y_t \in [0,1]$, every integer $N \ge 1$, every $\delta \in (0,1)$, and every $\gamma \in (0,1)$, on the event of Lemma 2,*

$$\mathrm{SReg}_T(\mathcal{F}) \le \frac{1}{1-\gamma}\,\mathrm{PSReg}_T(\mathcal{F}) + \frac{C_R N}{\gamma}\left(d\log(C_R T) + \log\frac{N+1}{\delta}\right).$$

Consequently, this bound holds with probability at least $1-\delta$.

Combining Lemma 1 and Lemma 3, and choosing $N \approx (T/(d\log(C_R T) + \log(T/\delta)))^{1/3}$, yields the main high-probability contextual swap-regret guarantee.

Theorem 4 (Main bound)**.** *For every oblivious sequence $(x_t, y_t)_{t=1}^T$ with $\|x_t\|_2 \le 1$ and $y_t \in [0,1]$, every $T \ge 2$, and every $\delta \in (0,1)$, there exists a learner such that, with probability at least $1-\delta$,*

$$\mathrm{SReg}_T(\mathcal{F}) \le C_R T^{1/3}\left(d\log(C_R T) + \log\frac{T}{\delta}\right)^{2/3} = \widetilde{O}\left(T^{1/3} d^{2/3}\right).$$

Thus, at fixed dimension, the algorithm improves the previous realized horizon dependence from $\widetilde{O}(T^{3/5})$ to $\widetilde{O}(T^{1/3})$; Table 1 gives the headline horizon comparison.

For comparison with the expectation lower bound below, fixing the grid independently of δ and integrating the tail gives a single learner with $\mathbb{E}\,\mathrm{SReg}_T(\mathcal{F}) = \widetilde{O}(T^{1/3} d^{2/3})$; Appendix A gives the grid and the explicit bound.

The bound separates cleanly into a horizon part and a dimension part. At constant dimension it reads $\widetilde{O}(T^{1/3})$, matching the pseudo-regret baseline, so the localized conversion costs nothing in the exponent relative to the distribution-weighted quantity it converts. The $d^{2/3}$ factor comes from the Nd estimation term, the price of fitting one affine comparator per bucket, which Theorem 7 shows is unavoidable for $R \ge 1$ in its stated large-horizon regime. Because the guarantee holds with high probability over the learner's draws, it certifies the realized sequence an auditor actually sees.

Corollary 5 (Swap multicalibration for affine tests)**.** *For each prediction bucket, let $n_p = |\{t : z_t = p\}|$ and $a_p = \sup_{f\in\mathcal{F}} |\sum_{t:z_t=p}(y_t - p) f(x_t)|$. Writing $S = \mathrm{SReg}_T(\mathcal{F})$ and $C_{\mathrm{mc}} = 3 + \sqrt{2}\,\tilde{R}$, every realized sequence satisfies*

$$\mathrm{MC}_{1,T} := \frac{1}{T}\sum_p a_p \le C_{\mathrm{mc}}\sqrt{\frac{S}{T}}, \qquad \mathrm{MC}_{2,T} := \frac{1}{T}\sum_{p:n_p>0}\frac{a_p^2}{n_p} \le C_{\mathrm{mc}}^2\frac{S}{T}.$$

Here $\mathrm{MC}_{2,T}$ denotes the weighted squared error, without a square root. On the event of Theorem 4, $\mathrm{MC}_{1,T} \le \widetilde{O}(T^{-1/3} d^{1/3})$ and $\mathrm{MC}_{2,T} \le \widetilde{O}(T^{-2/3} d^{2/3})$. If $n_p \ge \rho T > 0$, then $a_p/n_p \le C_{\mathrm{mc}}\sqrt{S/(\rho T)}$.

Each bucket may select its own test function, so these swap multicalibration bounds also control ordinary multicalibration with one common test across buckets. The proof retains each bucket's regret before summing, which avoids a factor of $\sqrt{N}$. The synthetic experiments also report the maximum bucket moment $\mathrm{MC}_{\max,T} := \max_p a_p/T \le \mathrm{MC}_{1,T}$. The experiments on real data and the ONS comparison also report both aggregated errors above. These guarantees apply to affine tests in the stated feature representation, including group indicators represented in $\mathcal{F}$.

5 Lower Bounds and Expected Minimax Rate

We complement Theorem 4 with a fixed-confidence horizon lower bound and an expected minimax characterization for every fixed radius $R \ge 1$. The lower bounds charge independent sources of uncertainty, first across contexts and then across coordinates within each prediction interval.

Benchmark for arbitrary learners. For a learner with arbitrary predictions, let $\mathcal{V}_T = \{z_t : t \le T\}$ be its realized prediction atoms and define

$$\mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F}) = \sup_{(f_v)_{v\in\mathcal{V}_T}\in\mathcal{F}^{\mathcal{V}_T}} \sum_{t=1}^T \left[(z_t - y_t)^2 - (f_{z_t}(x_t) - y_t)^2\right].$$

For a grid learner this equals the regret in Section 3. Distinct predictions can create singleton atoms and a strong hindsight benchmark; learners can control this fragmentation by reusing values. Appendix B describes the comparator families used in the proofs.

The scalar hard instance. Set $m = \lfloor T^{1/3} \rfloor \geq 2$ and take equispaced contexts $x_i = \frac{1}{4} + \frac{i}{2(m-1)}$, $i = 0, \ldots, m-1$, with spacing $\Delta = 1/[2(m-1)]$. Cycle through them for all T rounds, so each appears $q_i \in \{q, q+1\}$ times, where $q = \lfloor T/m \rfloor \geq m^2$. Before play, sample independent labels $y_t \sim \text{Bernoulli}(x_t)$. This defines a distribution over oblivious sequences, independent of the learner's randomization. Predicting near a context leaves label noise that a hindsight constant can exploit; moving away incurs displacement that the common comparator $f(x) = x$ can exploit.

Theorem 6 (Matching lower bound at fixed confidence)**.** *There are absolute constants $c, C > 0$ such that for every $\delta \in (0, \frac{1}{2}]$, every T with $\lfloor T^{1/3} \rfloor \geq C\delta^{-1} \log(4/\delta)$, and every randomized learner, the above oblivious instance satisfies*

$$\mathbb{P}\left[\mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F}) \geq c\,\delta\,T^{1/3}\right] \geq 1 - \delta.$$

The probability includes the instance and learner randomization. For each learner and δ, there is a deterministic oblivious sequence with the same guarantee over that learner's draws. Taking $\delta = 1/2$ gives an $\Omega(T^{1/3})$ expectation lower bound.

Proof mechanism. Both $f(x) = x$ and a separate empirical constant at each prediction atom are admissible. Their gains control squared displacement from the contexts and realized label noise. Group predictions into disjoint intervals of half-width $\Delta/8$ around the contexts. A linear number of independent context noise sums have squared size $\Omega(q)$. The learner can change these sums by dropping own-context increments or inserting foreign ones. These changes are predictable martingale transforms, and their number is controlled by displacement. Small displacement therefore leaves $\Omega(m)$ noise energy; large displacement gives the lower bound directly. Keeping the noise intact requires $q \gtrsim m^2$, which balances $T/m \asymp m^2$ and yields the exponent $1/3$. Appendix B gives the proof and confidence dependence. Proposition 10 further shows that the same horizon lower bound holds with one shared constant per fixed interval and the common contextual comparator $f(x) = x$, so independent fitting at every realized atom is unnecessary.

Dimension dependence, including radius one. Each of $m \asymp (T/d)^{1/3}$ cells carries d independent hidden signs. The construction sets the anchor coefficient to $1/2$, encodes one sign in the intercept, and encodes the remaining signs in coefficient differences. This leaves sufficient joint parameter norm at $R = 1$, including when $d = 1$. Each sign controls only its own Bernoulli labels. A comparator shared within a fixed prediction interval recovers those signs, and the KL chain rule gives an additive estimation cost of $\Omega(d)$ per cell.

Theorem 7 (Dimension lower bound)**.** *There is an absolute $c > 0$ such that for every $d \geq 1$, $R \geq 1$, and $T \geq 8\max\{d, d^{5/2}/R^3\}$, there is a distribution over oblivious instances with $\|x_t\|_2 \leq 1$ such that every learner satisfies*

$$\mathbb{E}_{\text{instance,learner}}[\mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F})] \geq c\,T^{1/3}d^{2/3}.$$

For each learner, a deterministic oblivious sequence retains this lower bound in expectation over its internal randomization.

Expected regret across horizons. Let $V_R(T, d) = \inf_A \sup_{(x_t, y_t)_{t=1}^T} \mathbb{E}_A\, \mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F})$, where the supremum ranges over the oblivious sequences in Section 3 and the learner knows T, d, R. Embedding the preceding construction in fewer coordinates covers horizons below its full-dimensional threshold. A weighted online gradient descent (OGD) variant of the master provides the matching upper branch.

Corollary 8 (Expected minimax rate at fixed radius)**.** *For every fixed $R \geq 1$, all $d \geq 1$, and $T \geq 8$,*

$$V_R(T, d) = \widetilde{\Theta}_R\left(\min\{T^{1/3}d^{2/3}, T^{3/5}\}\right).$$

The notation hides logarithmic factors and constants depending only on R.

The two branches meet at $d \asymp T^{2/5}$, up to radius constants and logarithms. The upper bound selects between ONS and OGD before play using their guarantees; the OGD branch has $O_R(T^{3/5})$ expected regret independent of dimension. Appendix C proves both bounds with explicit parameter dependence. Corollary 8 concerns expectation; Theorem 4 supplies the separate high-probability guarantee.

Table 2: **Headline results.** Finer versus coarse grid within the shared weighted-ridge implementation. Median realized contextual swap regret over 20 seeds at $T = 10^5$, with bootstrap 95% CIs on the median in brackets and the grid resolution N each analysis prescribes. Both arms use matched stream indices; learner seeds include an offset for each method. The last column counts the seeds on which the coarse grid is worse, and an exact two-sided sign test on the paired differences gives $p \le 2 \times 10^{-6}$ in every dimension. Lower is better (**bold**).

	Grid N		Realized swap regret ↓		Paired	
d	Loc.	Coarse	**Localized**	Coarse	Ratio	Coarse worse
2	14	5	**84.2** [81.6,85.8]	669 [657,681]	7.9×	20/20
5	12	5	**119** [117,122]	679 [675,687]	5.7×	20/20
10	10	4	**174** [172,177]	1074 [1065,1077]	6.2×	20/20
20	8	4	**269** [267,271]	1050 [1047,1055]	3.9×	20/20

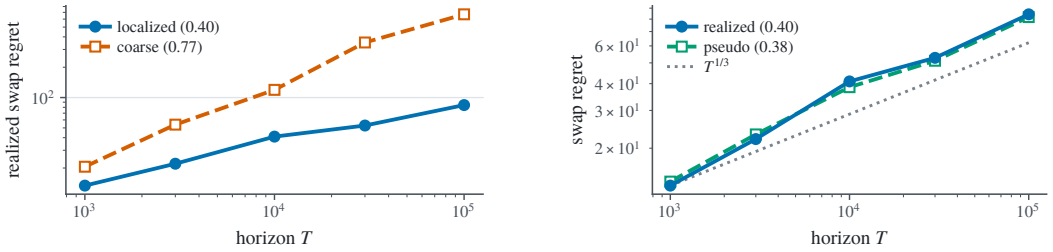

Figure 3: **Horizon scaling** ($d = 2$; medians over 20 seeds). Left. Realized regret under both ridge grids. Right. Pseudo and realized regret under the finer grid, with a $T^{1/3}$ reference.

6 Experiments

6.1 Experimental setup

We report realized contextual swap regret by solving the norm-constrained least-squares problem in each bucket. Synthetic streams use contexts in the unit ball and logistic labels with block drift, retaining their generated order. Adult and ACSIncome use shuffled prefixes without record reuse. The synthetic sweep has $d \in \{2, 5, 10, 20\}$ and $T \in \{10^3, \ldots, 10^5\}$; horizons for real data are 3000, 10000, and 30000. Results are medians over 20 seeds unless stated otherwise.

The grid comparison uses a shared Blum and Mansour implementation with weighted online-ridge bases. The coarse grid follows $N \sim (T/d)^{1/5}$ from Luo et al. (2025), and the finer grid follows $N \sim (T/d)^{1/3}$ from our analysis. We compare a single ridge predictor over 400 runs and weighted ONS with ten seeds for each grid at $d = 2, 10$ and $T = 10^3, 10^4, 10^5$. The ONS update uses projection in the local matrix norm and fixed parameters; its implementation and complete results appear in Appendix D.1.

6.2 Regret and grid resolution

Horizon scaling. Table 2 and Figure 3 compare the two ridge grids. At $T = 10^5$, the finer grid lowers median regret by 3.9 to 7.9× across dimensions, with lower regret on all 20 paired seeds in each dimension. The fitted horizon exponents are 0.4 versus 0.77 at $d = 2$. Both exceed their asymptotic exponents; removing the respective logarithmic factors gives slopes 0.32 and 0.59 to 0.72 across dimensions. At $T = 10^6$ (10 seeds), the finer grid remains 15.5× lower. Figure 4 compares ONS and ridge at $d = 2, 10$. The finer grid lowers ONS regret on every paired seed in all six settings (Table 9).

Additional comparisons. Against a single ridge predictor at $T = 10^5$, finer bucketed ridge improves regret for $d = 2, 5$; the methods are close at $d = 10$, and single ridge is slightly better at $d = 20$ (Table 7).

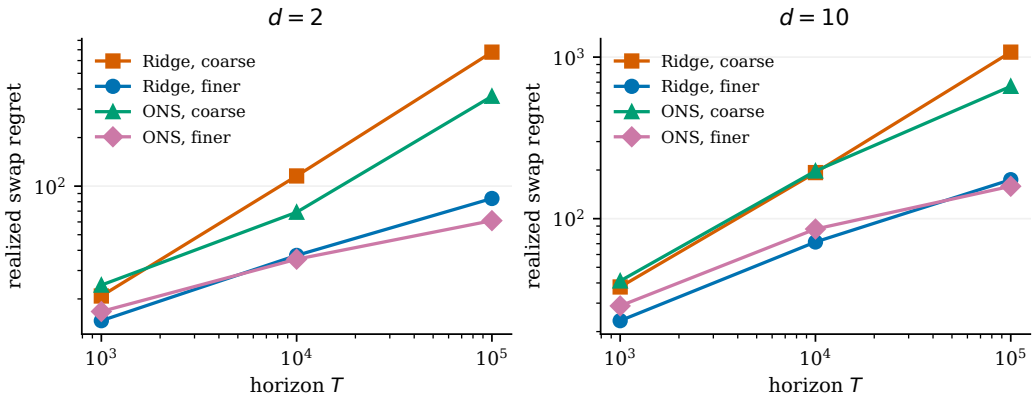

Figure 4: Weighted ONS and ridge at $d = 2$ (left) and $d = 10$ (right). Medians over ten seeds; all four arms share each stream and learner seed.

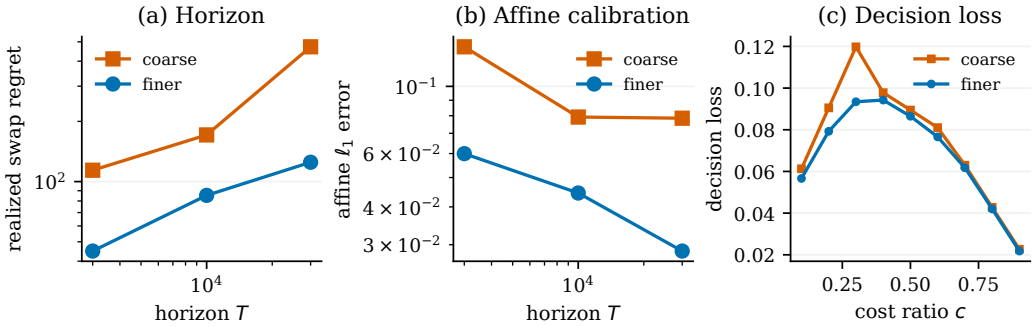

Figure 5: **Adult single-pass stream** (20 seeds). Realized regret and affine ℓ_1 error across horizons, followed by decision loss at $T = 30000$.

6.3 Calibration and issued forecasts

Adult forecasts. We measure the affine ℓ_1 and weighted squared errors of Corollary 5 on the issued predictions. At $T = 30000$, finer ridge reduces the affine ℓ_1 error from 0.0785 to 0.0286. Figure 5 reports regret, affine ℓ_1 error, and threshold decision loss. Table 6 gives decision losses at all nine cost ratios.

External coverage and timing. On Adult and 2018 California ACSIncome, we compare the ridge methods and the finite-group multicalibration algorithm of Lee et al. (2022) using the same affine regret and calibration metrics (Appendix D.1). In a separate synthetic stress test, labels chosen after observing the sampled prediction increase the finer grid's maximum affine bucket moment by 4.2×.

7 Conclusion

Localized sampling deviations let square-loss regret absorb the conversion cost, yielding $\widetilde{O}(T^{1/3}d^{2/3})$ realized contextual swap regret for unclipped bounded affine comparators under an oblivious adversary. The horizon exponent is optimal at fixed confidence. For every fixed $R \geq 1$, matching expected bounds give $\widetilde{\Theta}_R(\min\{T^{1/3}d^{2/3}, T^{3/5}\})$ across dimensions and horizons. A deterministic reduction gives swap multicalibration guarantees for affine tests. Experiments compare the grid prescriptions with ridge and ONS bases, supplemented by a single ridge predictor and a direct calibration baseline under the affine benchmark.

AI USE STATEMENT

We used generative AI tools to assist with drafting, editing, and presentation during manuscript preparation. The authors reviewed and revised all AI-assisted material and take responsibility for the accuracy, claims, and final content of this work.

ETHICS STATEMENT

This work is theoretical, and its experiments use simulated and public data only. The streams of real data are public Adult and ACSIncome census extracts (Becker & Kohavi, 1996; Ding et al., 2021); we report prediction loss, affine regret, and affine calibration on these benchmarks. No human subjects were involved and no new data were collected.

REPRODUCIBILITY STATEMENT

Complete proofs and all assumptions appear in the paper and appendix. The experimental appendix records data construction, implemented grids, evaluation definitions, and reported configurations. The grid scan uses common stream and learner seeds across resolutions; the supplementary comparisons specify their prediction grids, base learners, and evaluation metrics. Reported statistics use the seed counts specified with each experiment.

REFERENCES

Katy S. Azoury and Manfred K. Warmuth. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning*, 43(3):211–246, 2001. doi: 10.1023/A:1010896012157.

Barry Becker and Ronny Kohavi. Adult. *UCI Machine Learning Repository*, 1996. doi: 10.24432/C5XW20.

Alina Beygelzimer, John Langford, Lihong Li, Lev Reyzin, and Robert Schapire. Contextual bandit algorithms with supervised learning guarantees. In Geoffrey Gordon, David Dunson, and Miroslav Dudík (eds.), *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pp. 19–26, Fort Lauderdale, FL, USA, 2011. PMLR. URL https://proceedings.mlr.press/v15/beygelzimer11a.html.

Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8(47):1307–1324, 2007. URL https://www.jmlr.org/papers/v8/blum07a.html.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006. doi: 10.1017/CBO9780511546921.

A. Philip Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77 (379):605–610, 1982. doi: 10.1080/01621459.1982.10477856.

Frances Ding, Moritz Hardt, John Miller, and Ludwig Schmidt. Retiring Adult: New datasets for fair machine learning. In *Advances in Neural Information Processing Systems*, volume 34, pp. 6478–6490, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/32e54441e6382a7fbacbbbaf3c450059-Abstract.html.

Cynthia Dwork, Michael P. Kim, Omer Reingold, Guy N. Rothblum, and Gal Yona. Outcome indistinguishability. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pp. 1095–1108. ACM, 2021. doi: 10.1145/3406325.3451064.

Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*, volume 57 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, 1993. doi: 10.1007/978-1-4899-4541-9.

Gabriele Farina and Juan Carlos Perdomo. An efficient black-box reduction from online learning to multicalibration, and a new route to Φ-regret minimization. *arXiv preprint arXiv:2604.19592*, 2026. URL https://arxiv.org/abs/2604.19592v2.

Maxwell Fishelson, Robert Kleinberg, Princewill Okoroafor, Renato Paes Leme, Jon Schneider, and Yifeng Teng. Full swap regret and discretized calibration. In *Proceedings of the 36th International Conference on Algorithmic Learning Theory*, volume 272 of *Proceedings of Machine Learning Research*, pp. 444–480. PMLR, 2025. URL https://proceedings.mlr.press/v272/fishelson25a.html.

Dean P. Foster and Sergiu Hart. Forecast hedging and calibration. *Journal of Political Economy*, 129(12):3447–3490, 2021. doi: 10.1086/716559.

Dean P. Foster and Sham M. Kakade. Calibration via regression. In *2006 IEEE Information Theory Workshop*, pp. 82–86. IEEE, 2006. doi: 10.1109/ITW.2006.1633786.

Dean P. Foster and Rakesh V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1–2):40–55, 1997. doi: 10.1006/game.1997.0595.

Dean P. Foster and Rakesh V. Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998. doi: 10.1093/biomet/85.2.379.

David A. Freedman. On tail probabilities for martingales. *The Annals of Probability*, 3(1):100–118, 1975. doi: 10.1214/aop/1176996452.

Sumegha Garg, Christopher Jung, Omer Reingold, and Aaron Roth. Oracle efficient online multicalibration and omniprediction. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 2725–2792. SIAM, 2024. doi: 10.1137/1.9781611977912.98. URL https://epubs.siam.org/doi/10.1137/1.9781611977912.98.

Rohan Ghuge, Vidya Muthukumar, and Sahil Singla. Improved and oracle-efficient online ℓ_1-multicalibration. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 19437–19457. PMLR, 2025. URL https://proceedings.mlr.press/v267/ghuge25a.html.

Parikshit Gopalan, Adam Tauman Kalai, Omer Reingold, Vatsal Sharan, and Udi Wieder. Omnipredictors. In *13th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 215 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 79:1–79:21. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi: 10.4230/LIPIcs.ITCS.2022.79.

Varun Gupta, Christopher Jung, Georgy Noarov, Mallesh M. Pai, and Aaron Roth. Online multivalid learning: Means, moments, and prediction intervals. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, volume 215 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 82:1–82:24. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi: 10.4230/LIPIcs.ITCS.2022.82. URL https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ITCS.2022.82.

Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000. doi: 10.1111/1468-0262.00153.

Elad Hazan. Introduction to online convex optimization, 2019. URL https://arxiv.org/abs/1909.05207v1. arXiv:1909.05207v1.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007. doi: 10.1007/s10994-007-5016-8.

Úrsula Hébert-Johnson, Michael P. Kim, Omer Reingold, and Guy N. Rothblum. Multicalibration: Calibration for the (computationally-identifiable) masses. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, volume 80 of *Proceedings of Machine Learning Research*, pp. 1939–1948. PMLR, 2018. URL https://proceedings.mlr.press/v80/hebert-johnson18a.html.

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830.

Michael P. Kim, Amirata Ghorbani, and James Zou. Multiaccuracy: Black-box post-processing for fairness in classification. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES)*, pp. 247–254. ACM, 2019. doi: 10.1145/3306618.3314287.

Daniel Lee, Georgy Noarov, Mallesh Pai, and Aaron Roth. Online minimax multiobjective optimization: Multicalibeating and other applications. In *Advances in Neural Information Processing Systems*, volume 35, pp. 29051–29063. Curran Associates, Inc., 2022. doi: 10.52202/068431-2106. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/ba942323c447c9bbb9d4b638eadefab9-Abstract-Conference.html.

Tengyuan Liang, Alexander Rakhlin, and Karthik Sridharan. Learning with square loss: Localization through offset Rademacher complexity. In *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pp. 1260–1285. PMLR, 2015. URL https://proceedings.mlr.press/v40/Liang15.html.

Haipeng Luo, Spandan Senapati, and Vatsal Sharan. Improved bounds for swap multicalibration and swap omniprediction. In *Advances in Neural Information Processing Systems*, volume 38, pp. 20425–20467. Curran Associates, Inc., 2025. doi: 10.52202/085713-0689. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/1d6a7d85fc445891ef9caea6876875aa-Abstract-Conference.html.

Alexander Rakhlin and Karthik Sridharan. Online non-parametric regression. In *Proceedings of the 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pp. 1232–1264. PMLR, 2014. URL https://proceedings.mlr.press/v35/rakhlin14.html.

Guy N. Rothblum and Gal Yona. Multi-group agnostic PAC learnability. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 9107–9115. PMLR, 2021. URL https://proceedings.mlr.press/v139/rothblum21a.html.

Gilles Stoltz and Gábor Lugosi. Internal regret in on-line portfolio selection. *Machine Learning*, 59 (1–2):125–159, 2005. doi: 10.1007/s10994-005-0465-4.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, 2009. doi: 10.1007/b13794.

Volodya Vovk. Competitive on-line statistics. *International Statistical Review*, 69(2):213–248, 2001. doi: 10.1111/j.1751-5823.2001.tb00457.x.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML)*, pp. 928–936, 2003.

A Proof of Main Theorem

Throughout this appendix, let $Z_N = \{0, 1/N, \ldots, 1\}$. We assume every comparator $f \in \mathcal{F}$ is an affine function over $\tilde{x} = (x, 1) \in \mathbb{R}^{d+1}$ with parameter norm bounded by $\tilde{R} = \max(R, 1)$. Outcomes satisfy $y_t \in [0, 1]$. Constants depending only on R are absorbed into C_R.

A.1 Proof of Lemma 1. Pseudo Swap Regret

For each grid point $p_i = i/N$, instantiate an independent online Newton step (ONS) learner A_i over the enlarged domain $\|\theta\|_2 \le 4\tilde{R}$. On round t, after observing $\tilde{x}_t$, learner A_i outputs $\theta_{t,i}$, forms $\tilde{w}_{t,i} = \langle \theta_{t,i}, \tilde{x}_t \rangle$, and clips to $[0, 1]$ via $w_{t,i} = \mathrm{Proj}_{[0,1]}(\tilde{w}_{t,i})$. Let $q_{t,i} \in \Delta(Z_N)$ be the two-point randomized rounding of $w_{t,i}$ with $\mathbb{E}_{z\sim q_{t,i}}[z] = w_{t,i}$ and support on at most two adjacent grid points. Since $\mathrm{Var}_{z\sim q_{t,i}}(z) \le 1/N^2$,

$$\mathbb{E}_{z\sim q_{t,i}}(z - y_t)^2 \le (w_{t,i} - y_t)^2 + \frac{1}{N^2}. \tag{A.1}$$

Form Q_t with column i equal to $q_{t,i}$ and let P_t be any stationary distribution ($Q_t P_t = P_t$). The master predicts by sampling $z_t \sim P_t$.

Weighted ONS updates. Following the Blum and Mansour reduction (Blum & Mansour, 2007, Theorem 5) and its instantiation in Luo et al. (2025, Algorithm 1), after y_t is revealed, each base learner A_i is updated on the weighted square loss

$$\phi_{t,i}(\theta) = P_t(p_i) \cdot (\langle \theta, \tilde{x}_t \rangle - y_t)^2.$$

The scalar weight $P_t(p_i) \in [0,1]$ is fixed before the loss is supplied to A_i; it may depend on the current context and on the vector of base predictions through the stationary solve. This causes no extra probabilistic condition in the ONS step, because the ONS regret guarantee is pathwise for any adaptively revealed sequence of exp-concave losses.

Weighted ONS regret bound. Write $w = P_t(p_i) \in [0,1]$ and $e = \langle \theta, \tilde{x}_t \rangle - y_t$. On the ONS domain, $|e| \le B_R := 4\sqrt{2}\,\tilde{R} + 1$. The gradient and Hessian are $\nabla \phi_{t,i} = 2we\tilde{x}_t$ and $\nabla^2 \phi_{t,i} = 2w\tilde{x}_t\tilde{x}_t^\top$, respectively. For $\eta = (2B_R^2)^{-1}$,

$$\nabla^2 \phi_{t,i} - \eta \nabla \phi_{t,i} (\nabla \phi_{t,i})^\top = 2w(1 - 2\eta w e^2)\tilde{x}_t\tilde{x}_t^\top \succeq 0. \tag{A.2}$$

Here $2\eta w e^2 \le 1$ uses $w \le 1$, and $w = 0$ gives the zero matrix. The loss is uniformly η-exp-concave and $\|\nabla \phi_{t,i}\|_2 \le 2\sqrt{2} B_R$, including rounds with zero weight.

The weights $P_t(p_i)$ are time-varying and may be chosen after seeing the current vector of base predictions, but the per-round η-exp-concavity holds with $\eta = \Omega(1/\tilde{R}^2)$ *uniformly* in $P_t(p_i) \in [0,1]$. The ONS regret bound depends only on this uniform η and the gradient bound $\|\nabla \phi_{t,i}\| \le C_R$, both of which are insensitive to the realized weights. It first gives the bound with the unclipped prediction $\tilde{w}_{t,i} = \langle \theta_{t,i}, \tilde{x}_t \rangle$ on the learner side. Since $y_t \in [0,1]$ and $w_{t,i} = \mathrm{Proj}_{[0,1]}(\tilde{w}_{t,i})$, projection onto $[0,1]$ cannot increase square loss. $(w_{t,i} - y_t)^2 \le (\tilde{w}_{t,i} - y_t)^2$. The ONS bound (Hazan et al., 2007) therefore gives, for each i and every $f \in \mathcal{F}$,

$$\begin{aligned} \sum_{t=1}^{T} P_t(p_i) \big[(w_{t,i} - y_t)^2 - (f(\tilde{x}_t) - y_t)^2 \big] \\ \le C_R\,(d+1)\log(C_R T). \end{aligned} \tag{A.3}$$

From weighted ONS to pseudo swap regret. By stationarity,

$$\begin{aligned} \mathbb{E}_{z \sim P_t} (z - y_t)^2 &= \sum_{i=0}^{N} P_t(p_i)\, \mathbb{E}_{z \sim q_{t,i}} (z - y_t)^2 \\ &\le \sum_{i=0}^{N} P_t(p_i)(w_{t,i} - y_t)^2 + \frac{1}{N^2}. \end{aligned} \tag{A.4}$$

For any swap rule assigning $f_i \in \mathcal{F}$ to grid point p_i,

$$\mathbb{E}_{z \sim P_t} (f_z(\tilde{x}_t) - y_t)^2 = \sum_{i=0}^{N} P_t(p_i)(f_i(\tilde{x}_t) - y_t)^2. \tag{A.5}$$

Taking the difference, summing over t, and applying equation A.3 for each i, write $R_{t,i}(f) := (w_{t,i} - y_t)^2 - (f(\tilde{x}_t) - y_t)^2$ to get

$$\begin{aligned} \mathrm{PSReg}_T &\le \sum_{i=0}^{N} \sup_{f \in \mathcal{F}} \sum_{t=1}^{T} P_t(p_i) R_{t,i}(f) + \frac{T}{N^2} \\ &\le C_R \Big(N(d+1)\log(C_R T) + \frac{T}{N^2} \Big). \end{aligned}$$

This proves the pseudo swap regret bound.

A.2 Proof of Lemmas 2 and 3

Quadratic offsets localize fluctuations in online square-loss regression (Rakhlin & Sridharan, 2014, Section 3.1) and statistical regression (Liang et al., 2015). Here the offset measures replacement distance within each sampled prediction bucket. The proof below establishes the required martingale event directly for the affine parameter ball.

For the concentration step, write

$$B_{t,p} := \mathbf{1}\{z_t = p\}, \qquad q_{t,p} := P_t(p).$$

For $p \in Z_N$ and $f \in \mathcal{F}$, define

$$\begin{aligned} r_{t,p} &:= p - y_t, \\ v_{t,p,f} &:= p - f(\tilde{x}_t), \\ \Delta_{t,p,f} &:= (p - y_t)^2 - (f(\tilde{x}_t) - y_t)^2. \end{aligned}$$

Then

$$\Delta_{t,p,f} = r_{t,p}^2 - (r_{t,p} - v_{t,p,f})^2 = 2r_{t,p}v_{t,p,f} - v_{t,p,f}^2.$$

Since $p, y_t \in [0,1]$ and $\|\tilde{x}_t\|_2 \le \sqrt{2}$, we have $|f(\tilde{x}_t)| \le \sqrt{2}\,\tilde{R}$, hence $|r_{t,p}| \le 1$ and $|v_{t,p,f}| \le 1 + \sqrt{2}\,\tilde{R}$. It follows that

$$|\Delta_{t,p,f}| \le 2|r_{t,p}||v_{t,p,f}| + v_{t,p,f}^2 \le C_R,$$

and

$$\Delta_{t,p,f}^2 = v_{t,p,f}^2(2r_{t,p} - v_{t,p,f})^2 \le C_R v_{t,p,f}^2.$$

The filtration. Because the adversary is oblivious, the whole sequence $(x_t, y_t)_{t=1}^T$ is fixed before any randomization, and the only fresh randomness at round t is the master's draw $z_t \sim P_t$. The rounding kernel Q_t and the stationary distribution P_t it induces (with $Q_t P_t = P_t$) are computed from the contexts and base learner states up to round t, before z_t is drawn. Define the pre-draw analysis filtration

$$\mathcal{F}_{t-1} := \sigma\big(x_{1:t}, y_{1:t}, Q_{1:t}, P_{1:t}, z_{1:t-1}\big),$$

where P_s denotes the round-s master distribution. This analysis filtration includes fixed outcomes that the learner has not yet observed. under obliviousness, the x_s and y_s are fixed constants on the probability space, even when y_t has not yet been revealed to the learner. Thus x_t, y_t and P_t are $\mathcal{F}_{t-1}$-measurable, while $B_{t,p} = \mathbf{1}\{z_t = p\}$ is not, and the master draw satisfies

$$\mathbb{E}[B_{t,p} \mid \mathcal{F}_{t-1}] = P_t(p) = q_{t,p}.$$

For every $p \in Z_N$ and every *fixed* $f \in \mathcal{F}$, the offset $\Delta_{t,p,f}$ depends only on x_t, y_t, p, f and is therefore $\mathcal{F}_{t-1}$-measurable. Consequently

$$\xi_t := (B_{t,p} - q_{t,p})\,\Delta_{t,p,f}$$

is a martingale difference, $\mathbb{E}[\xi_t \mid \mathcal{F}_{t-1}] = \Delta_{t,p,f}\,\mathbb{E}[B_{t,p} - q_{t,p} \mid \mathcal{F}_{t-1}] = 0$, with $|\xi_t| \le C_R$ and, using $\Delta_{t,p,f}^2 \le C_R v_{t,p,f}^2$,

$$\begin{aligned} \mathbb{E}[\xi_t^2 \mid \mathcal{F}_{t-1}] &= \Delta_{t,p,f}^2\ \mathrm{Var}(B_{t,p} \mid \mathcal{F}_{t-1}) \\ &\le q_{t,p}\,\Delta_{t,p,f}^2 \le C_R\, q_{t,p} v_{t,p,f}^2. \end{aligned}$$

Lemma 2 (Localized martingale bound, restated). *For every $\gamma \in (0,1)$, with probability at least $1-\delta$, simultaneously for all $p \in Z_N$ and all $f \in \mathcal{F}$,*

$$\begin{aligned} &\sum_{t=1}^{T}(B_{t,p} - q_{t,p})\Delta_{t,p,f} \\ &\qquad \le \gamma \sum_{t=1}^{T} q_{t,p} v_{t,p,f}^2 \\ &\qquad\quad + \frac{C_R}{\gamma}\left[(d+1)\log(C_R T) + \log\frac{N+1}{\delta}\right]. \end{aligned}$$

Proof. Single comparator. Fix $p \in Z_N$ and $f \in \mathcal{F}$ and let ξ_t be as above. We use the Freedman-type inequality in the multiplicative form in the proof of Beygelzimer et al. (2011, Theorem 1) (a corollary of Freedman (1975)). For a martingale-difference sequence with $|\xi_t| \le b$ and any $\eta \in (0, 1/b]$, with probability at least $1 - \delta'$,

$$\sum_{t=1}^{T} \xi_t \le (e-2)\,\eta \sum_{t=1}^{T} \mathbb{E}[\xi_t^2 \mid \mathcal{F}_{t-1}] + \frac{\log(1/\delta')}{\eta}.$$

Here $b = C_R$ and $\sum_t \mathbb{E}[\xi_t^2 \mid \mathcal{F}_{t-1}] \le C_R \sum_t q_{t,p} v_{t,p,f}^2$. Choose $\eta = \min\{\gamma/((e-2)C_R),\, 1/C_R\} \in (0, 1/C_R]$. Then the variance coefficient satisfies $(e-2)\eta C_R \le \gamma$ in both branches (it equals γ when $\gamma \le e-2$ and equals $e-2 < \gamma$ otherwise), while $1/\eta \le C_R/\gamma$ for every $\gamma \in (0, 1)$ (when $\gamma \le e-2$, $1/\eta = (e-2)C_R/\gamma \le C_R/\gamma$ since $e-2 < 1$; when $\gamma > e-2$, $1/\eta = C_R \le C_R/\gamma$ since $\gamma < 1$). Hence with probability $\ge 1 - \delta'$

$$\sum_{t=1}^{T} (B_{t,p} - q_{t,p})\Delta_{t,p,f} \le \gamma \sum_{t=1}^{T} q_{t,p} v_{t,p,f}^2 + \frac{C_R}{\gamma} \log \frac{1}{\delta'}. \tag{A.6}$$

Uniformity over p and f. Inequality equation A.6 is for one fixed comparator. Let $\mathcal{N} \subset \{\tilde{\theta} : \|\tilde{\theta}\|_2 \le \tilde{R}\}$ be a Euclidean ε-net of the parameter ball with $\varepsilon = T^{-2}$; the standard volumetric bound gives $|\mathcal{N}| \le (1 + 2\tilde{R}/\varepsilon)^{d+1} = (1 + 2\tilde{R}T^2)^{d+1} \le (C_R T)^{2(d+1)}$, so $\log|\mathcal{N}| \le 2(d+1)\log(C_R T)$. Apply equation A.6 to each of the $(N+1)|\mathcal{N}|$ pairs $(p, g) \in Z_N \times \mathcal{N}$ at confidence $\delta' = \delta/\big((N+1)|\mathcal{N}|\big)$; a union bound makes all of them hold simultaneously with probability $\ge 1 - \delta$, with

$$\begin{aligned}\log \frac{1}{\delta'} &= \log \frac{(N+1)|\mathcal{N}|}{\delta} \\ &\le 2(d+1)\log(C_R T) + \log \frac{N+1}{\delta},\end{aligned}$$

and the harmless factor 2 is absorbed into C_R in the final bound. *Extension from the net to the full class.* Fix any $f \in \mathcal{F}$ with parameter $\tilde{\theta}$ and choose $g \in \mathcal{N}$ with $\|\tilde{\theta} - \tilde{\theta}_g\|_2 \le \varepsilon$. Since $\|\tilde{x}_t\|_2 \le \sqrt{2}$ we have $|f(\tilde{x}_t) - g(\tilde{x}_t)| \le \sqrt{2}\,\varepsilon$, so with $|r_{t,p}| \le 1$ and $|v_{t,p,\cdot}| \le C_R$,

$$\begin{aligned}|v_{t,p,f}^2 - v_{t,p,g}^2| &\le 2C_R\,|v_{t,p,f} - v_{t,p,g}| \le C_R\varepsilon, \\ |\Delta_{t,p,f} - \Delta_{t,p,g}| &\le 2|r_{t,p}|\,|v_{t,p,f} - v_{t,p,g}| + |v_{t,p,f}^2 - v_{t,p,g}^2| \\ &\le C_R\varepsilon.\end{aligned}$$

Summing over t and using $\sum_t |B_{t,p} - q_{t,p}| \le T$ and $T\varepsilon = T^{-1} \le 1$,

$$\begin{aligned}\Big|\sum_t (B_{t,p} - q_{t,p})(\Delta_{t,p,f} - \Delta_{t,p,g})\Big| &\le C_R T\varepsilon \le C_R, \\ \gamma \sum_t q_{t,p}\,|v_{t,p,f}^2 - v_{t,p,g}^2| &\le C_R T\varepsilon \le C_R.\end{aligned}$$

Starting from equation A.6 at the net point g, replacing g by f perturbs the left side by at most C_R and the variance term on the right by at most C_R; both are absorbed into the additive log term. This yields the stated inequality simultaneously over all $p \in Z_N$ and all $f \in \mathcal{F}$, with the constant C_R enlarged to absorb the factor $e - 2$. □

Lemma 3 (Offset absorption, restated). *Fix $\gamma \in (0, 1)$. On the event of Lemma 2 for this γ,*

$$\begin{aligned}\mathrm{SReg}_T \le (1-\gamma)^{-1}\,&\mathrm{PSReg}_T \\ &+ \frac{C_R(N+1)}{\gamma}\left[(d+1)\log(C_R T) + \log\frac{N+1}{\delta}\right].\end{aligned}$$

Proof. Fix $p \in Z_N$. Define $A_p(f) := \sum_t q_{t,p}\Delta_{t,p,f}$ and $S_p(f) := \sum_t B_{t,p}\Delta_{t,p,f}$. Lemma 2 gives, simultaneously for all $f \in \mathcal{F}$,

$$\begin{aligned} S_p(f) &= A_p(f) + \sum_t (B_{t,p} - q_{t,p})\Delta_{t,p,f} \\ &\le \sum_t q_{t,p}[2r_{t,p}v_{t,p,f} - v_{t,p,f}^2] + \gamma \sum_t q_{t,p}v_{t,p,f}^2 + E_p \\ &= \sum_t q_{t,p}[2r_{t,p}v_{t,p,f} - (1-\gamma)v_{t,p,f}^2] + E_p, \end{aligned}$$

where $E_p := \frac{C_R}{\gamma}[(d+1)\log(C_R T) + \log\frac{N+1}{\delta}]$.

Let $\alpha := 1 - \gamma$. Since $\mathcal{F}$ is convex and contains constant functions, $f_\alpha := (1-\alpha)p + \alpha f \in \mathcal{F}$. Moreover, $v_{t,p,f_\alpha} = p - f_\alpha(\tilde{x}_t) = \alpha(p - f(\tilde{x}_t)) = \alpha v_{t,p,f}$. Therefore

$$\begin{aligned} A_p(f_\alpha) &= \sum_t q_{t,p}[2r_{t,p}v_{t,p,f_\alpha} - v_{t,p,f_\alpha}^2] \\ &= \alpha \sum_t q_{t,p}[2r_{t,p}v_{t,p,f} - \alpha v_{t,p,f}^2]. \end{aligned}$$

Since $\alpha = 1 - \gamma$, we obtain $S_p(f) \le \alpha^{-1}A_p(f_\alpha) + E_p \le \alpha^{-1}\sup_{g\in\mathcal{F}} A_p(g) + E_p$. Taking $\sup_f$ and summing over p gives

$$\mathrm{SReg}_T \le (1-\gamma)^{-1}\,\mathrm{PSReg}_T + \sum_p E_p,$$

which is the claimed bound. □

A.3 Proof of Theorem 4

Define

$$L_{T,\delta} := (d+1)\log(C_R T) + \log\frac{T}{\delta}.$$

Apply Lemma 3 with $\gamma = 1/2$ and combine with Lemma 1. With probability at least $1 - \delta$,

$$\mathrm{SReg}_T \le C_R\left[\frac{T}{N^2} + N L_{T,\delta}\right].$$

Here we used $N \le T$ without loss of generality, so $\log\frac{N+1}{\delta} \le C\log\frac{T}{\delta}$. If $L_{T,\delta} > T$ the claim is immediate. Each round contributes at most C_R to SReg_T, so $\mathrm{SReg}_T \le C_R T \le C_R T^{1/3} L_{T,\delta}^{2/3}$. Assume therefore $L_{T,\delta} \le T$ and choose $N = \lceil (T/L_{T,\delta})^{1/3}\rceil \ge 1$. This balances the two terms, yielding

$$\begin{aligned} \mathrm{SReg}_T &\le C_R T^{1/3}\big((d+1)\log(C_R T) + \log\tfrac{T}{\delta}\big)^{2/3} \\ &= \widetilde{O}\Big(T^{1/3} d^{2/3}\Big). \end{aligned}$$

For an expectation bound, fix the learner before integrating its tail. The grid above depends on δ, so we use a single grid, independent of the confidence parameter, defined by $L_0 := (d+1)\log(C_R T) + \log T$, i.e. $N_0 = \lceil (T/L_0)^{1/3}\rceil$, and apply Lemmas 2 and 3 to that one algorithm at confidence $\delta = e^{-u}$. If $L_0 > T$, the bound follows from $\mathrm{SReg}_T \le T$. Otherwise, for this fixed learner, the deviation term is at most $C_R N_0 u$ in addition to $C_R(T/N_0^2 + N_0 L_0)$. Integrating over $u > 0$ gives

$$\mathbb{E}\ \mathrm{SReg}_T \le C_R T^{1/3}\big((d+1)\log(C_R T) + \log T\big)^{2/3}.$$

This proves the main theorem. The reduction uses Blum and Mansour (Blum & Mansour, 2007), the online learning component uses ONS (Hazan et al., 2007), and high-probability bounds use Freedman (Freedman, 1975). The prior contextual swap regret bound is that of Luo et al. (2025).

A.4 Comparison of the grid balances

For $1 \le N \le T$, write $H_{N,\delta} = (d+1)\log(C_R T) + \log((N+1)/\delta)$. The uniform conversion of Luo et al. (2025, Lemma 8 and Appendix C.2), applied to a parameter net, gives an upper bound of the form

$$C_R\left(\frac{T}{N^2} + NH_{N,\delta} + \sqrt{NTH_{N,\delta}}\right). \tag{A.7}$$

The square-root term includes both $\sqrt{NT(d+1)\log(C_R T)}$ and $\sqrt{NT\log((N+1)/\delta)}$, up to constants. Replacing $H_{N,\delta}$ by $L_{T,\delta}$ and balancing that term with T/N^2 selects $N \asymp (T/L_{T,\delta})^{1/5}$. For $L_{T,\delta} \le T$, the resulting scale is $T^{3/5}L_{T,\delta}^{2/5}$. Offset absorption removes this square-root cost. The remaining balance $T/N^2 + NL_{T,\delta}$ selects $N \asymp (T/L_{T,\delta})^{1/3}$ and gives $T^{1/3}L_{T,\delta}^{2/3}$. This comparison suppresses constants depending on the radius, while retaining the dimension complexity that is lost in a fixed-dimension shorthand $\sqrt{NT}$.

A.5 Proof of Corollary 5. Swap multicalibration

The reduction is deterministic. Fix a realized sequence and define its regret at bucket p by

$$s_p := \sup_{f\in\mathcal{F}} \sum_{t:z_t=p} \left[(p-y_t)^2 - (f(x_t)-y_t)^2\right]. \tag{A.8}$$

Comparators are chosen independently at each bucket, so $S = \sum_p s_p$. The constant p is admissible and losses are nonnegative, giving $0 \le s_p \le \sum_{t:z_t=p}(p-y_t)^2 \le n_p$. An empty bucket has $s_p = a_p = 0$ and contributes zero throughout. Fix a nonempty bucket and put

$$r_p = \sum_{t:z_t=p}(y_t - p), \quad H_p(f) = \sum_{t:z_t=p}(y_t-p)f(x_t),$$
$$U_p(f) = H_p(f) - pr_p, \quad V_p(f) = \sum_{t:z_t=p}(f(x_t)-p)^2.$$

With $K_R := 1 + \sqrt{2}\,\tilde{R}$, we have $V_p(f) \le K_R^2 n_p$.

Constant residual. The empirical mean $\bar{y}_p = n_p^{-1}\sum_{t:z_t=p} y_t$ is an admissible constant. Its square-loss improvement is $n_p(p-\bar{y}_p)^2 = r_p^2/n_p$, hence

$$|r_p| \le \sqrt{n_p s_p}. \tag{A.9}$$

Affine moment. For $\alpha \in [0,1]$, convexity and the constant p make $g_\alpha = (1-\alpha)p + \alpha f$ admissible. Expanding its improvement gives

$$2\alpha U_p(f) - \alpha^2 V_p(f) \le s_p. \tag{A.10}$$

If $V_p(f) = 0$, then $f(x_t) = p$ on this bucket and $U_p(f) = 0$. If $U_p(f) \le 0$, an upper bound is immediate. Otherwise $V_p(f) > 0$ and we choose $\alpha = \min\{1, U_p(f)/V_p(f)\}$. When $U_p(f) < V_p(f)$, equation A.10 yields $U_p(f)^2/V_p(f) \le s_p$. When $U_p(f) \ge V_p(f)$, it yields $U_p(f) \le 2U_p(f) - V_p(f) \le s_p$. In every case $U_p(f) \le K_R\sqrt{n_p s_p} + s_p$.

The class is symmetric and $H_p(-f) = -H_p(f)$. Applying the last bound to both f and $-f$, then using equation A.9, gives

$$|H_p(f)| \le K_R\sqrt{n_p s_p} + s_p + pr_p \le (K_R+1)\sqrt{n_p s_p} + s_p \le (K_R+2)\sqrt{n_p s_p}.$$

The final inequality uses $s_p \le n_p$. Taking the supremum over f proves

$$a_p \le C_{\mathrm{mc}}\sqrt{n_p s_p}, \qquad C_{\mathrm{mc}} = K_R + 2 = 3 + \sqrt{2}\,\tilde{R}. \tag{A.11}$$

Aggregation and populated buckets. Cauchy Schwarz and $\sum_p n_p = T$ imply

$$\frac{1}{T}\sum_p a_p \le \frac{C_{\mathrm{mc}}}{T}\sum_p \sqrt{n_p s_p} \le C_{\mathrm{mc}}\sqrt{\frac{S}{T}}, \qquad \frac{1}{T}\sum_{p:n_p>0}\frac{a_p^2}{n_p} \le \frac{C_{\mathrm{mc}}^2}{T}\sum_p s_p.$$

For $n_p \ge \rho T > 0$, equation A.11 also gives $a_p/n_p \le C_{\mathrm{mc}}\sqrt{s_p/n_p} \le C_{\mathrm{mc}}\sqrt{S/(\rho T)}$. Substituting Theorem 4 proves the probability statements. For ordinary multicalibration, moving a common supremum outside the sum can only decrease either error. No union bound over buckets is needed for this deterministic reduction. □

B Proof of the Lower Bound

We prove Theorem 6. Throughout, $m = \lfloor T^{1/3} \rfloor$ and the m contexts are cycled for *all* T rounds, so context i occurs $q_i \in \{q, q+1\}$ times with $q := \lfloor T/m \rfloor$ and $\sum_i q_i = T$; no round is discarded. Since $m^3 \le T$ we have $q \ge m^2$. Write $h := \Delta/8 = \frac{1}{16(m-1)}$ for the half-width of the honest intervals $I_i = [x_i - h, x_i + h]$, so that

$$h^{-2} = 256(m-1)^2 \le 256\, q. \tag{B.1}$$

We use the following constants.

$$\lambda_0 = \tfrac{3}{32}, \qquad p_0 = \tfrac{9}{448}, \qquad \gamma^2 = \tfrac{p_0\lambda_0}{4} = \tfrac{27}{57344}, \qquad b_\delta = \tfrac{\delta\,\gamma^2}{2^{16}}.$$

It suffices to prove the bound for a deterministic adaptive learner. Condition on the learner's internal seed, prove the statement for each fixed seed, and average. Set

$$a_t := z_t - x_t, \qquad \xi_t := x_t - y_t, \qquad B_s := \sum_{t=1}^{s} a_t^2, \qquad L_s := \sum_{t=1}^{s} a_t\xi_t,$$

for $0 \le s \le T$, with $B_0 = L_0 = 0$. Thus a_t is predictable and $\mathbb{E}[\xi_t \mid \mathcal{F}_{t-1}] = 0$ for the pre-outcome filtration.

Let $\mathcal{V}_T := \{z_t : 1 \le t \le T\}$ be the finite set of realized prediction atoms. For each $v \in \mathcal{V}_T$, define

$$S_v := \{t : z_t = v\}, \qquad n_v := |S_v|, \qquad M_v := \sum_{t\in S_v}(v - y_t).$$

For any $I \subseteq [0,1]$, set

$$C(I) := \sum_{v\in\mathcal{V}_T\cap I}\frac{M_v^2}{n_v}, \qquad C_T := C([0,1]) = \sum_{v\in\mathcal{V}_T}\frac{M_v^2}{n_v}.$$

Our scalar proof compares with two admissible families, $f_v(x) = x$ for every atom and $f_v \equiv \bar{y}_v$ chosen separately at each atom. The interval inequality below aggregates these atomwise improvements, with the constants selected independently at individual prediction values.

B.1 A pathwise interval inequality

Index by *prediction interval*. Let $J_t = i$ if $z_t \in I_i$, and $J_t = \perp$ if z_t lies in no honest interval. Put

$$N_i = \#\{t : J_t = i\}, \quad U_i = \sum_{t:J_t=i}\xi_t, \quad V_i = \sum_{t:J_t=i}a_t, \quad B_i = \sum_{t:J_t=i}a_t^2.$$

Since $z_t - y_t = a_t + \xi_t$, the aggregate moment in I_i is $M_i := \sum_{v\in\mathcal{V}_T\cap I_i} M_v = U_i + V_i$. For $N_i > 0$, Cauchy Schwarz gives $C(I_i) \ge M_i^2/N_i$ and $B_i \ge V_i^2/N_i$, while $(U_i+V_i)^2 + V_i^2 \ge U_i^2/2$ because $U_i = (U_i+V_i) - V_i$. Hence, pathwise, $C(I_i) + B_i \ge U_i^2/(2N_i)$. The I_i are disjoint, so $\sum_i C(I_i) \le C_T$ and $\sum_i B_i \le B_T$, and with the convention $0/0 = 0$,

$$B_T + C_T \;\ge\; \tfrac{1}{2}\sum_i U_i^2/N_i. \tag{B.2}$$

This single inequality replaces a case analysis over honest bias, foreign bias and bucket dilution. Every deterministic displacement is charged to B_T, and only label noise survives in U_i.

B.2 A LINEAR NUMBER OF CONTEXT NOISE SUMS ARE LARGE

Let $Z_i := \sum_{t:x_t=x_i} \xi_t$ be the *full* noise sum of context i, over all its rounds. The Z_i are independent, since they use disjoint independent label streams. For $x \in [\frac{1}{4}, \frac{3}{4}]$ a centred Bernoulli increment satisfies $\mathbb{E}\xi^2 = x(1-x) \in [\sigma_0^2, \frac{1}{4}]$ and $\mathbb{E}\xi^4 \le \frac{1}{4}$, so

$$\mathbb{E}[Z_i^2] \ge \sigma_0^2 q_i, \qquad \mathbb{E}[Z_i^4] = q_i \mathbb{E}\xi^4 + 3q_i(q_i-1)(\mathbb{E}\xi^2)^2 \le \tfrac{7}{16} q_i^2.$$

Paley Zygmund applied to Z_i^2 gives $\mathbb{P}(Z_i^2 \ge \frac{1}{2}\mathbb{E}Z_i^2) \ge \frac{1}{4}(\mathbb{E}Z_i^2)^2/\mathbb{E}Z_i^4 \ge p_0$, and $\frac{1}{2}\mathbb{E}Z_i^2 \ge \lambda_0 q_i$, so

$$\mathbb{P}\big(Z_i^2 \ge \lambda_0 q_i\big) \;\ge\; p_0. \tag{B.3}$$

These events are independent across i, so a multiplicative Chernoff bound gives

$$\mathbb{P}\Big(\#\{i : Z_i^2 \ge \lambda_0 q_i\} < \tfrac{p_0 m}{2}\Big) \;\le\; e^{-p_0 m/8}. \tag{B.4}$$

B.3 A SMALL B_T PERMITS ONLY A SMALL PERTURBATION

Call a round of context i a *deviation* if $z_t \notin I_i$; let D be their number. Each has $|a_t| \ge h$, so by equation B.1, $D \le B_T/h^2 \le 256\, q\, B_T$. Writing f_i for the number of *foreign* rounds routed into I_i (true context $\neq i$) and $F = \sum_i f_i$, every foreign round is a deviation from its own context, so $F \le D$.

Define two $\mathbb{R}^m$-valued martingale transforms,

$$R_i = \sum_{t:x_t=x_i,\, z_t \notin I_i} \xi_t, \qquad W_i = \sum_{t:J_t=i,\, x_t \neq x_i} \xi_t,$$

so that coordinatewise $U = Z - R + W$. Here R is noise removed from a context's own coordinate, W is foreign noise inserted into an interval. Both the selection and, for W, the target coordinate are determined by z_t, hence before y_t; increments have Euclidean norm at most one and are orthogonal in L^2, so a transform stopped after at most L selected increments obeys $\mathbb{E}\|R^{(L)}\|_2^2 \le L$ and $\mathbb{E}\|W^{(L)}\|_2^2 \le L$.

Suppose $B_T \le b_\delta m$; otherwise $B_T + C_T \ge b_\delta m$ holds outright. Then $D \le 256\, b_\delta\, m\, q$. Call I_i *overloaded* if $f_i > q$ and let J collect those indices; since $F \le D$, $|J| \le D/q \le 256 b_\delta m \le p_0 m/4$. Let G be the complement. On the complement of the event in equation B.4, at least $p_0 m/4$ indices of G satisfy $Z_i^2 \ge \lambda_0 q_i$, whence

$$\|Z_G\|_2^2 \;\ge\; \tfrac{p_0 m}{4} \lambda_0 q \;=\; \gamma^2 m q. \tag{B.5}$$

Take $L = \lceil 256 b_\delta m q \rceil \le 512 b_\delta m q$ and stop both transforms there. Markov's inequality gives $\mathbb{P}(\|R^{(L)}\|_2 > \frac{\gamma}{4}\sqrt{mq}) \le 8192\, b_\delta/\gamma^2 = \delta/8$ and the same for $W^{(L)}$. On $\{B_T \le b_\delta m\}$ the stopped transforms equal the actual ones, so outside a further event of probability $\delta/4$,

$$\|U_G\|_2 \;\ge\; \|Z_G\|_2 - \|R\|_2 - \|W\|_2 \;\ge\; \tfrac{\gamma}{2}\sqrt{mq}.$$

For $i \in G$ we have $N_i \le q_i + q \le 3q$, so $\sum_i U_i^2/N_i \ge \|U_G\|_2^2/(3q) \ge \gamma^2 m/12$, and equation B.2 yields $B_T + C_T \ge \gamma^2 m/24 \ge b_\delta m$. In both cases, therefore,

$$B_T + C_T \;\ge\; b_\delta\, m. \tag{B.6}$$

The conclusion is global and pathwise given three high-probability events; no context-indexed filtration is needed.

B.4 FROM THE TWO ENERGIES TO SWAP REGRET

Both bounds below are pathwise. For each atom v, let $\bar{y}_v = n_v^{-1} \sum_{t \in S_v} y_t \in [0,1]$. The constant comparator $f_v \equiv \bar{y}_v$ belongs to $\mathcal{F}$ and satisfies

$$\sum_{t \in S_v} \big[(v - y_t)^2 - (\bar{y}_v - y_t)^2\big] = \frac{M_v^2}{n_v}.$$

Choosing this constant separately at every atom gives $\mathrm{SReg}_T^{\mathrm{atom}} \geq C_T$; the identity comparator in every atom gives $\mathrm{SReg}_T^{\mathrm{atom}} \geq \sum_t [(z_t - y_t)^2 - (x_t - y_t)^2] = B_T + 2L_T$. Since $a_t \xi_t$ has conditional mean zero and conditional range of length $|a_t|$, Hoeffding's lemma (Hoeffding, 1963) makes $\exp(-2L_t - B_t/2)$ a nonnegative supermartingale, so with probability at least $1 - \delta/4$ we have $L_T \geq -B_T/4 - \frac{1}{2}\log(4/\delta)$. Combining,

$$\mathrm{SReg}_T^{\mathrm{atom}} \;\geq\; \max\Big\{C_T,\; \tfrac{B_T}{2} - \log\tfrac{4}{\delta}\Big\} \;\geq\; \tfrac{B_T + C_T}{4} - \tfrac{1}{2}\log\tfrac{4}{\delta}. \tag{B.7}$$

B.5 Conclusion

Assume $m \geq 4b_\delta^{-1}\log(4/\delta)$. Since $b_\delta \leq p_0/2$ this also makes the failure probability in equation B.4 at most $\delta/4$; together with the two bounds for the stopped transforms and the supermartingale bound, the total failure probability is at most $3\delta/4$. On the complement, equation B.6 and equation B.7 give $\mathrm{SReg}_T^{\mathrm{atom}} \geq b_\delta m/4 - \frac{1}{2}\log(4/\delta) \geq b_\delta m/8$, and $m \geq T^{1/3}/2$ for $T \geq 8$ yields

$$\mathrm{SReg}_T^{\mathrm{atom}} \;\geq\; \tfrac{b_\delta}{16} T^{1/3} \;=\; \tfrac{\gamma^2}{2^{20}}\,\delta\, T^{1/3}$$

with probability at least $1 - \delta$. This proves Theorem 6 with $c = \gamma^2/2^{20}$ and $C = 2^{18}/\gamma^2$. The swap map $f_v \equiv v$ is admissible, so $\mathrm{SReg}_T^{\mathrm{atom}} \geq 0$ on every realization. Taking $\delta = \frac{1}{2}$, with the same horizon condition, gives directly

$$\mathbb{E}\,\mathrm{SReg}_T^{\mathrm{atom}} \geq \tfrac{c}{2} T^{1/3} \mathbb{P}\Big(\mathrm{SReg}_T^{\mathrm{atom}} \geq \tfrac{c}{2} T^{1/3}\Big) \geq \tfrac{c}{4} T^{1/3}.$$

This is the expectation lower bound. □

Remark 9 (Horizon exponent and confidence dependence). The construction isolates the horizon exponent with explicit constants and the sufficient threshold $T^{1/3} \geq C\delta^{-1}\log(4/\delta)$. The second-moment control of the two vector transforms yields $c(\delta) = \Theta(\delta)$, while the upper bound in Theorem 4 depends logarithmically on $1/\delta$. Thus the two bounds match in the horizon exponent. Sharper confidence dependence would require a stronger Hilbert-space concentration step for the stopped vector transforms.

B.6 A restricted hybrid comparator family

Retain the scalar instance and fixed intervals $I_i = [x_i - h, x_i + h]$ above. Let $\mathcal{H}$ contain the no-change map $\psi(p, x) = p$, the common contextual map $\psi(p, x) = x$, and all maps

$$\psi_c(p, x) = \begin{cases} c_i, & p \in I_i, \\ x, & p \notin \bigcup_i I_i, \end{cases} \qquad c \in [0,1]^m.$$

Define $S_{\mathcal{H}} = \sup_{\psi \in \mathcal{H}} \sum_t [(z_t - y_t)^2 - (\psi(z_t, x_t) - y_t)^2]$. The no-change map guarantees $S_{\mathcal{H}} \geq 0$. At an atom v, the possible affine parameters are $(0, v)$, $(1, 0)$, and $(0, c_i)$, all of norm at most one. It follows that $S_{\mathcal{H}} \leq \mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F})$ for $R \geq 1$. All atoms within an interval share one constant, and the only context-dependent function is x.

Proposition 10 (Hybrid lower bound)**.** *There are absolute $c_{\mathcal{H}}, C_{\mathcal{H}} > 0$ such that for every $\delta \in (0, 1/2]$ and $m = \lfloor T^{1/3} \rfloor \geq C_{\mathcal{H}}\delta^{-1}\log(4/\delta)$, every randomized learner on the scalar oblivious instance satisfies*

$$\mathbb{P}\{S_{\mathcal{H}} \geq c_{\mathcal{H}}\delta T^{1/3}\} \geq 1 - \delta.$$

The probability includes instance and learner randomness. For each learner and δ, a deterministic oblivious sequence retains the same probability bound over the learner's seed. At $\delta = 1/2$, this also gives an $\Omega(T^{1/3})$ expectation lower bound.

Proof. Use $a_t, \xi_t, B_T, L_T, J_t, N_i, U_i$ defined above, and set $E = \sum_{i:N_i>0} U_i^2/N_i$. Let V be the learner's private random seed, independent of the labels, and use $\mathcal{F}_{t-1} = \sigma(V, y_1, \ldots, y_{t-1})$ with the fixed context schedule. On rounds with $J_t = i$, put $b_t = x_t - x_i$, $r_t = z_t - x_i = a_t + b_t$, and $D_i = \sum_{t:J_t=i} b_t$. The best shared constant in that interval is $c_i = \bar{y}_i = x_i + (D_i - U_i)/N_i \in [0, 1]$.

Write $s_t = r_t$ inside the intervals and $s_t = a_t$ outside, and put $M = \sum_t s_t \xi_t$. Completing the square for the interval means gives the exact realized gain of this hybrid map,

$$G = B_T - \sum_{t:J_t \neq \perp} b_t^2 + 2M + \sum_{i:N_i>0} \frac{(D_i - U_i)^2}{N_i}. \tag{B.8}$$

On a visit routed into another context's interval, $|b_t| \geq 8h$ and $|r_t| \leq h$. Therefore $b_t^2 \leq 64 a_t^2/49$ and $r_t^2 \leq a_t^2/49$. On an own-context visit inside its interval, $b_t = 0$ and $s_t = a_t$. Together with the outside rule, these facts imply

$$\sum_t s_t^2 \leq B_T, \qquad \sum_{t:J_t \neq \perp} b_t^2 \leq \frac{64}{49} B_T.$$

Using $(D_i - U_i)^2 \geq U_i^2/2 - D_i^2$ and $D_i^2/N_i \leq \sum_{t:J_t=i} b_t^2$ in equation B.8 yields

$$G \geq \frac{E}{2} - \frac{79}{49} B_T + 2M.$$

The coefficients s_t, a_t are predictable, while ξ_t is centered with conditional range of length one. Hoeffding's lemma (Hoeffding, 1963) makes both $\exp(-2M_t - B_t/2)$ and $\exp(-2L_t - B_t/2)$ nonnegative supermartingales. With failure probability at most 2ε, putting $\ell = \log(1/\varepsilon)$ gives

$$S_{\mathcal{H}} \geq \frac{B_T}{2} - \ell, \qquad S_{\mathcal{H}} \geq \frac{E}{2} - \frac{207}{98} B_T - \ell.$$

Multiply the first inequality by $256/49$, add the second, and divide by $305/49$ to obtain

$$S_{\mathcal{H}} \geq \frac{49}{610}(B_T + E) - \ell. \tag{B.9}$$

The noise argument leading to equation B.6 also controls $B_T + E$ directly. If $B_T > b_\delta m$, the bound is immediate. Otherwise its independent context sums and stopped transforms give $E \geq \gamma^2 m/12 \geq b_\delta m$ except on probability $e^{-p_0 m/8} + \delta/4$. This conclusion uses only the prediction process and the label noise. Under $m \geq (2^{21}/\gamma^2)\delta^{-1}\log(4/\delta)$, the Chernoff failure is at most $\delta/4$, and $256 b_\delta m q \geq 1$ justifies rounding the stopping threshold. We obtain $B_T + E \geq b_\delta m$ with probability at least $1 - \delta/2$. Take $\varepsilon = \delta/4$ in equation B.9. Since $b_\delta m \geq 32\log(4/\delta)$ and $m \geq T^{1/3}/2$, a union bound gives

$$S_{\mathcal{H}} \geq \frac{49}{1220} b_\delta m \geq \frac{\gamma^2}{2^{22}} \delta T^{1/3}$$

with probability at least $1 - \delta$. This proves the proposition with $C_{\mathcal{H}} = 2^{21}/\gamma^2$ and $c_{\mathcal{H}} = \gamma^2/2^{22}$. Averaging over the finite label space provides the deterministic sequence for each learner and confidence level. Nonnegativity gives the expectation consequence. □

C Dimension Lower Bound and Expected Minimax Rate

We prove Theorem 7 using an oblivious distribution that forces independent sign estimation within each cell. The intercept encodes one of the signs, so the construction applies to $d = 1$ and $R = 1$.

C.1 A construction feasible at radius one

Fix integers $m, n \geq 1$, set $\lambda = 1/2$, and choose

$$0 < \gamma \leq \frac{1}{100m}, \qquad \frac{1}{4} + [1 + 16(d-1)]\gamma^2 \leq R^2. \tag{C.1}$$

For $i = 1, \ldots, m$, let

$$a_i = \frac{1}{4} + 10\gamma\left(i - \frac{m+1}{2}\right).$$

For each cell i, sample independent uniform signs $\sigma_{ij} \in \{-1, +1\}$, $j = 0, \ldots, d-1$. Define contexts and label means by

$$x_{i0} = (2a_i, 0, \ldots, 0), \qquad x_{ij} = (2a_i, \lambda e_j) \quad (j \geq 1), \qquad \mu_{ij} = a_i + \gamma\sigma_{ij},$$

where $e_j \in \mathbb{R}^{d-1}$. Visit every pair (i, j) exactly n times in a fixed order. Conditional on the signs, all labels are independent Bernoulli variables of means μ_{ij}. The signs and labels are sampled before learner randomization. The anchors lie in $(1/5, 3/10)$, the means lie in $[1/8, 3/8]$, and $\|x_{ij}\|_2^2 \le (3/5)^2 + (1/2)^2 = 61/100$.

Use the baseline $f_0(x) = x_1/2$ and the cell comparators

$$f_i^\sigma(x) = \frac{x_1}{2} + \gamma\sigma_{i0} + \frac{\gamma}{\lambda}\sum_{j=1}^{d-1}(\sigma_{ij} - \sigma_{i0})x_{j+1}.$$

The intercept is included in the parameter norm. Consequently,

$$\|\tilde{\theta}_i\|_2^2 = \frac{1}{4} + \gamma^2 + 4\gamma^2\sum_{j=1}^{d-1}(\sigma_{ij} - \sigma_{i0})^2 \le \frac{1}{4} + [1 + 16(d-1)]\gamma^2 \le R^2.$$

Thus all these comparators belong to $\mathcal{F}$. For every cell k and coordinate j, including $j = 0$, $f_i^\sigma(x_{kj}) = a_k + \gamma\sigma_{ij}$. When $d = 1$, the sum is empty and $f_i^\sigma(x) = x/2 + \gamma\sigma_{i0}$.

C.2 A fixed interval witness controls estimation risk

The intervals $I_i = [a_i - 3\gamma, a_i + 3\gamma]$ are disjoint and fixed before play. Assign to each prediction p the comparator

$$\phi_\sigma(p) = \begin{cases} f_i^\sigma, & p \in I_i, \\ f_0, & p \notin \bigcup_i I_i. \end{cases} \tag{C.2}$$

This is an admissible atom swap, with one shared affine function per interval. On a round with pair (i, j), condition on the signs, past labels, and current prediction z. Write $\mu = \mu_{ij}$ and $q = \phi_\sigma(z)(x_{ij})$. Conditional independence gives

$$\mathbb{E}_y[(z - y)^2 - (q - y)^2] = (z - \mu)^2 - (q - \mu)^2.$$

If $z \in I_i$, then $q = \mu$. If z lies outside all intervals, then $q = a_i$ and $|z - \mu| \ge 2\gamma$, giving gain at least $3\gamma^2$. If $z \in I_k$ for $k \ne i$, then $|z - \mu| \ge 6\gamma$ and $|q - \mu| \le 2\gamma$, giving gain at least $32\gamma^2$. In all cases,

$$\mathbb{E}_y[(z - y)^2 - (q - y)^2] \ge \min\{(z - \mu)^2, \gamma^2\}. \tag{C.3}$$

This is a conditional expectation bound. On each realized transcript, atom regret is at least the witness's total realized gain.

C.3 Independent sign tests have additive cost

Fix a pair (i, j) and condition on all other signs. Include the learner's independent random seed in its observation history. On the lth visit to this pair, compare the two history laws P_+^l, P_-^l under $\sigma_{ij} = +1, -1$. The learner uses the same decision rule under both hypotheses, and all other pairs have identical conditional label laws. The KL chain rule therefore gives

$$\mathrm{KL}(P_+^l \| P_-^l) \le (l - 1)C_0\gamma^2, \qquad C_0 = 256/7.$$

Indeed, $\mathrm{kl}(p\|q) \le (p - q)^2/[q(1 - q)]$ and $q \in [1/8, 3/8]$ give the stated constant. Estimate the sign by $+1$ exactly when $z \ge a_i$. On an error, $\min\{(z - \mu_{ij})^2, \gamma^2\} \ge \gamma^2$. For $L = \lfloor (2C_0\gamma^2)^{-1} \rfloor + 1$, Pinsker's inequality (Tsybakov, 2009, Chapter 2) bounds total variation by $1/2$ on each of the first L visits. The equal-prior testing error is therefore at least $1/4$, so

$$\mathbb{E}\sum_{l=1}^{n}\min\{(z_l - \mu_{ij})^2, \gamma^2\} \ge \frac{\gamma^2}{4}\min\{n, L\} \ge c_1 \min\{n\gamma^2, 1\}, \qquad c_1 = \frac{1}{8C_0}. \tag{C.4}$$

Average over the other signs and sum over all md pairs. Equations equation C.3 and equation C.4 imply

$$\mathbb{E}\,\mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F}) \ge c_1 md \min\{n\gamma^2, 1\} \tag{C.5}$$

for a horizon of mnd rounds.

C.4 Parameters and completion of the horizon

Assume

$$T \geq 8\max\{d, d^{5/2}/R^3\}, \qquad R \geq 1, \tag{C.6}$$

and set

$$m = \lfloor (T/d)^{1/3} \rfloor, \qquad n = \lfloor T/(md) \rfloor, \qquad \gamma = 1/(100m).$$

Then $m \geq 2$, $m \geq (T/d)^{1/3}/2$, and $n \geq m^2$, because $m^3 d \leq T$. The parameter norm is feasible even at radius one, since

$$[1 + 16(d-1)]\gamma^2 \leq \frac{16d}{10000m^2} \leq \frac{64d^{5/3}}{10000T^{2/3}} \leq 0.0016R^2 \leq R^2 - \frac{1}{4}.$$

Also $n\gamma^2 \geq 1/10000$. Use mnd rounds for the complete visits, and fill the remaining rounds by revisiting existing pairs with fresh conditionally independent labels of their original means, in an order fixed before play. Equation equation C.3 gives nonnegative conditional expected witness gain on every extra round. This completion preserves the lower bound despite the comparators' nonzero intercepts. Substituting in equation C.5 proves

$$\mathbb{E}\,\mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F}) \geq cT^{1/3}d^{2/3}$$

with an absolute $c > 0$. The distribution is the same for every learner. Averaging over its finite support gives, for each learner separately, a deterministic oblivious sequence retaining the expectation lower bound over that learner's seed. This proves Theorem 7.

C.5 The lower envelope at every horizon

For $T \geq 8$, $R \geq 1$, and $d \geq 1$, choose

$$k = \min\left\{d, \left\lfloor \frac{T}{8} \right\rfloor, \left\lfloor \left(\frac{TR^3}{8}\right)^{2/5} \right\rfloor\right\} \geq 1.$$

This choice satisfies $T \geq 8\max\{k, k^{5/2}/R^3\}$. Apply Theorem 7 in dimension k and append zero context coordinates and zero coefficients to embed the construction in $\mathbb{R}^d$. The intercept is unchanged, so the input and joint parameter norms are preserved. A learner in the larger space induces a valid learner on this embedded instance. For every $u \geq 1$, $\lfloor u \rfloor \geq u/2$; hence, for an absolute $c > 0$,

$$V_R(T, d) \geq c\min\{T, T^{1/3}d^{2/3}, R^{4/5}T^{3/5}\}. \tag{C.7}$$

C.6 A dimension-independent expected upper bound

Use the same stationary master as Algorithm 1 with $K = N + 1$ buckets. Replace each ONS base by projected online gradient descent (Zinkevich, 2003) on $\Theta_R = \{\theta : \|\theta\|_2 \leq R\}$, initialized at zero. Write $u_t = (x_t, 1)$ and set

$$g_{t,i} = 2P_t(p_i)(\theta_{t,i}^\top u_t - y_t)u_t,$$

$$\theta_{t+1,i} = \mathrm{Proj}_{\Theta_R}(\theta_{t,i} - \eta g_{t,i}), \qquad \eta = \frac{R\sqrt{K}}{G_R\sqrt{T}},$$

$$G_R = 2\sqrt{2}(\sqrt{2}R + 1).$$

The master uses the clipped base prediction and its two-point rounding onto Z_N. Gradients are computed from the raw affine predictions after the label is revealed.

Weighted learning cost. Since $\|u_t\|_2 \leq \sqrt{2}$ and the current bucket probabilities sum to one,

$$\sum_{t,i} \|g_{t,i}\|_2^2 \leq G_R^2 \sum_{t,i} P_t(p_i)^2 \leq G_R^2 T.$$

For any comparator profile $(\theta_i)_{i=0}^N \in \Theta_R^K$, the projection inequality and convexity of the weighted losses yield

$$\sum_{t,i} P_t(p_i)\left[(\theta_{t,i}^\top u_t - y_t)^2 - (\theta_i^\top u_t - y_t)^2\right] \leq \frac{KR^2}{2\eta} + \frac{\eta}{2}\sum_{t,i}\|g_{t,i}\|_2^2 \leq RG_R\sqrt{KT}.$$

This deterministic bound permits weights computed from the current base predictions. Clipping to $[0, 1]$ reduces square loss. Mean-preserving rounding onto adjacent grid values adds variance at most $1/(4N^2)$. Stationarity, as in equation A.4, therefore gives the pathwise bound

$$\mathrm{PSReg}_T(\mathcal{F}) \le \frac{T}{4N^2} + RG_R\sqrt{KT}. \tag{C.8}$$

A uniform bound on the sampling gap. Use the Hilbert space consisting of a scalar, a vector, and a matrix, equipped with the sum of their squared Euclidean and Frobenius norms. For a fixed grid value p, define

$$Z_{t,p} = (p^2 - 2py_t,\ 2y_tu_t,\ -u_tu_t^\top), \qquad c_\theta = (1, \theta, \theta\theta^\top).$$

The square-loss difference has the exact representation

$$(p - y_t)^2 - (\theta^\top u_t - y_t)^2 = \langle c_\theta, Z_{t,p}\rangle, \qquad \|Z_{t,p}\|^2 \le 13, \quad \|c_\theta\|^2 \le 1 + R^2 + R^4.$$

Let $M_p = \sum_t (B_{t,p} - P_t(p))Z_{t,p}$, and let $M = (M_p)_{p\in Z_N}$ belong to the direct sum of K copies of this space. For each fixed oblivious sequence, $Z_{t,p}$ and $P_t(p)$ are measurable before the master's draw. The increments have conditional mean zero. Temporal orthogonality gives

$$\mathbb{E}\|M\|^2 = \sum_t \mathbb{E}\sum_p P_t(p)(1 - P_t(p))\|Z_{t,p}\|^2 \le 13T. \tag{C.9}$$

The direct sum norm has no cross terms between distinct buckets, so their dependence within a round causes no additional term. On every path, simultaneously for all hindsight comparator profiles,

$$\sum_p \langle c_{\theta_p}, M_p\rangle \le \sqrt{1 + R^2 + R^4}\sum_p \|M_p\| \le \sqrt{K(1 + R^2 + R^4)}\|M\|.$$

Apply this uniform bound to the realized maximizer, then use equation C.8, equation C.9, and Jensen's inequality. We obtain

$$\mathbb{E}\,\mathrm{SReg}_T^{\mathrm{atom}}(\mathcal{F}) \le \frac{T}{4N^2} + \left(RG_R + \sqrt{13(1 + R^2 + R^4)}\right)\sqrt{(N + 1)T}. \tag{C.10}$$

Choosing $N = \lceil T^{1/5}\rceil$ gives $C_RT^{3/5}$ expected regret, independent of dimension. This argument establishes an expectation guarantee.

C.7 Combining the algorithms and proving Corollary 8

The ONS learner with the grid independent of confidence in Appendix A satisfies $\mathbb{E}\,\mathrm{SReg}_T \le C'_RT^{1/3}[(d + 1)\log(C'_RT)]^{2/3}$, after enlarging C'_R to absorb $\log T$. The atomwise identity is admissible, learner losses are at most one, and comparator losses are nonnegative. Hence every learner satisfies $0 \le \mathrm{SReg}_T^{\mathrm{atom}} \le T$. Choose between the ONS and OGD learners before play according to their proved bounds, using only T, d, R. This yields

$$V_R(T, d) \le \min\left\{T,\ C_RT^{3/5}, C'_RT^{1/3}[(d + 1)\log(C'_RT)]^{2/3}\right\}. \tag{C.11}$$

For every fixed $R \ge 1$, the lower envelope equation C.7 is at least $c\min\{T^{1/3}d^{2/3}, T^{3/5}\}$. The upper envelope equation C.11 matches it up to logarithms and radius-dependent constants, proving Corollary 8.

Table 3: Setup for the ridge grid comparison. Both arms use the same weighted-ridge Blum and Mansour implementation and differ only in grid resolution.

Synthetic stream	Oblivious; contexts in the unit ball; logistic labels with block drift
Real stream	Adult census (48,842 records, 11 features), single shuffled pass
Synthetic horizon T	$\{10^3, 3{\cdot}10^3, 10^4, 3{\cdot}10^4, 10^5\}$
Synthetic dimension d	$\{2, 5, 10, 20\}$
Adult horizon T	$\{3 \cdot 10^3, 10^4, 3 \cdot 10^4\}$
Seeds	20 (main sweeps and Adult); 10 at the $T = 10^6$ extension
Learner	Blum and Mansour grid reduction; one weighted online-ridge learner per bucket; stationary master
Grid (localized / coarse)	$N \sim (T/d)^{1/3}$ / $N \sim (T/d)^{1/5}$
Metrics	Realized swap regret (exact per-bucket sup over $\mathcal{F}$); pseudo regret; maximum affine bucket moment
Compute	4 CPU cores, ≈ 3 h total

Table 4: Diagnostic summary, $d = 2$ unless noted. Regret and calibration curves use medians over 20 seeds; the envelope uses the draw counts stated below. "Loc." is the localized grid $N \sim (T/d)^{1/3}$, "Coarse" the grid from uniform conversion $N \sim (T/d)^{1/5}$.

Quantity	Prediction	Outcome
SReg vs T exponent (Loc.)	$\approx 1/3$ (+poly-log)	fit 0.4 (0.37 to 0.45 over d)
SReg vs T exponent (Coarse)	$\approx 3/5$ (+poly-log)	fit 0.77; Loc. $\ll$ Coarse
realized regret at T=10^5	Loc. $\ll$ Coarse	Coarse is 7.9× larger
SReg vs d at T=10^5	sublinear in d	sublinear growth at radius one
grid ablation = $a/N^2 + bN + c$	descriptive tradeoff fit	unified 600-run scan; Table 5
envelope fitted to the sample maximum	empirical transfer across horizons	$1/300$ at 10^4; $0/120$ at 10^5
maximum affine bucket moment	smaller measured residual	Loc. 5.4× better; adaptive labels inflate error by 4.2×

D Additional Experimental Details

We complement Theorem 4 with a numerical study of the choice of grid behind its analysis. Both arms use our implementation of the Blum and Mansour swap reduction over Z_N, with one weighted online-ridge learner per bucket, mean-preserving two-point rounding, and a master using a stationary distribution that samples z_t. Algorithm 1 presents an ONS instantiation of the second-order base learner, while the grid comparison uses weighted online ridge; the master, evaluator, and all other experimental components are shared across the two arms. After each run, we evaluate $\mathrm{SReg}_T(\mathcal{F})$ and $\mathrm{PSReg}_T(\mathcal{F})$ by solving the per-bucket norm-constrained least-squares problems to numerical tolerance. The synthetic and Adult grid comparisons use fixed data streams with $\|x_t\|_2 \le 1$, $y_t \in [0, 1]$, square loss, and bounded affine comparators. A separate stress test chooses labels after observing the sampled prediction. Unless otherwise stated, reported quantities are medians over 20 independent seeds; the separate $T = 10^6$ extension uses 10 seeds. Fitted exponents are free log-log slopes of the median curves. To probe the *exponents* we sweep the horizon over two decades, $T \in \{10^3, \ldots, 10^5\}$, and the dimension over $d \in \{2, 5, 10, 20\}$.

Horizon scaling under the prescription for a finer grid. Fixing d and the localized grid $N \sim (T/d)^{1/3}$, we sweep T over two decades. The realized swap regret grows as $T^{0.4}$ at $d = 2$ (and in the range 0.37 to 0.45 across $d \in \{2, 5, 10, 20\}$; Table 4), with the pseudo-regret curve tracking the same law. Figure 3 shows the representative $d = 2$ curve. These measured exponents are consistent with the $T^{1/3}$ bound up to logarithmic factors over the tested range. The relevant comparison is therefore between the slow realized rate and the coarser scale obtained under a global conversion, which we isolate next by changing only the grid.

Slopes over the observed range are consistent with the poly-log factors. Both fitted exponents exceed the asymptotic values associated with their analyses. We examine the contribution of loga-

rithmic factors through corrected slopes and a longer horizon. Dividing each arm by the factor its own analysis predicts ($(\log T)^{2/3}$ for ours, $(\log T)^{2/5}$ for the prior bound) brings the localized slope to 0.32, matching $T^{1/3}$, and the coarse slope to 0.72 (0.59 to 0.72 over the swept dimensions), which is near or above the $3/5$ scale its analysis predicts. The theorem's exact form $T^{1/3}(\log T)^{2/3}$ fits the localized medians with $R^2 = 0.985$, comparable to a free power law with one fewer parameter. Extending the sweep a further decade to $T = 10^6$ ($d = 2$, 10 seeds) then moves *both* free exponents toward their predictions without any correction (localized $0.40 \to 0.36$, coarse $0.77 \to 0.70$), which is also consistent with this interpretation over this range. At that horizon the localized grid is 15.5× lower on 10/10 seeds. Bootstrap 95% intervals for the two headline exponents are [0.37, 0.40] and [0.76, 0.78].

Runtime of the ridge grid comparison. A full ridge run with the finer grid takes 0.98 to 1.25× the wall-clock of the coarse one, the extra buckets entering only through an $O(Nd^2)$ update at $N \le 29$. Each transition distribution has at most two nonzero entries. A sparse stationary solve scales better than a dense solve at large N (measured exponents $N^{1.8}$ versus $N^{2.4}$, and up to 6.7× speedup), but the crossover occurs only near $N = 512$. At the resolutions prescribed by the theorem used here, the dense solve costs below 0.06 ms per round and is faster, so it is used for the ridge grid comparison.

Grid prescriptions within a shared implementation. The localized and uniform conversion analyses select the finer $N \sim (T/d)^{1/3}$ and coarser $N \sim (T/d)^{1/5}$ resolutions, respectively. We instantiate both within the same weighted-ridge Blum and Mansour implementation, hold streams, paired seeds, and the evaluator fixed, and vary only N. The comparison therefore measures the effect of the two grid prescriptions within this implementation. Under each grid (Figure 3) the localized scale holds the slow $T^{0.4}$ rate while the scale from uniform conversion grows faster on the very same streams. Its fitted exponent is 0.77, at or above the $3/5$ scale of the prior analysis, and at $T = 10^5$ its realized swap regret is 7.9× larger. Thus the prescription for a finer grid reduces realized regret by a clear factor on identical instances, with a gap that widens over the swept horizon range.

The dimension dependence is sublinear. Fixing the largest horizon and sweeping $d \in \{2, 5, 10, 20\}$ with $N \sim (T/d)^{1/3}$, the realized swap regret grows sublinearly in d (Figure 7), consistent with the $d^{2/3}$ factor the bound carries over the tested range.

Unified grid ablation. We evaluate 600 configurations on shared streams and learner seeds across resolutions. The rule for selecting the stationary distribution is fixed for every point, as described in Appendix D.1. For each of the four cells, we fit $\mathrm{SReg} \approx a/N^2 + bN + c$ with nonnegative coefficients to the median curve. The minimum occurs at $4N_0$ in all four cells (Table 5). At $d = 20, T = 10^5$, the descriptive fit sets $b = 0$ despite the observed rise after the minimum. The large regrets at small grids dominate the overall R^2.

Table 5: Unified grid scan, 20 seeds per point. N_0 is the finer prescription; N_* minimizes the observed median. The slope uses the three smallest grid resolutions, and the final column is the increase from the minimum to the largest scanned resolution. Fits constrain a, b, c to be nonnegative.

d	T	N_0	N_*	N_*/N_0	R^2	Slope	Increase
2	10,000	7	28	4	0.998	-1.82	395.2%
20	10,000	4	16	4	0.982	-2.24	204.1%
2	100,000	14	56	4	0.998	-2.11	131.6%
20	100,000	8	32	4	0.992	-1.74	80.2%

Sampling diagnostic with a fixed stream. To isolate the randomness appearing in Theorem 4, we fix the oblivious synthetic sequence separately at each horizon and vary only the learner's draws. At $d = 2$, we use 500, 300, and 120 draws at $T = 10^3, 10^4, 10^5$. For $\delta = 0.05$, the empirical 95th quantiles of realized regret are 16.68, 43.85, and 91.20; after division by $T^{1/3}(d \log T + \log(T/\delta))^{2/3}$, the corresponding ratios are 0.202, 0.208, and 0.175. Calibrating the multiplier to the 95th quantile at $T = 10^3$ gives 0.2019980661, with 44/300 and 0/120 exceedances at the larger horizons. Separately, calibrating to the maximum of the same 500 draws gives 0.2364223177, with 1/300 and 0/120 exceedances. The two fitting rules yield different empirical exceedance frequencies.

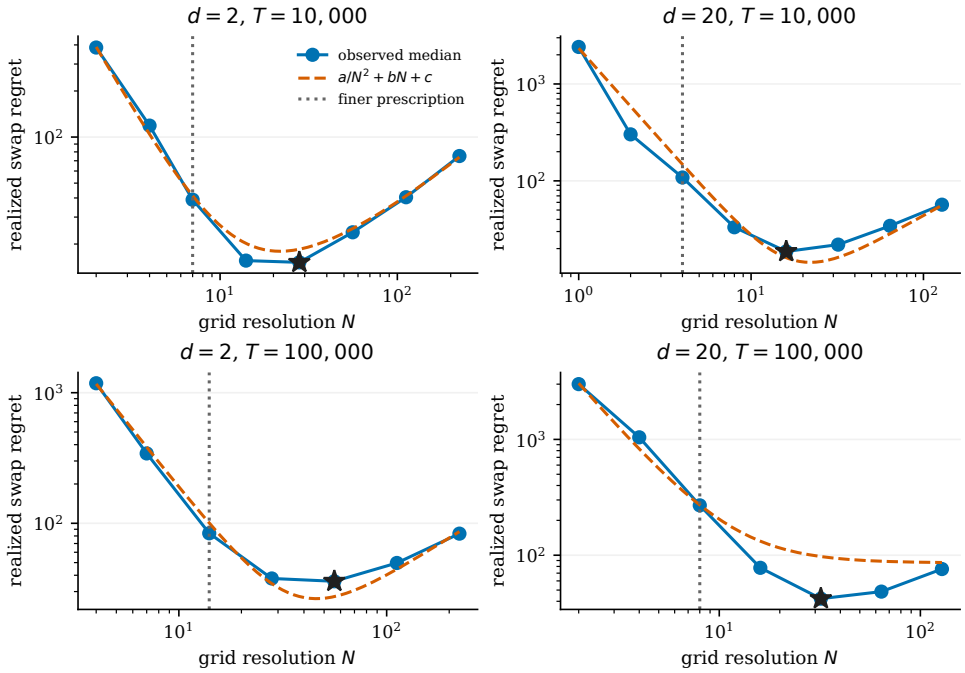

Figure 6: **Unified grid ablation.** Each curve uses 20 shared stream and learner seeds across all grid resolutions. Dashed curves are nonnegative fits of $a/N^2 + bN + c$; dotted lines mark the finer prescription and stars mark observed minima.

Table 6: **Adult decision loss at $T = 30000$.** Medians over 20 seeds with bootstrap 95% CIs at each cost ratio. Every finer/coarse pair has nonoverlapping intervals.

	Finer	Coarse	Gain
$c = 0.1$	**0.0566** [0.0564, 0.0568]	0.0613 [0.0606, 0.0615]	7.6%
$c = 0.2$	**0.0793** [0.0788, 0.0795]	0.0905 [0.0901, 0.0910]	12.4%
$c = 0.3$	**0.0934** [0.0928, 0.0939]	0.1198 [0.1194, 0.1206]	22.0%
$c = 0.4$	**0.0943** [0.0940, 0.0951]	0.0978 [0.0974, 0.0986]	3.6%
$c = 0.5$	**0.0865** [0.0862, 0.0868]	0.0895 [0.0890, 0.0902]	3.3%
$c = 0.6$	**0.0766** [0.0763, 0.0769]	0.0810 [0.0805, 0.0815]	5.5%
$c = 0.7$	**0.0618** [0.0614, 0.0620]	0.0630 [0.0629, 0.0634]	2.0%
$c = 0.8$	**0.0420** [0.0417, 0.0422]	0.0428 [0.0427, 0.0431]	1.8%
$c = 0.9$	**0.0218** [0.0217, 0.0218]	0.0227 [0.0225, 0.0228]	3.8%

Maximum affine bucket moment. For the affine ball, the measured diagnostic has the closed form

$$\mathrm{MC}_{\max,T} = \frac{\tilde{R}}{T} \max_{p} \left\| \sum_{t:z_t=p} (y_t - p)\tilde{x}_t \right\|_2 .$$

This statistic measures the largest residual moment of an individual bucket and is bounded by the aggregated error in Corollary 5. Its value depends on the residuals and their allocation across prediction buckets. The real data and ONS comparisons report the aggregated ℓ_1 and weighted squared errors. On oblivious streams it decays with T, with slope -0.49 at $d = 2$ (Figure 8). The finer grid has a smaller maximum moment by 5.4× at the largest horizon, and 8.6× at $d = 10$, on 20/20 paired seeds in each dimension ($p \le 2 \times 10^{-6}$). When labels depend on the sampled prediction, the finer grid's moment increases by 4.2× at the largest horizon. The increase measures sensitivity to labels that respond to issued forecasts.

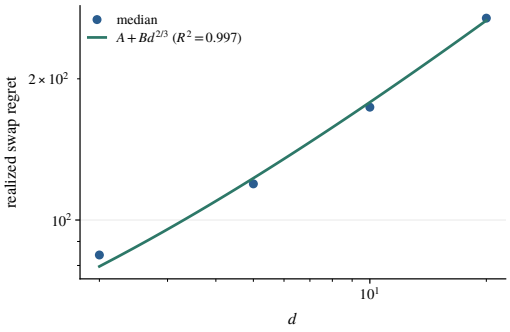

Figure 7: Realized swap regret versus dimension at $T = 10^5$, with the descriptive $A + Bd^{2/3}$ fit at radius one.

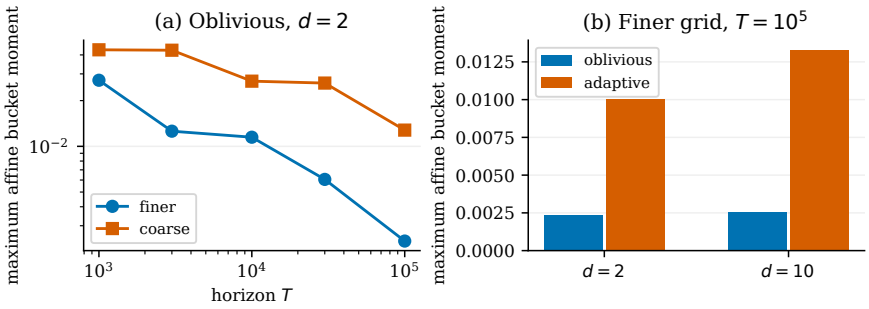

Figure 8: Maximum affine bucket moment. Left. the moment of the produced sequence decays with T on oblivious streams, localized well below coarse. Right. the localized predictor's maximum affine bucket moment at the largest horizon under oblivious and adaptive labels.

Adult results in detail. Table 6 gives the measurements for each cost ratio underlying the Adult results in Section 6.

D.1 Additional baselines and affine calibration

A single ridge predictor. We evaluate a single online-ridge predictor, rounded to the finer grid, over 400 runs. This predictor shares the affine feature representation and evaluator with the bucketed method, but maintains one coefficient vector. Table 7 includes every dimension and horizon. Its intervals for median paired differences use the percentile bootstrap (Efron & Tibshirani, 1993, Chapter 13). At $T = 10^5$, finer bucketed ridge has lower median regret for $d = 2, 5$. The methods are close at $d = 10$, where the interval for their paired difference includes zero. At $d = 20$, single ridge has the lower median. Learner seeds include a separate offset for each method; pairing uses the stream seed index. All sign test p values in the supplementary comparisons are descriptive and unadjusted for multiple comparisons.

Affine calibration metrics. For every run on real data, we compute $\mathrm{MC}_{1,T}$ and $\mathrm{MC}_{2,T}$ from Corollary 5. Each prediction atom selects its own affine test, and the metrics aggregate over all occupied atoms. The evaluator uses the same feature representation and comparator radius one for every method. Table 8 reports both metrics alongside realized regret, pseudo regret, and prediction loss.

ACSIncome and a direct calibration baseline. We use the 2018 California ACS person file from the annual survey with the ACSIncome target and filtering rule of Ding et al. (2021). The filter is age above 16, personal income above 100, positive weekly work hours, and person weight at least one; the target is income above 50,000. It leaves 195,665 records. Our custom 13-dimensional representation consists of standardized age, education level, and weekly hours, a female indicator, and nine race indicators, followed by row normalization to the unit ball. As for Adult, standardization uses

Table 7: Single ridge comparison over 400 runs, paired by dimension, horizon, and stream seed. Δ is finer bucketed minus single ridge. Brackets are percentile bootstrap intervals for the median paired difference (10,000 resamples). The last columns give finer wins and unadjusted two-sided sign test p values. Learner randomizations use a separate seed offset for each method.

d	T	Finer	Single	Δ	95% interval	Wins	p
2	1,000	13.38	13.32	0.42	[-0.93, 2.55]	9/20	0.824
2	3,000	22.06	23.38	-1.08	[-2.85, -0.07]	14/20	0.115
2	10,000	40.97	45.43	-4.51	[-7.40, -3.06]	20/20	1.91×10^{-6}
2	30,000	52.78	84.03	-31.39	[-34.59, -27.44]	20/20	1.91×10^{-6}
2	100,000	84.23	199.01	-115.27	[-117.15, -113.14]	20/20	1.91×10^{-6}
5	1,000	21.28	20.34	1.19	[-0.76, 3.01]	7/20	0.263
5	3,000	36.10	34.70	1.16	[0.23, 3.36]	5/20	0.041
5	10,000	51.58	50.22	0.10	[-1.58, 2.05]	10/20	1.000
5	30,000	83.84	84.76	0.01	[-4.03, 3.30]	10/20	1.000
5	100,000	119.42	133.24	-11.70	[-17.64, -10.61]	19/20	4.01×10^{-5}
10	1,000	22.58	21.08	2.07	[0.90, 2.66]	3/20	0.003
10	3,000	36.24	34.84	2.67	[-0.43, 3.33]	7/20	0.263
10	10,000	71.62	70.27	2.59	[0.01, 3.56]	6/20	0.115
10	30,000	107.96	106.40	3.17	[0.35, 6.11]	4/20	0.012
10	100,000	174.01	173.50	-0.09	[-1.87, 5.25]	10/20	1.000
20	1,000	31.91	30.99	0.77	[-0.68, 3.42]	7/20	0.263
20	3,000	66.27	63.96	2.36	[0.81, 4.83]	5/20	0.041
20	10,000	108.51	105.06	3.57	[1.90, 6.03]	3/20	0.003
20	30,000	144.01	144.04	-0.09	[-0.99, 1.26]	10/20	1.000
20	100,000	269.36	263.81	3.24	[0.74, 7.45]	5/20	0.041

the full unlabeled feature table; survey weights are not applied. We use $T \in \{3000, 10000, 30000\}$ and 20 shuffled prefixes without record reuse. The permutation seed is $1000 + s$ and the learner seed is s, for $s = 0, \ldots, 19$.

On both real datasets we compare coarse and finer bucketed ridge, single ridge on the finer grid, and the direct algorithm for finite groups of Lee et al. (2022, Appendix D, Algorithm D.1). For the latter, we fix $n = 10$ calibration intervals, $r = 10$, and $\eta = \sqrt{\log(2 \cdot 7 \cdot 10)/(4T)}$, with the six demographic groups and the whole population supplied to the learner. The six groups are female, male, White alone, Black alone, age at most 35, and age above 50. The prediction grid has 101 points. At a sign crossing we select the first qualifying adjacent pair and use the algorithm's stated convention $0/0 = 1$. This method accesses raw group membership and controls bin moments for a finite family of groups. We evaluate all four methods using the same affine regret and calibration metrics. All labels are fixed before learner randomization. At $T = 30000$, finer ridge has lower median affine regret and affine ℓ_1 error than the other three methods on both datasets. Table 8 reports every selected configuration.

Weighted ONS implementation. For the weighted raw square loss, we use $g_{t,i} = 2P_t(p_i)(\langle \theta_{t,i}, \tilde{x}_t \rangle - y_t)\tilde{x}_t$, $A_{t,i} = A_{t-1,i} + g_{t,i}g_{t,i}^\top$, and

$$\theta_{t+1,i} = \operatorname*{argmin}_{\|\theta\| \le 4} \left\| \theta - \left(\theta_{t,i} - \gamma^{-1} A_{t,i}^{-1} g_{t,i} \right) \right\|_{A_{t,i}}^2 .$$

Set $B = 4\sqrt{2} + 1$, $D = 8$, $G = 2\sqrt{2}B$, $\alpha = 1/(2B^2)$, $\gamma = \frac{1}{2}\min\{1/(4GD), \alpha\}$, and $A_{0,i} = (\gamma D)^{-2} I$, with $\theta_{1,i} = 0$. These conservative ONS parameters are fixed before the runs, following Hazan (2019, Theorem 4.4); the ONS method originates in Hazan et al. (2007). We compute the projection by its KKT form $(A_{t,i} + \lambda I)^{-1} A_{t,i} v$, using an eigendecomposition and 60 bisection steps when the unconstrained point v is outside the ball.

The comparison uses $d \in \{2, 10\}$, $T \in \{10^3, 10^4, 10^5\}$, and ten seeds for each ONS/ridge and finer/coarse combination, giving 240 runs. The stream seed is $10000d + 101T + s$ and the learner seed is s in all four arms. The practical grid constants are specified in Appendix E. All four synthetic arms use the same explicit stationary rule. We identify closed strongly connected components of the transition graph with at most two outgoing edges per state, choose the component containing the smallest state index among closed components, and solve within that irreducible component. Probability is zero elsewhere, and no damping is used. This selects an admissible stationary distribution even when the full chain has several stationary distributions. The stationary residual tolerance is

Table 8: Results on real data over 20 seeds. S and P are realized and pseudo affine swap regret; M_1 and M_2 are the affine metrics over prediction atoms of Corollary 5. Brier is mean square loss; seconds measure the learner loop on the same host, excluding evaluation and storage. Lee optimizes moments for finite groups in ten bins, so its affine metrics over prediction atoms assess a different objective.

Data	T	Method	N	S	P	M_1	M_2	Brier	Sec.
Adult	3,000	Coarse	2	114.0	115.3	0.1351	0.0188	0.160	0.68
Adult	3,000	Finer	4	44.9	43.1	0.0600	0.0038	0.137	0.69
Adult	3,000	Single	4	48.0	47.2	0.0589	0.0038	0.137	0.27
Adult	3,000	Lee	100	141.2	136.8	0.0823	0.0093	0.167	0.37
Adult	10,000	Coarse	3	171.1	169.8	0.0792	0.0064	0.139	2.13
Adult	10,000	Finer	5	85.3	80.4	0.0445	0.0021	0.129	2.33
Adult	10,000	Single	5	116.8	112.5	0.0484	0.0026	0.132	0.88
Adult	10,000	Lee	100	420.4	414.4	0.0763	0.0071	0.164	1.21
Adult	30,000	Coarse	3	472.1	473.4	0.0785	0.0062	0.137	6.64
Adult	30,000	Finer	7	124.9	121.9	0.0286	0.0009	0.124	6.98
Adult	30,000	Single	7	261.3	258.3	0.0468	0.0025	0.129	2.85
Adult	30,000	Lee	100	1182.7	1178.5	0.0748	0.0064	0.163	3.66
ACS	3,000	Coarse	2	133.3	131.9	0.1222	0.0192	0.197	0.89
ACS	3,000	Finer	3	71.5	68.3	0.0723	0.0073	0.175	0.96
ACS	3,000	Single	3	74.1	71.8	0.0637	0.0061	0.175	0.27
ACS	3,000	Lee	100	237.7	231.1	0.1216	0.0163	0.228	0.35
ACS	10,000	Coarse	3	196.0	193.5	0.0677	0.0069	0.172	3.30
ACS	10,000	Finer	5	93.6	90.5	0.0369	0.0021	0.161	3.34
ACS	10,000	Single	5	124.5	117.8	0.0493	0.0031	0.164	0.89
ACS	10,000	Lee	100	716.7	710.3	0.1151	0.0137	0.225	1.18
ACS	30,000	Coarse	3	542.9	543.6	0.0660	0.0067	0.171	9.83
ACS	30,000	Finer	6	171.1	168.4	0.0266	0.0012	0.157	9.50
ACS	30,000	Single	6	283.7	279.2	0.0475	0.0027	0.161	2.90
ACS	30,000	Lee	100	2054.6	2044.9	0.1117	0.0127	0.223	3.60

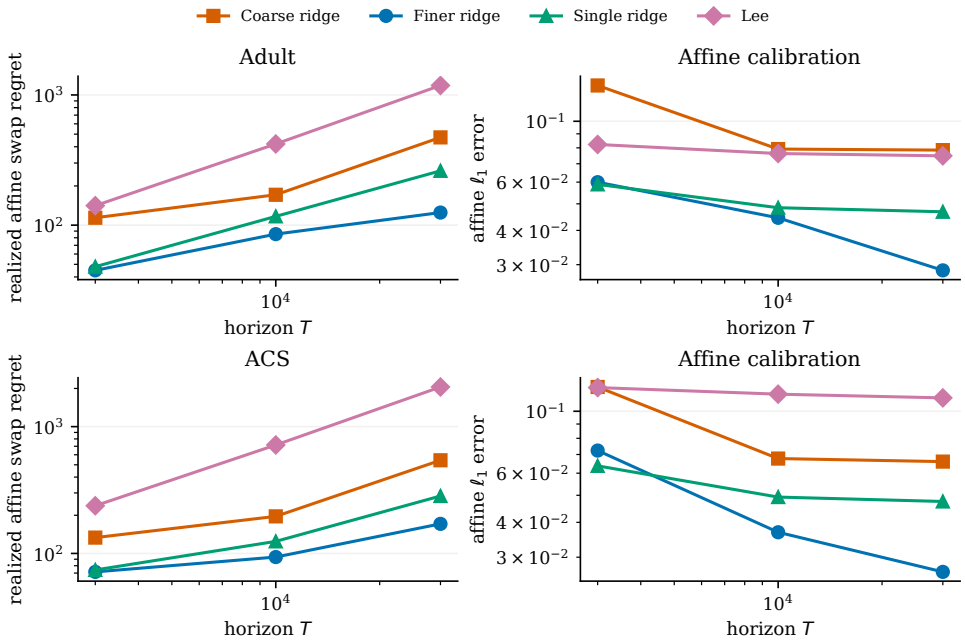

Figure 9: Comparisons on real data over 20 seeds. Left. Realized affine swap regret over prediction atoms. Right. Affine ℓ_1 calibration error over the same atoms. Rows correspond to Adult and ACSIncome. Lee's native objective is calibration for a finite family of groups.

10^{-8}. Both ridge and ONS loops use Numba with `fastmath=False`; timings exclude compilation. Table 9 reports all selected settings. Finer ONS reduces realized regret on all ten paired seeds in each of the six cells. Its affine ℓ_1 error is lower at the two larger horizons, but higher at $d = 10, T = 1000$. ONS and ridge exchange advantages across metrics and horizons.

Table 9: Weighted ONS and ridge on the same fixed streams and learner seeds. Medians over ten seeds for every selected configuration. All methods use comparator radius one and learner radius four. Timing excludes post hoc evaluation.

d	T	Base	Grid	N	S	P	M_1	M_2	Sec.
2	1,000	Ridge	coarse	3	20.85	19.68	0.1013	0.0127	0.01
2	1,000	Ridge	finer	4	14.66	14.06	0.0794	0.0087	0.01
2	1,000	ONS	coarse	3	24.36	23.43	0.1136	0.0157	0.01
2	1,000	ONS	finer	4	16.72	17.66	0.0815	0.0107	0.01
2	10,000	Ridge	coarse	4	115.64	119.87	0.0739	0.0070	0.07
2	10,000	Ridge	finer	7	37.22	38.93	0.0356	0.0018	0.08
2	10,000	ONS	coarse	4	68.90	68.35	0.0561	0.0044	0.13
2	10,000	ONS	finer	7	35.24	34.95	0.0339	0.0016	0.18
2	100,000	Ridge	coarse	5	676.67	671.71	0.0413	0.0042	0.74
2	100,000	Ridge	finer	14	83.85	82.64	0.0114	0.0002	1.08
2	100,000	ONS	coarse	5	360.45	359.49	0.0324	0.0016	1.55
2	100,000	ONS	finer	14	61.12	60.65	0.0128	0.0003	3.37
10	1,000	Ridge	coarse	2	37.78	37.15	0.0984	0.0291	0.01
10	1,000	Ridge	finer	3	23.36	22.21	0.1074	0.0131	0.01
10	1,000	ONS	coarse	2	41.16	41.29	0.1027	0.0312	0.01
10	1,000	ONS	finer	3	28.84	28.51	0.1215	0.0174	0.01
10	10,000	Ridge	coarse	3	192.85	191.13	0.1026	0.0115	0.08
10	10,000	Ridge	finer	5	71.71	72.03	0.0412	0.0027	0.09
10	10,000	ONS	coarse	3	197.32	197.51	0.1047	0.0125	0.15
10	10,000	ONS	finer	5	86.25	85.90	0.0477	0.0037	0.22
10	100,000	Ridge	coarse	4	1073.47	1068.85	0.0546	0.0051	0.79
10	100,000	Ridge	finer	10	174.30	172.95	0.0119	0.0003	0.99
10	100,000	ONS	coarse	4	659.16	657.81	0.0496	0.0040	1.76
10	100,000	ONS	finer	10	158.69	157.56	0.0172	0.0005	3.52

E Experimental Protocol

This appendix records the protocol behind the experiments section.

Instances (oblivious). Contexts are independent and uniform on the unit ball, generated by normalizing independent Gaussian directions and multiplying by $U^{1/d}$ for $U \sim \mathrm{Uniform}[0, 1]$. For the tested dimensions $d \geq 2$, let b be the unit normalization of $(1, -0.7, 0, \ldots, 0)$. Let a be a cyclic shift of b by one coordinate, add 0.8 to its third coordinate when $d > 2$, and normalize again. With $B = \max\{50, \mathrm{round}(T^{0.55})\}$ and $k_t = \lfloor (t-1)/B \rfloor$, set

$$\theta_t = (-1)^{k_t} b + 0.65a, \qquad y_t = \frac{1}{1 + \exp(-2.4\langle \theta_t, x_t \rangle - \nu_t)},$$
$$\nu_t = 0.35 \sin(4x_{t,1}) + 0.25 \,\mathrm{sign}(x_{t,2})|x_{t,1}|.$$

This logistic sigmoid produces soft labels. The entire sequence is fixed before learner randomization, and its block order is preserved.

Learner and master. Each bucket uses a weighted variant of online least-squares prediction (Azoury & Warmuth, 2001; Vovk, 2001), with unit regularization. Before round t, bucket i maintains $A_{t,i} = I + \sum_{s<t} P_s(p_i)\tilde{x}_s\tilde{x}_s^\top$ and $b_{t,i} = \sum_{s<t} P_s(p_i) y_s \tilde{x}_s$; the coefficient $A_{t,i}^{-1} b_{t,i}$ is projected in Euclidean norm to radius $4\tilde{R}$ before forming and clipping the prediction. The per-bucket predictions are rounded onto Z_N while preserving their means to form the transition kernel, and the master samples z_t from its stationary distribution. This radial projection of the ridge coefficient differs from the projection in the local matrix norm in ONS. Algorithm 1 uses ONS, and the ridge comparisons use the update specified here.

Evaluation. The realized and pseudo swap regret use the comparator ball of radius $\tilde{R} = \max(R, 1) = 1$. Per bucket, the inner supremum over $\mathcal{F}$ is the exact ℓ_2-ball-constrained least-squares value, solved to numerical tolerance. The maximum affine bucket moment uses the closed-form $\sup_{\|\theta\| \leq \tilde{R}}$ described in the experiments section.

Grids and sweep. In the implementation, $L_{\text{impl}} = d \log T + \log(T/0.05)$, with grid constant $c_{\text{grid}} = 1$ and $\delta = 0.05$. The localized grid is $N = \lceil (T/L_{\text{impl}})^{1/3} \rceil$ and the coarse grid is $N = \lceil (T/L_{\text{impl}})^{1/5} \rceil$, the resolution Luo et al. (2025) select for the same Blum and Mansour template in proving their $\widetilde{O}(T^{3/5} d^{2/5})$ bound (their Theorem 3). Both resolutions use our shared weighted-ridge implementation, isolating the choice of N. These practical grids use d, while Algorithm 1 uses $d + 1$ and a sufficiently large theoretical C_R. At $T = 30000$, Adult uses $N = 7$ and $N = 3$; at $T = 10^5, d = 2$, the synthetic arms use $N = 14$ and $N = 5$. We sweep $T \in \{10^3, 3 \cdot 10^3, 10^4, 3 \cdot 10^4, 10^5\}$ and $d \in \{2, 5, 10, 20\}$ with 20 seeds, plus a point in a third decade at $T = 10^6$ for $d = 2$ with 10 seeds. The multicalibration track sweeps $d \in \{2, 10\}$ with 20 seeds under both the oblivious labels and the adaptive rule $y_t = \text{clip}\big((1-\kappa)\mu_t + \kappa(1 - z_t)\big)$, $\kappa = 0.5$, where μ_t is the oblivious base mean and z_t the sampled prediction.

Sampling diagnostic with a fixed stream. For each $T \in \{10^3, 10^4, 10^5\}$ at $d = 2$, the synthetic sequence uses stream seed 0 and is held fixed while only the learner seed varies. The respective numbers of learner draws are 500, 300, and 120. Both envelope multipliers are calibrated using only the $T = 10^3$ draws and are not refit at the two larger horizons.

Real stream (Adult). The track on real data uses the Adult census dataset (Becker & Kohavi, 1996), 48,842 records. We take five standardized numeric features (age, education-num, capital-gain, capital-loss, hours-per-week), a binary sex indicator, and a one-hot race encoding, giving $d = 11$; each row is scaled to norm at most one and the binary income label is kept unchanged. For each seed we draw a uniformly shuffled prefix of length T; since $T \le 30 \cdot 10^3 < 48{,}842$, every horizon is a *single pass with no record reused*, and the whole stream is fixed before the learner's randomization. The two arms share the stream, the learner and the evaluator, changing only N. We sweep $T \in \{3 \cdot 10^3, 10^4, 3 \cdot 10^4\}$ with 20 seeds.

Decision loss on the real stream. The threshold-decision loss at cost ratio c is $\frac{1}{T} \sum_t \big[(1 - c)\, y_t \mathbf{1}\{z_t \le c\} + c\,(1 - y_t)\mathbf{1}\{z_t > c\}\big]$, evaluated at $c \in \{0.1, \ldots, 0.9\}$ using the sampled predictions z_t. Comparisons between arms are paired by seed and tested with an exact two-sided sign test.

Grid ablation. The resolution sweep multiplies the finer grid by $\{\frac{1}{4}, \frac{1}{2}, 1, 2, 4, 8, 16\}$ for $(d, T) \in \{2, 20\} \times \{10^4, 10^5\}$, adding factor 32 at $T = 10^4$. We set $N = \max\{1, \text{round}(f N_0)\}$, use 20 seeds per point, and hold both stream seed $10000d + 101T + s$ and learner seed s fixed across factors f. The master uses the explicit selection rule for the closed component in Appendix D.1.