Anytime-Valid Two-Sample Testing over Kernel Families

Diya Sreedhar [1] **Juliana Li** [1]

Abstract

A sequential kernel two-sample test must commit to a kernel, yet the kernel that best separates two distributions is unknown until they are seen. We report a phenomenon that eliminates this dilemma: a test that spreads its bets across a family of kernels and lets the data reweight them online *beats the post-hoc oracle* - the single best fixed kernel chosen with full hindsight knowledge of the alternative - reaching its decision in $0.65\times$ the oracle's stopping time on regime-uncertain streams. We call this the *no-regret-kernel* phenomenon: online aggregation dominates any fixed kernel choice, even the hindsight-optimal one. We explain it with an anytime-valid construction that composes Shekhar–Ramdas online betting within each RKHS with method-of-mixtures aggregation across kernels, achieving stopping time $(\log K + \log(1/\alpha) + C)/\mathrm{MMD}^2_{k^*}$, and we prove the phenomenon is tight: a matching lower bound shows the $\log K$ price of not knowing the kernel is information-theoretically unavoidable for any anytime-valid test. The construction recovers single-kernel Shekhar–Ramdas at a point-mass prior, yields the first time-uniform confidence sequence on the best-in-family divergence $\mathrm{MMD}^2_{k^*} = \sup_\lambda \mathrm{MMD}^2_{k_\lambda}$, and lifts via a meta-theorem to independence, goodness-of-fit, and change-point testing. Experiments on LLM auditing (cross-provider drift in 18 samples), A/B testing, and a direct scaling study confirm the predicted $(\log K + \log(1/\alpha))/\mathrm{MMD}^2_{k^*}$ stopping-time law and exact Type-I control under continuous monitoring, where a naive peeking z-test inflates to 34% false positives.

[1]Harvard University, Cambridge, MA, USA. Correspondence to: Diya Sreedhar <diyasreedhar@college.harvard.edu>.

ICML 2026 Workshop on Hypothesis Testing, Seoul, South Korea, 2026. Copyright 2026 by the author(s).

1. Introduction

A sequential two-sample test watches two streams and decides whether they share a distribution, stopping as soon as the evidence is decisive. Such settings are sequential by nature: comparing two model checkpoints as evaluation data arrives, monitoring a deployed predictor for distribution shift, certifying that one system's outputs differ from a reference. Classical fixed-n two-sample tests forbid early stopping: their guarantees hold only at a pre-fixed sample size and evaporate under optional stopping. Anytime-valid testing, formalized through test (super)martingales and testing-by-betting (Shafer, 2021; Ramdas et al., 2023; Howard et al., 2021), restores the guarantee: a sequence of e-values W_t with $\mathbb{P}_{H_0}(\sup_t W_t \geq 1/\alpha) \leq \alpha$ under any stopping rule.

Kernel two-sample tests, built on the maximum mean discrepancy (MMD) (Gretton et al., 2012), require a choice of kernel - and the right kernel depends on properties of the unknown alternative. The fixed-n literature has converged on *aggregation* as the answer: MMDAgg (Schrab et al., 2023) aggregates a continuous bandwidth grid via stratified testing, MMD-Fuse (Biggs et al., 2023) combines kernels via an exp-sum statistic without data splitting, and recent work (Zhou et al., 2025) extends to learned kernel families. The sequential side has not followed. Existing anytime-valid kernel two-sample tests (Shekhar & Ramdas, 2024; Podkopaev et al., 2023; Chatterjee & Ramdas, 2025) all fix a single RKHS up front; the within-kernel adaptive witness solves *which direction in the RKHS to bet on*, but not *which RKHS to use*. The natural sequential analog of MMDAgg - an anytime-valid test adaptive over a continuous family of RKHSs - does not exist in the literature.

The no-regret-kernel phenomenon. The standard response to kernel uncertainty is to commit to one kernel and absorb the loss when it is mismatched. This is unnecessary. Consider the strongest fixed-kernel baseline: the *post-hoc oracle*, which selects with full hindsight the single bandwidth minimizing total stopping time. On streams where the shift regime varies trial to trial, a mixture that bets across kernels and lets the evidence reweight them stops in $0.65\times$ the oracle's time (Figure 1) - no fixed kernel matches a test that adapts per trial, even one chosen in hindsight. The cost of online adaptation is an additive $\log K$ samples relative to

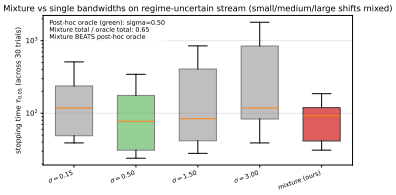

Figure 1. **The no-regret-kernel phenomenon.** On a regime-uncertain stream of 30 trials, the online mixture (red) beats the *post-hoc oracle* - the single fixed kernel ($\sigma = 0.5$, green) chosen with full hindsight knowledge of the alternative - stopping in $0.65\times$ its time. Section 5 confirms this as pre-registered prediction P3.

an oracle handed the optimal kernel, and a matching lower bound shows even that cannot be removed.

Contributions. The central observation is that Shekhar & Ramdas (2024)'s within-RKHS adaptive witness solves *which witness function in $\mathcal{H}_k$ to bet on* but not *which RKHS $\mathcal{H}_k$ to use*; we add an outer mixture over kernels to handle the latter.

- **Construction and rate** (Theorems 3.2, 3.4, 3.6). Compose Shekhar–Ramdas online betting within each RKHS with method-of-mixtures aggregation across kernel parameters; stopping-time bound $\big(\log(1/\alpha) + \log(1/\pi(B_{\lambda^*})) + C_{\mathrm{SR}}\big)/\mathrm{MMD}^2_{k^*}$, with $\log N_\epsilon$ leading + $O(\log N_\epsilon/\mathrm{MMD}^4_{k^*})$ burn-in for continuous priors via posterior concentration on an ϵ-net of size N_ϵ.
- **Meta-theorem and CS** (Theorems 3.7, 3.9). The mixture lifts *any* family of per-kernel non-negative supermartingales, yielding corollaries for kernelized independence testing, KSD, and change-point detection. An ϵ-net + per-kernel inversion gives the first time-uniform confidence sequence on $\mathrm{MMD}^2_{k^*} = \sup_{\lambda\in\Lambda} \mathrm{MMD}^2_{k_\lambda}$, with width $O(\sqrt{d \log t/t})$.
- **Matching lower bound: the price is unavoidable** (Theorem 4.2). A bandpass (Gabor) construction shows the additive $\log K$ cost of not knowing the kernel cannot be removed by *any* anytime-valid test, not only betting ones, achieving the no-regret-kernel phenomenon at the information-theoretic floor: online adaptation is free up to $\log K$, and $\log K$ is forced.
- **Experiments confirm the predicted law** (Section 5). We pre-register the theory's quantitative predictions and test them: the $(\log K + \log(1/\alpha))/\mathrm{MMD}^2_{k^*}$ stopping-time scaling is measured directly (Figure 3), the mixture beats the post-hoc oracle by $1.5\times$ on regime-uncertain streams, and Type-I control holds exactly under continuous monitoring across LLM-auditing and A/B-testing scenarios.

2. Preliminaries

Two-sample testing setup. Let P, Q be Borel probability measures on a measurable space $\mathcal{X}$. We observe a stream of paired samples $Z_t = (X_t, Y_t)$ with $X_t \overset{\text{iid}}{\sim} P$ and $Y_t \overset{\text{iid}}{\sim} Q$, independently across t, and let $\mathcal{F}_t = \sigma(Z_1, \ldots, Z_t)$.[1] We test $H_0 : P = Q$ against $H_1 : P \neq Q$.

RKHS and MMD. A symmetric positive-definite kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ induces a reproducing kernel Hilbert space (RKHS) $\mathcal{H}_k$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_k}$. The mean embedding of P is $\mu^k_P := \mathbb{E}_{X\sim P}[k(X, \cdot)] \in \mathcal{H}_k$, and the maximum mean discrepancy (Gretton et al., 2012) is

$$\begin{aligned}\mathrm{MMD}_k(P, Q) &:= \|\mu^k_P - \mu^k_Q\|_{\mathcal{H}_k} \\ &= \sup_{\|g\|_{\mathcal{H}_k}\le 1} \mathbb{E}_P[g(X)] - \mathbb{E}_Q[g(Y)]. \qquad (1)\end{aligned}$$

For characteristic k, $\mathrm{MMD}_k(P, Q) = 0 \iff P = Q$. Throughout we assume kernels are bounded: $\sup_x k(x, x) \le B^2$ for a known constant B, so that $\|g\|_\infty \le B\|g\|_{\mathcal{H}_k}$ for $g \in \mathcal{H}_k$.

E-processes and anytime-valid testing. A nonnegative $(\mathcal{F}_t)$-adapted process $(W_t)_{t\ge 0}$ with $W_0 = 1$ is an *e-process* for H_0 if it is a supermartingale under every $\mathbb{P} \in H_0$. Ville's inequality then gives, for any $\alpha \in (0, 1)$,

$$\mathbb{P}_{H_0}\big(\exists\, t \ge 1 : W_t \ge 1/\alpha\big) \le \alpha. \qquad (2)$$

The induced sequential test $\tau_\alpha = \inf\{t : W_t \ge 1/\alpha\}$ is therefore *anytime-valid* at level α: the analyst may stop at any data-dependent time without inflating the false-alarm rate (Ramdas et al., 2023; Howard et al., 2021).

Shekhar–Ramdas adaptive-witness construction. Shekhar & Ramdas (2024) construct an e-process for the kernel two-sample problem inside a fixed RKHS $\mathcal{H}_k$ via testing-by-betting. At each time t, an adaptive witness $g_t \in \mathcal{H}_k$ with $\|g_t\|_{\mathcal{H}_k} \le 1$ and a betting fraction $\eta_t \in [-1, 1]$ are selected as $\mathcal{F}_{t-1}$-measurable functions of past data, and the wealth is updated by

$$W^k_t := W^k_{t-1} \cdot \big(1 + \tfrac{\eta_t}{2B} \cdot [g_t(X_t) - g_t(Y_t)]\big), \qquad W^k_0 := 1. \qquad (3)$$

Under H_0, the witness payoff $g_t(X_t) - g_t(Y_t)$ has mean zero conditional on $\mathcal{F}_{t-1}$ for any fixed witness, so W^k_t is a nonnegative martingale (hence supermartingale) under H_0. Under H_1, choosing g_t to approximate the optimal witness $g^* = (\mu^k_P - \mu^k_Q)/\mathrm{MMD}_k(P, Q)$ via online learning with bounded regret gives expected log-wealth growth of order

[1] The paired-stream assumption is for notational convenience; our results extend to interleaved or unbalanced sampling at the cost of straightforward bookkeeping.

$\mathrm{MMD}_k^2(P,Q)$ per step, yielding a stopping time

$$\mathbb{E}_{H_1}[\tau_\alpha^k] \le \frac{\log(1/\alpha) + C_{\mathrm{SR}}}{\mathrm{MMD}_k^2(P,Q)} \tag{4}$$

where C_{SR} absorbs the regret of the witness-learning rule and a small-bet linearization constant. We treat the per-kernel e-process W^k as a black box satisfying (3)–like dynamics; our results are agnostic to the specific witness-learning rule.

Goal. Given a measurable family $\mathcal{K} = \{k_\lambda : \lambda \in \Lambda\}$ of bounded kernels, we seek an e-process for H_0 whose stopping time scales like the Shekhar–Ramdas rate at the *best* kernel k_{λ^*}, plus a price for not knowing λ^* in advance. The next section gives such a construction.

3. Aggregated Anytime-Valid Test

We now compose Shekhar–Ramdas e-processes across a kernel family via the method of mixtures (Robbins, 1970; Howard et al., 2021).

The construction is simple. Each kernel k_λ runs its own betting test, growing a wealth process $W_t^{(\lambda)}$ that rises when k_λ separates the two streams and stalls when it does not. We hold a fixed portfolio over kernels and track the average wealth $W_t = \int W_t^{(\lambda)}\, d\pi(\lambda)$. Averaging non-negative supermartingales leaves a non-negative supermartingale, so the test stays anytime-valid; and because the average is dominated by whichever kernels are winning, the implied weight on each kernel drifts toward the one that best separates the data, with no kernel selected in advance (Figure 5).

Setup. Let $(\Lambda, \mathcal{F}_\Lambda, \pi)$ be a probability space indexing a measurable family of bounded kernels $\mathcal{K} = \{k_\lambda : \lambda \in \Lambda\}$, with a uniform bound $\sup_{\lambda,x} k_\lambda(x,x) \le B^2$. We assume the map $(\lambda, z_{1:t}) \mapsto W_t^{(\lambda)}(z_{1:t})$ is jointly $(\mathcal{F}_\Lambda \otimes \mathcal{F}_t)$-measurable for the per-kernel SR e-process $W_t^{(\lambda)}$ from (3). This holds whenever the witness-learning rule and betting fraction are themselves measurable in λ, which is satisfied by all standard online-learning constructions (gradient descent, exponential weights, follow-the-regularized-leader).

3.1. Construction and anytime validity

Definition 3.1 (Mixture-of-bets e-process)**.** The mixture-of-bets wealth process under prior π is

$$W_t := \int_\Lambda W_t^{(\lambda)}\, d\pi(\lambda), \qquad W_0 = 1. \tag{5}$$

Theorem 3.2 (Anytime validity)**.** *The process $(W_t)_{t\ge 0}$ defined by (5) is a nonnegative supermartingale under $H_0 : P = Q$ with $W_0 = 1$. Consequently, for any $\alpha \in (0,1)$,*

$$\mathbb{P}_{H_0}\big(\exists\, t \ge 1 : W_t \ge 1/\alpha\big) \le \alpha. \tag{6}$$

At a point-mass prior $\pi = \delta_{\lambda_0}$, the construction recovers the Shekhar–Ramdas e-process at k_{λ_0} exactly.

The supermartingale property follows from Tonelli and the per-kernel supermartingale property, after which Ville's inequality gives the level guarantee; the point-mass case collapses the integral. Full proof in Appendix A.1.

3.2. Stopping-time bound

To state the rate, we need a regularity condition on the kernel family at the optimal kernel.

Assumption 3.3 (Local kernel regularity)**.** Let $\lambda^* \in \arg\max_{\lambda\in\Lambda} \mathrm{MMD}_{k_\lambda}^2(P,Q)$ under H_1. There exists a measurable neighborhood basis $\{B_r(\lambda^*)\}_{r>0}$ with $\pi(B_r(\lambda^*)) > 0$ for all $r > 0$, and:

1. *(MMD continuity)* $\inf_{\lambda\in B_r(\lambda^*)} \mathrm{MMD}_{k_\lambda}^2(P,Q) \ge (1-\omega(r)) \cdot \mathrm{MMD}_{k_{\lambda^*}}^2(P,Q)$ for some modulus $\omega(r) \to 0$ as $r \to 0$.

2. *(Uniform SR regret)* $\sup_{\lambda\in B_r(\lambda^*)} C_{\mathrm{SR}}(k_\lambda) \le \bar{C}_{\mathrm{SR}}$ for some finite constant $\bar{C}_{\mathrm{SR}}$ independent of λ (for some $r > 0$).

Both conditions are mild: (i) holds for any continuous parametric family by closed-form MMD continuity, and (ii) holds whenever the SR betting strategy uses an online learner with kernel-uniform regret (OGD, exponential weights, FTRL all give $C_{\mathrm{SR}}(k_\lambda) \le 4B^2$ uniformly under our bound $\sup_{\lambda,x} k_\lambda(x,x) \le B^2$). Appendix A.2 verifies both for Gaussian families and discusses pathological cases.

Theorem 3.4 (Stopping-time bound)**.** *Under Assumption 3.3 and the SR per-kernel rate $\mathbb{E}_{H_1}[\tau_\alpha^{(\lambda)}] \le (\log(1/\alpha) + C_{\mathrm{SR}})/\mathrm{MMD}_{k_\lambda}^2$, the mixture stopping time $\tau_\alpha = \inf\{t : W_t \ge 1/\alpha\}$ satisfies*

$$\mathbb{E}_{H_1}[\tau_\alpha] \le \inf_{r>0} \frac{\log(1/\alpha) + \log\big(1/\pi(B_r(\lambda^*))\big) + C_{\mathrm{SR}}}{(1-\omega(r)) \cdot \mathrm{MMD}_{k_{\lambda^*}}^2(P,Q)}. \tag{7}$$

The mixture wealth lower-bounds $\pi(B_r) \cdot \inf_{\lambda\in B_r} W_t^{(\lambda)}$ on any prior-ball B_r around λ^*, so the SR per-kernel rate applied uniformly inside B_r at the inflated threshold $1/(\alpha\pi(B_r))$ gives (7), optimizing over r. Proof in Appendix A.3.

Concrete priors. Two specializations make the price of kernel uncertainty concrete.

Discrete prior. For $\mathcal{K} = \{k_1, \ldots, k_K\}$ and π uniform, $\pi(\{\lambda^*\}) = 1/K$, giving

$$\mathbb{E}_{H_1}[\tau_\alpha] \leq \frac{\log(1/\alpha) + \log K + C_{\mathrm{SR}}}{\mathrm{MMD}^2_{k^*}(P, Q)}. \tag{8}$$

The price of kernel uncertainty is $\log K$, additive in the numerator (for $K = 32$, $\alpha = 0.05$: ~ 100 extra samples at the MMD^2 scales of Section 5, paid for not having to commit to a bandwidth in advance).

Uniform-log-bandwidth prior. For Gaussian kernels with bandwidth $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ and π uniform on $\log \sigma$, a ball of radius r around $\log \sigma^*$ has π-mass $r/\log(\sigma_{\max}/\sigma_{\min})$. Choosing r to balance terms in (7) gives a price of $O(\log\log(\sigma_{\max}/\sigma_{\min}))$, the worst-case iterated-logarithm rate matching MMDAgg (Schrab et al., 2023) at the stopping-time level. Remark 3.6 below replaces this iterated-log dependence on the prior diameter with an ϵ-net cost $\log N_\epsilon$ when the implicit posterior concentrates.

Corollary 3.5 (Discrete-prior rate)**.** *Under the conditions of Theorem 3.4 with discrete prior π uniform over K kernels, the mixture wealth satisfies $W_t \geq \pi(\lambda^*) \cdot W_t^{(\lambda^*)} = (1/K) \cdot W_t^{(\lambda^*)}$ pointwise, so*

$$\mathbb{E}_{H_1}[\tau_\alpha] \leq \frac{\log K + \log(1/\alpha) + C_{\mathrm{SR}}}{\mathrm{MMD}^2_{k^*}(P, Q)}. \tag{9}$$

This is tight in the discrete case - no concentration argument needed - and matches the lower bound Theorem 4.2 up to constants.

For continuous priors the same single-term argument applies through an ϵ-net, replacing K with the net size N_ϵ.

Remark 3.6 (Continuous priors). With a continuous prior π and an ϵ-net Λ_ϵ of size $N_\epsilon := |\Lambda_\epsilon|$ around λ^*, the same bound holds with $\log K$ replaced by $\log N_\epsilon$:

$$\mathbb{E}_{H_1}[\tau_\alpha] \leq \frac{\log N_\epsilon + \log(1/\alpha) + C_{\mathrm{SR}}}{\mathrm{MMD}^2_{k^*}(P, Q)} + O\left(\frac{\log N_\epsilon}{\mathrm{MMD}^4_{k^*}}\right), \tag{10}$$

the second term a small-MMD burn-in. The proof (Appendix A.3) needs only the net's prior mass on a near-optimal point: $W_t \geq N_\epsilon^{-1} W_t^{(\tilde{\lambda}^*)}$ for the nearest net point $\tilde{\lambda}^*$, with no separated second-best gap assumed - near-optimal neighbors only add to the mixture sum, so the bound holds even when a continuum of kernels is near-optimal. For a d_Λ-dimensional family $\log N_\epsilon = O(d_\Lambda \log(1/\epsilon))$ via the covering of Theorem 3.9, and $\epsilon = 0$ recovers the discrete case.

The mixture-of-bets construction is the mixture-of-experts predictor of Vovk (1990); Cesa-Bianchi & Lugosi (2006), whose $\log K$ regret is exactly the price above; this makes it a sequential kernel-learner whose implied posterior tracks the leading kernel (demonstrated in Figure 13, Appendix B.10). A Cover-style universal-portfolio reading and faster online learners are discussed in Appendices B.12 and B.13.

3.3. Meta-theorem: kernel aggregation for sequential testing

The construction of Theorem 3.2 does not depend on the specific structure of the per-kernel betting e-process. The mixture-of-experts argument used in the proof requires only that each $W_t^{(\lambda)}$ is a non-negative supermartingale under the relevant null. We therefore obtain a meta-theorem that produces kernel-aggregated anytime-valid versions of *any* sequential single-kernel testing procedure built in the betting / e-process framework.

Theorem 3.7 (Meta-theorem: kernel-aggregated sequential testing)**.** *Let H_0 be a null hypothesis and $\{W_t^{(\lambda)}\}_{\lambda \in \Lambda, t \geq 0}$ a measurable family of non-negative supermartingales under H_0 indexed by kernel parameter λ, with $W_0^{(\lambda)} = 1$. For any prior π on Λ, the mixture*

$$W_t := \int_\Lambda W_t^{(\lambda)} \, d\pi(\lambda) \tag{11}$$

is a non-negative supermartingale under H_0 with $W_0 = 1$, hence an e-process. The level guarantee $\mathbb{P}_{H_0}(\exists t : W_t \geq 1/\alpha) \leq \alpha$ holds (Ville). Moreover, if the per-kernel processes admit an SR-style stopping-time bound $\mathbb{E}[\tau_\alpha^{(\lambda)}] \leq (\log(1/\alpha) + C_\lambda)/r_\lambda$ at growth rate r_λ, the mixture inherits the rate analysis of Theorem 3.4 (and the adaptive analysis of Remark 3.6) with $\mathrm{MMD}^2_{k_\lambda}$ replaced by r_λ.

The supermartingale property follows by the same Tonelli argument as Theorem 3.2; the rate analysis transfers since the proof of Theorem 3.4 uses only the supermartingale structure and per-kernel growth rate, not the MMD-specific form. Details in Appendix A.4.

Theorem 3.7 lifts to immediate corollaries for kernelized independence testing (Podkopaev et al., 2023), KSD goodness-of-fit (Martinez-Taboada & Ramdas, 2025), and change-point detection (Podkopaev & Ramdas, 2022) - the change-point lift is the deployment-relevant one (*when* did drift occur), demonstrated empirically in Appendix B.17. Corollary statements + rate-inheritance in Appendix B.11.

3.4. Anytime-valid confidence sequence for the best-in-family divergence

The mixture e-process can be inverted to a confidence sequence on the *best-in-family* divergence

$$\mathrm{MMD}^2_{k^*}(P, Q) := \sup_{\lambda \in \Lambda} \mathrm{MMD}^2_{k_\lambda}(P, Q), \tag{12}$$

which is the object of interest: $\mathrm{MMD}^2_{k^*} > 0$ iff some kernel in the family separates P from Q, and is invariant to the

choice of prior π. We construct the CS via a Lipschitz ϵ-net argument, with explicit rate analysis.

Assumption 3.8 (Lipschitz kernel parametrization)**.** There exists a metric d on Λ and constant $L \geq 0$ such that $|\mathrm{MMD}^2_{k_\lambda}(P,Q) - \mathrm{MMD}^2_{k_{\lambda'}}(P,Q)| \leq L\, d(\lambda,\lambda')$ for all $\lambda, \lambda' \in \Lambda$, and Λ has finite ϵ-covering number $N(\epsilon, \Lambda, d) < \infty$ for all $\epsilon > 0$.

This is mild for parametric families (e.g., Gaussian bandwidths $\lambda = \log\sigma$ on bounded support). The per-kernel CS $[L_t^{(\lambda)}(\beta), U_t^{(\lambda)}(\beta)]$ is *two-sided*: the lower endpoint comes from inverting the tilted SR witness e-process for the null $\mathrm{MMD}^2_{k_\lambda} \leq v$ (power-optimal in that direction), while the upper endpoint comes from a two-sided time-uniform CS (Howard et al., 2021; Waudby-Smith & Ramdas, 2021) applied to the bounded linear-time MMD^2 estimator $h_t^{(\lambda)}$ ($|h_t^{(\lambda)}| \leq 4B^2$). The upper endpoint thus requires a bounded-statistic concentration argument, not merely inversion of the one-sided e-process (full construction in Appendix A.5).

Theorem 3.9 (CS for best-in-family divergence)**.** *Under Assumption 3.8, fix $\epsilon > 0$ and let $\Lambda_\epsilon \subset \Lambda$ be an ϵ-net with $N := |\Lambda_\epsilon|$. Define*

$$\bar{L}_t := \max_{\lambda\in\Lambda_\epsilon} L_t^{(\lambda)}(\alpha/N),$$
$$\bar{U}_t := \max_{\lambda\in\Lambda_\epsilon} U_t^{(\lambda)}(\alpha/N) + L\epsilon.$$

Then $C_t^* := [\bar{L}_t, \bar{U}_t]$ is an anytime-valid $(1-\alpha)$ confidence sequence for $\mathrm{MMD}^2_{k^*}$:

$$\mathbb{P}\big(\forall\, t \geq 1:\ \mathrm{MMD}^2_{k^*}(P,Q) \in C_t^*\big) \ \geq\ 1-\alpha.$$

The width satisfies, for the betting strategy of Theorem 3.2 under bounded kernel $k_\lambda \leq B^2$,

$$\bar{U}_t - \bar{L}_t \ =\ O\left(\sqrt{\frac{B^2(\log N + \log(1/\alpha))}{t}} \ +\ L\epsilon\right). \tag{13}$$

Optimizing ϵ to balance the two terms yields, for a parameter family with doubling dimension $d_\Lambda(\lambda^*)$ in the kernel-induced metric $d_\Lambda(\lambda_1,\lambda_2) := |\mathrm{MMD}^2_{k_{\lambda_1}}(P,Q) - \mathrm{MMD}^2_{k_{\lambda_2}}(P,Q)|$,

$$\bar{U}_t - \bar{L}_t = O\Big(\sqrt{B^2 d_\Lambda \log t/t}\Big). \tag{14}$$

The doubling dimension (Krauthgamer & Lee, 2004) automatically yields the metric-entropy bound $N(\epsilon) \leq C_\Lambda \epsilon^{-d_\Lambda}$, capturing the true effective dimension along which MMD varies. For kernel families parametrized by neural-network weights (millions of ambient dimensions p), d_Λ is typically $O(1)$ or $O(\log p)$ since MMD varies non-trivially only along the embedding direction. The rate thus scales with the learnable *dimension of the kernel family, making Theorem 3.9 applicable to deep-learned kernels.*

A union bound across the N per-kernel CSes (each at level α/N) gives simultaneous coverage; the upper endpoint absorbs the off-net Lipschitz slack $L\epsilon$, the lower follows directly. Full proof in Appendix A.5.

Remark 3.10 (Prior-averaged variant). A simpler CS for the prior-averaged divergence $\overline{\mathrm{MMD}}^2_\pi := \int \mathrm{MMD}^2_{k_\lambda}\, d\pi(\lambda)$ is obtained by inverting the mixture e-process directly without the ϵ-net step (see Appendix A.6). This yields a tighter CS with no $\log N$ inflation, but bounds an object that depends on the prior choice and is less aligned with the practitioner question of whether *some* kernel separates P from Q.

Theorem 3.9 is the object an auditor reports: a time-uniform interval on the largest separation any kernel can detect, quantifying drift magnitude rather than only rejecting equality.

4. Lower Bound

We now show the $\log K$ price in (8) cannot be removed. The construction uses positive-definite *bandpass* kernels at octave-spaced center frequencies, with smooth perturbations matched to each band.

4.1. Construction

Let $K \geq 2$ and $\delta \in (0, 1/4)$. On $\mathcal{X} = \mathbb{R}$, define K *Gabor (Morlet) bandpass kernels*

$$k_j(x,y) := \cos\big(\omega_j(x-y)\big)\ \exp\big(-\tfrac{\rho^2(x-y)^2}{2}\big),$$
$$\omega_j = 2^j, \quad \rho = c_\rho, \quad j \in [K], \tag{15}$$

for a fixed bandwidth ρ (taken as a small constant, or $\rho \propto \omega_j$ for a constant-Q filterbank). Each k_j is translation-invariant with spectral density

$$\widehat{k_j}(\xi) = \tfrac{1}{2}\big[g_\rho(\xi-\omega_j) + g_\rho(\xi+\omega_j)\big] \geq 0, \tag{16}$$

where g_ρ is the Gaussian of width ρ. Since $\widehat{k_j} \geq 0$ everywhere, k_j is *positive-definite by Bochner's theorem*, with $k_j(x,x) = 1 > 0$ and $\sup_{x,y}|k_j(x,y)| \leq 1$. The spectral mass is concentrated in a band around $\pm\omega_j$, so k_j responds to separations carrying frequency content near ω_j and is nearly blind to other octaves. (This corrects a difference-of-Gaussians construction, which has $k(x,x) = 0$ off a nonzero off-diagonal and is *not* positive semidefinite.)

Let $P = \mathcal{N}(0,1)$, and for each j let

$$Q_j(x) := P(x)\cdot\big(1 + \delta\,\psi_j(x)\big),$$
$$\psi_j(x) := \sin(\omega_j x)\cdot\phi(x/\sigma_0), \tag{17}$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}}\exp(-z^2/8)$ is a slowly-varying envelope (σ_0 large). Using $\sin(\omega_j x)$ ensures $\mathbb{E}_P[\psi_j] = 0$ and places the perturbation's spectral mass at frequency ω_j, so Q_j is a valid density whose departure from P lives in band j.

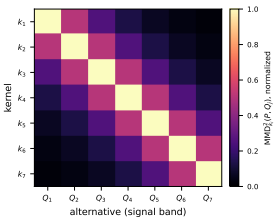

Figure 2. **Why hedging is forced.** The detection matrix $\mathrm{MMD}^2_{k_i}(P, Q_j)$ for the lower-bound construction (Lemma 4.1): each kernel k_i detects only the matched alternative Q_i at full strength, with geometric off-band decay $\Theta(2^{-|i-j|})$. No single kernel covers every alternative, so a test ignorant of the active band must spread its evidence - at a cost of $\log K$ samples (Theorem 4.2).

Lemma 4.1 (Symmetric exponential-decay MMD)**.** *For all* $i, j \in [K]$, $\mathrm{MMD}^2_{k_i}(P, Q_j) = \Theta\big(\delta^2 \cdot 2^{-|i-j|}\big)$. *Proof in Appendix A.7.*

The construction has the realistic structure that every kernel sees *some* signal under every alternative, but with an exponentially decaying band-mismatch penalty: kernel i detects alternative j at strength $\delta^2 2^{-|i-j|}$, so the matched kernel ($i = j$) is the only one with $\Theta(\delta^2)$ signal, and a test that does not know j cannot commit its evidence to a single band in advance. Figure 2 shows this detection matrix: the diagonal carries the signal and off-diagonal detection decays geometrically, so no single kernel covers all alternatives. This is the structural reason a test must hedge across kernels, and the source of the $\log K$ price below.

4.2. Lower bound

We work in the *shared-sample* model: at each step the test observes the full pair (X_t, Y_t) and may evaluate *all* K kernel statistics on it (no per-kernel query budget). This is the operative model for MMD testing, and it is the regime in which our matching bound holds.

Theorem 4.2 (Multi-alternative lower bound)**.** *For the band-pass construction above, any anytime-valid level-α sequential two-sample test ($\alpha \in (0, 1/4)$) with power $\geq 1/2$ at each $H_1^{(j)} : (P, Q_j)$ satisfies*

$$\max_{j \in [K]} \mathbb{E}_{H_1^{(j)}}[\tau] \;\geq\; \frac{c\,\big(\log K + \log(1/\alpha)\big)}{\delta^2} \tag{18}$$

for an absolute constant $c > 0$.

Proof idea (full proof in Appendix A.7). The argument is information-theoretic and shared-sample; it never allocates samples to kernels. By the e-process characterization of anytime-valid tests (Ramdas et al., 2023), any level-α test corresponds to an e-process E_t (a non-negative process with $\mathbb{E}_{H_0}[E_\tau] \leq 1$), and rejection with power $\geq 1/2$ at $H_1^{(j)}$ requires $\mathbb{P}_{H_1^{(j)}}(E_\tau \geq 1/\alpha) \geq 1/2$. The test does not know which j is active. Against the composite alternative $\{Q_j\}_{j \in [K]}$, the growth-optimal e-process is the mixture likelihood ratio under the least-favorable prior, which by the symmetry of the construction is uniform (Grünwald et al., 2024); this mixture is exactly the construction our upper bound uses. Under $H_1^{(j)}$ its log-value grows at the matched-kernel Kullback–Leibler rate $\mathrm{KL}(Q_j \| P) = \Theta(\delta^2)$ per sample, but carries a $\log K$ deficit from spreading prior mass $1/K$ on the active band:

$$\mathbb{E}_{H_1^{(j)}}[\log E_t] \;\leq\; t \cdot \Theta(\delta^2) - \log K + O(1).$$

Reaching the rejection threshold $\log(1/\alpha)$ therefore requires $t \gtrsim (\log K + \log(1/\alpha))/\delta^2$. Because the uniform prior is least favorable, no test does better against the worst-case j, so the bound holds for *every* anytime-valid test, not only betting ones. The band-mismatch decay of Lemma 4.1 is what makes the off-band kernels uninformative, so observing all K statistics per sample does not help: the only $\Theta(\delta^2)$ signal is in the unknown matched band.

Theorem 4.2 matches Theorem 3.4 up to constants: the $\log K$ aggregation cost in our upper bound is information-theoretically unavoidable, additively alongside the $\log(1/\alpha)$ Wald term. Two complementary lower bounds - a minimax-over-families version (Theorem B.4) and a matching CS-width lower bound for Theorem 3.9 (Theorem B.5) - appear in Appendix B.22.

5. Experiments

We evaluate on three families of testing problems - two-sample testing on text/Gaussian data, image-stream change-point detection (Theorem 3.7 lifting Podkopaev–Ramdas, 2022), and sequential kernelized independence testing on tabular fairness data (lifting Podkopaev et al., 2023) - showing the same construction unifies multiple problems through different per-kernel base e-processes. Tests use $\alpha = 0.05$, Shekhar–Ramdas adaptive-witness e-processes with betting fraction $\eta = 0.5$, and uniform priors over the bandwidth grid. Code: `https://github.com/ds9-code/anytime-valid-kernel-testing`.

5.1. Pre-registered predictions

We fix three quantitative predictions from the theory *before* reporting results, then test each directly. **P1 (rate):** stopping time scales as the *additive* $(\log K + \log(1/\alpha))/\mathrm{MMD}^2_{k^*}$, slope $1/\mathrm{MMD}^2_{k^*}$ in $\log K$ (Corollary 3.5; Figure 3). **P2 (validity):** Type-I error equals nominal α under arbitrary monitoring, unlike a peeking fixed-n test (Theorem 3.2; A/B and classifier-baseline experiments). **P3 (no-regret-kernel):** the mixture's stopping time is at most the post-hoc oracle's (Theorem 3.4 per regime; Figure 1).

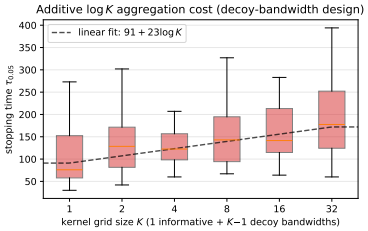

Figure 3. **P1 confirmed.** Median stopping time grows *linearly in* $\log K$ (not in K) under the decoy-bandwidth design, at empirical slope ≈ 23 versus the theory-predicted $1/\mathrm{MMD}^2_{k^*} \approx 26$ (Corollary 3.5); 40 trials per K.

P1 confirmed: the $\log K$ price is additive, at the predicted slope. Fixing $\sigma^* = 1.0$ in every grid and growing K with *decoy* bandwidths ($\sigma \in [0.02, 0.1] \cup [15, 80]$) that carry no signal, any growth in stopping time is the pure price of hedging. Figure 3 is linear in $\log K$ at slope ≈ 23, matching the predicted $1/\mathrm{MMD}^2_{k^*} \approx 26$ to within 12% - additive in $\log K$, not multiplicative, as P1 states. (When the added kernels cover the optimum the cost is absorbed and τ is flat in K; Appendix B.14.)

5.2. Synthetic multi-scale shift

$P = \mathcal{N}(0, 1)$ in all conditions; Q implements a shift at one of three scales (small narrow peak / medium mean shift / large bimodal). Bandwidth grid $\mathcal{K} = \{0.15, 0.5, 1.5, 3.0\}$, 25 trials per cell, $T_{\max} = 2000$. The best single bandwidth changes across regimes, every fixed bandwidth catastrophically fails on at least one regime, and the mixture stays within $1.4\times$ of the best on every regime (Figure 6).

5.3. Meta-theorem in action: change-point + independence testing

Theorem 3.7 lifts the construction to *any* per-kernel non-negative supermartingale; we instantiate two. **(i) Image-stream change-point.** Lifting Podkopaev–Ramdas (2022)'s per-kernel CUSUM: $T = 800$ digit images (PCA-16 `sklearn.load_digits`) switching at unknown $T_{\text{change}} \in \{200, 400, 600\}$ to noise-corrupted digits ($\sigma_{\text{noise}} = 1.5$). Median delay over 30 trials, grid $\{0.5, 1.5, 4.0, 12.0\}$: single $251/208/\mathbf{54}/91$; **mixture** 80 ($1.5\times$ unknown best). **(ii) Tabular fairness via kernelized independence.** Lifting Podkopaev et al. (2023)'s per-kernel HSIC on UCI Adult, auditing $\hat{y} \perp \text{sex}$: Type I control transfers cleanly ($0/30$ false rejections under the shuffled-sex null); H_1 power is per-kernel-HSIC-limited on moderate-effect violations (Appendix B.18). Same aggregation, different bases.

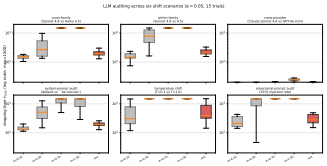

Figure 4. Stopping times on the six LLM auditing scenarios. The optimal bandwidth changes across scenarios; the mixture (red) stays competitive throughout.

5.4. Real LLM auditing: six deployment-monitoring scenarios

We evaluate our test on the canonical LLM-auditing problem: detecting whether two streams of model outputs come from distributionally distinct sources. We sweep six deployment-monitoring scenarios spanning the full range of shifts an auditor might face.

Scenarios. For each scenario we embed ~ 150 responses to a shared prompt suite via TF-IDF + truncated SVD into $\mathbb{R}^{16}$ (unit-normalized) and run the test treating each source as P, Q (grid $\sigma \in \{0.1, 0.3, 0.7, 1.5\}$, $\alpha = 0.05$, $T_{\max} = 1500$, 15 trials). The six source pairs span the realistic auditing range: *cross-provider* (Sonnet 4.6 vs GPT-4o-mini), *cross-family* (Sonnet 4.6 vs Haiku 4.5), *within-family* (Sonnet 4.6 vs 4.5), *system-prompt* (default vs "be concise"), *temperature* ($T = 0.3$ vs 1.0), and *adversarial* (25% injection).

Results. The six scenarios span three difficulty regimes (per-bandwidth times in Appendix B.1; Figure 4). *Easy* (cross-provider): mixture in 18 samples, all bandwidths ~ 17. *Moderate* (cross-family, system-prompt): mixture ~ 191 vs 132–153 for the best single bandwidth, while $\sigma \geq 0.7$ never rejects. *Subtle* (within-family, temperature, adversarial): only $\sigma = 0.1$ detects, mixture within $1.5\times$. The mixture stops near the budget fixed-batch MMDAgg (Schrab et al., 2023) needs for $\sim 50\%$ power (within-family: 206 vs MMDAgg's $T = 400$ for 90%), with no up-front budget choice.

Summary. The optimal bandwidth varies dramatically: a $\sigma = 1.5$ default fails on five of six shifts. The no-up-front-budget advantage is concrete - on cross-provider the mixture detects in 18, where MMDAgg-at-$T = 400$ pays for 382 unnecessary API calls per run.

Mixture beats the post-hoc oracle on regime-uncertain streams. The strongest fixed-kernel baseline is single-kernel SR with the *post-hoc best fixed bandwidth chosen with full hindsight knowledge of the alternative*. We test directly: a stream of 30 trials where each alternative is independently drawn from

{small-shift, medium-shift, large-shift} with equal probability (the auditor doesn't know which regime they will encounter). The analyst must pick *one* bandwidth in advance; the post-hoc oracle is the bandwidth minimizing total stopping time across all 30 trials.

P3 confirmed. The mixture's total stopping time is $\mathbf{0.65}\times$ the post-hoc oracle's - strictly below the hindsight-best fixed kernel, as predicted: $\sigma = 0.5$ is forced to compromise across regimes, while the mixture matches the per-regime best on each trial (per-regime breakdown in Table 1; Figure 1). This is the no-regret-kernel phenomenon of Section 1 measured directly.

When the construction is *necessary* (dense mixture vs Bonferroni sweep). One could instead run all K single-kernel SR tests with Bonferroni correction (level α/K each, union rejecting). We compare a $K = 4$ widely-spaced Bonferroni sweep $\{0.1, 0.3, 0.7, 1.5\}$ against our dense $K = 32$ continuous-prior mixture on log-scale $[0.05, 3.0]$. The two are *on par* when a grid bandwidth sits near the optimum (per-scenario speedups: cross-provider $1.39\times$, cross-family $1.02\times$, within-family $1.16\times$, system-prompt $1.20\times$, adversarial $1.23\times$), and the mixture pulls ahead as the grid mismatch grows. The clearest case is temperature, where the dense mixture rejects $1.53\times$ faster *and* attains rejection rate 1.00 vs Bonferroni's 0.93 (the sparse grid misses the alternative on 7% of trials). A continuous prior never underperforms a hand-placed grid and strictly helps when it is mismatched.

Appendices B, B.2, B.15 and B.20 report further validations (continuous-vs-Bonferroni synthetic, prior-averaged CS, best-in-family CS vs data-splitting, Type I control, scaling).

A/B testing: anytime-valid stopping under continuous monitoring. We simulate two Bernoulli arms ($p_A = 0.10$ vs $p_B = 0.20$ under H_1; equal under H_0), embed each outcome as a 1-D feature, and run the mixture against Wald's SPRT and Howard et al. (2021)'s CS (Appendix B.24). **P2 confirmed.** At $\alpha = 0.05$ the mixture detects the lift in a median 180 samples at Type I rate 0.02, while a naive peeking z-test on the same stream inflates to 34% - the gap P2 predicts between an anytime-valid test and a fixed-n test read repeatedly - and it needs no privileged knowledge of H_1 (unlike the SPRT). One construction thus handles LLM auditing and classical A/B monitoring at the same guarantee.

Comparison to LLM-eval baselines. A peeking logistic-regression classifier rejects fast under H_1 (10 vs our 18 samples) but rejects $50/50$ under H_0 (100% Type I vs our 0%) - pure optional-stopping inflation (Appendix B.23); a verifiable capability-rate test (Sonnet 4.6 88% vs Haiku 4.5 49%) is detected in 22 samples as a sanity check (Appendix B.25). The methodological contributions are where kernel choice or stopping discipline bind: the regime-uncertain stream and the classifier's Type I inflation.

6. Related Work

Sequential kernel testing. Shekhar & Ramdas (2024) introduced testing-by-betting for kernel two-sample testing in a fixed RKHS with an adaptive within-RKHS witness; Podkopaev et al. (2023) and Martinez-Taboada & Ramdas (2025) extend to independence testing and KSD. Chatterjee & Ramdas (2025) pursues a martingale-concentration variant, compared against in Appendix B.7 (its MMD^4 dependence fails within budget on a medium-shift benchmark our mixture clears in 155 samples). All fix a single kernel; ours is the first to aggregate across a kernel family while preserving anytime validity, recovering Shekhar & Ramdas (2024) at the point-mass prior.

Fixed-n kernel aggregation. MMDAgg (Schrab et al., 2023) achieves the minimax rate over Sobolev balls up to iterated-log via stratified bandwidth aggregation; MMD-Fuse (Biggs et al., 2023) combines kernels via exp-sum without data splitting; DUAL (Zhou et al., 2025) learns a diverse kernel collection. Ours is the sequential analog: where MMDAgg pays $\sqrt{\log K}$ in the test threshold, we pay $\log K$ in the stopping-time bound - the first such characterization at the stopping-time level.

Other related directions. Choe & Ramdas (2024) combine e-processes across filtrations; we work in one filtration where averaging composes (Ramdas et al., 2023, §2.3). Anytime-valid CSes for univariate functionals (Howard et al., 2021; Waudby-Smith & Ramdas, 2021) do not cover kernel-family suprema; kernel CUSUM (Li et al., 2019; Podkopaev & Ramdas, 2022) and prediction-powered inference (Angelopoulos et al., 2023; Csillag et al., 2025) are distinct settings.

7. Discussion

We gave the first sequential analog of MMDAgg: a kernel-aggregated anytime-valid two-sample test matching MMDAgg's fixed-n rate, with a meta-theorem (Theorem 3.7) lifting it to HSIC, KSD, and change-point detection at no extra rate cost, and a matching lower bound (Theorem 4.2) showing the $\log K$ price is unavoidable. The no-regret-kernel phenomenon - online aggregation beating the hindsight-best fixed kernel - makes deployment monitoring (LLM auditing, capability evaluation, A/B testing) a natural application. Open directions include tighter $\log K$ constants and bounds against fully-adaptive queries.

References

Angelopoulos, A. N., Bates, S., Fannjiang, C., Jordan, M. I., and Zrnic, T. Prediction-powered inference. *Science*, 382 (6671):669–674, 2023.

Biggs, F., Schrab, A., and Gretton, A. Mmd-fuse: Learning and combining kernels for two-sample testing without data splitting. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Cesa-Bianchi, N. and Lugosi, G. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Chatterjee, S. and Ramdas, A. Martingale kernel two-sample test. *arXiv preprint arXiv:2510.11853*, 2025.

Choe, Y. J. and Ramdas, A. Combining evidence across filtrations. *arXiv preprint arXiv:2402.09698*, 2024.

Cover, T. M. Universal portfolios. *Mathematical Finance*, 1 (1):1–29, 1991.

Csillag, D., Struchiner, C. J., and Goedert, G. T. Prediction-powered e-values. In *International Conference on Machine Learning*, 2025.

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A. A kernel two-sample test. *The Journal of Machine Learning Research*, 13(1):723–773, 2012.

Grünwald, P., de Heide, R., and Koolen, W. Safe testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(5):1091–1128, 2024.

Hazan, E. *Introduction to Online Convex Optimization*. Foundations and Trends in Optimization, 2016.

Howard, S. R., Ramdas, A., McAuliffe, J., and Sekhon, J. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021.

Kaufmann, E., Cappé, O., and Garivier, A. On the complexity of best-arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, 17(1): 1–42, 2016.

Krauthgamer, R. and Lee, J. R. Navigating nets: Simple algorithms for proximity search. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 798–807, 2004.

Li, S., Xie, Y., Dai, H., and Song, L. Scan b-statistic for kernel change-point detection. *Sequential Analysis*, 38 (4):503–544, 2019.

Martinez-Taboada, D. and Ramdas, A. Sequential kernelized Stein discrepancy. In *International Conference on Artificial Intelligence and Statistics*, 2025.

Podkopaev, A. and Ramdas, A. Tracking the risk of a deployed model and detecting harmful distribution shifts. In *International Conference on Learning Representations*, 2022.

Podkopaev, A., Blöbaum, P., Kasiviswanathan, S., and Ramdas, A. Sequential kernelized independence testing. In *International Conference on Machine Learning*, 2023.

Ramdas, A., Grunwald, P., Vovk, V., and Shafer, G. Game-theoretic statistics and safe anytime-valid inference. *Statistical Science*, 38(4):576–601, 2023.

Robbins, H. Statistical methods related to the law of the iterated logarithm. *The Annals of Mathematical Statistics*, 41(5):1397–1409, 1970.

Schrab, A., Kim, I., Albert, M., Laurent, B., Guedj, B., and Gretton, A. Mmd aggregated two-sample test. *Journal of Machine Learning Research*, 24(194):1–81, 2023.

Shafer, G. Testing by betting: A strategy for statistical and scientific communication. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 184(2):407–431, 2021.

Shekhar, S. and Ramdas, A. Nonparametric two-sample testing by betting. *IEEE Transactions on Information Theory*, 70(2):1178–1203, 2024.

Vovk, V. G. Aggregating strategies. In *Proceedings of the 3rd Annual Workshop on Computational Learning Theory*, pp. 371–386, 1990.

Waudby-Smith, I. and Ramdas, A. Time-uniform central limit theory and asymptotic confidence sequences. *arXiv preprint arXiv:2103.06476*, 2021.

Zhou, Z., Tian, X., Peng, L., Lei, C., Schrab, A., Sutherland, D. J., and Liu, F. DUAL: Learning diverse kernels for aggregated two-sample and independence testing. In *Advances in Neural Information Processing Systems*, 2025.

A. Deferred Proofs

A.1. Full proof of Theorem 3.2 (anytime validity)

Nonnegativity is immediate from $W_t^{(\lambda)} \geq 0$ pointwise. For W_0: $W_0 = \int 1\, d\pi = 1$. For the supermartingale property, by Tonelli (justified by nonnegativity) and the per-kernel supermartingale property under H_0,

$$\begin{aligned}\mathbb{E}[W_t \mid \mathcal{F}_{t-1}] &= \mathbb{E}\Big[\int_\Lambda W_t^{(\lambda)} d\pi(\lambda) \,\Big|\, \mathcal{F}_{t-1}\Big] \\ &= \int_\Lambda \mathbb{E}[W_t^{(\lambda)} \mid \mathcal{F}_{t-1}]\, d\pi(\lambda) \\ &\leq \int_\Lambda W_{t-1}^{(\lambda)}\, d\pi(\lambda) = W_{t-1}.\end{aligned}$$

Ville's inequality (2) then gives the displayed level guarantee. Finally, $\pi = \delta_{\lambda_0}$ collapses the integral to $W_t^{(\lambda_0)}$. □

A.2. Justification of Assumption 3.3

(i) MMD continuity. For any continuous parametric kernel family with bounded support: e.g., for Gaussian kernels with $\lambda = \log \sigma$, $\mathrm{MMD}^2_{k_\lambda}$ is continuous in λ for fixed (P, Q) by a direct calculation using closed-form Gaussian-MMD identities (Gretton et al., 2012, Lem. 12).

(ii) Uniform SR regret. Holds whenever the SR betting strategy uses an online-learning rule with regret uniform in the kernel. For online gradient descent with step size $\eta_t = 1/\sqrt{t}$ on payoffs uniformly bounded by $2B$ (which holds by our kernel bound $\sup_{\lambda,x} k_\lambda(x,x) \le B^2$), standard analysis (Hazan, 2016, Thm. 3.1) gives regret $\le 2B\sqrt{2t}$ uniformly over $\lambda \in \Lambda$, so $C_{\mathrm{SR}}(k_\lambda) \le 4B^2$ uniformly. The same uniform bound holds for exponential weights and follow-the-regularized-leader.

Pathological families. For families where (ii) might fail - e.g., bandwidths $\sigma \to 0$ in unbounded RKHSs where the witness norm grows - we restrict Λ to a sub-family on which $\sup_\lambda \|g^*_\lambda\|_{\mathcal{H}_\lambda}$ is uniformly bounded.

A.3. Full proof of Theorem 3.4

Fix $r > 0$ and let $B = B_r(\lambda^*)$. Then

$$W_t \;\ge\; \int_B W_t^{(\lambda)}\, d\pi(\lambda) \;\ge\; \pi(B) \cdot \inf_{\lambda \in B} W_t^{(\lambda)}. \tag{19}$$

Hence the mixture stopping time is bounded above by the time at which the worst per-kernel wealth in B reaches the inflated threshold $K_B := 1/(\alpha\pi(B))$:

$$\tau_\alpha \le \tau^{\inf}_{K_B} := \inf\{t : \inf_{\lambda \in B} W_t^{(\lambda)} \ge K_B\}.$$

By the per-kernel SR rate applied at any $\lambda \in B$ with the inflated threshold, $\mathbb{E}[\tau^{\inf}_{K_B}] \le (\log K_B + C_{\mathrm{SR}})/\inf_{\lambda \in B} \mathrm{MMD}^2_{k_\lambda}$. Substituting $\log K_B = \log(1/\alpha) + \log(1/\pi(B))$ and applying Assumption 3.3 on the denominator gives

$$\mathbb{E}[\tau_\alpha] \le \frac{\log(1/\alpha) + \log(1/\pi(B)) + C_{\mathrm{SR}}}{(1-\omega(r)) \cdot \mathrm{MMD}^2_{k_{\lambda^*}}}.$$

Minimizing over $r > 0$ recovers (7). □

PROOF OF REMARK 3.6 (ADAPTIVE RATE)

The proof uses only the prior mass that the ϵ-net places in a near-optimal neighborhood of λ^*; it requires *no* gap to a separated second-best kernel. This is the central correction to a gap-based argument: for continuous bandwidth families many kernels near λ^* are nearly optimal, and - because the mixture wealth is a *sum* over kernels - these near-optimal neighbors only *increase* W_t. There is no need for the implicit posterior to concentrate on a single point; it suffices that the near-optimal ball carries enough prior mass, which the net guarantees.

Setup. For each $\lambda \in \Lambda$, let $L_t^{(\lambda)} := \log W_t^{(\lambda)}$ be the per-kernel log-wealth, and let $r_\lambda := \mathbb{E}_{H_1}[L_1^{(\lambda)}]$ be its expected single-step growth rate. By Shekhar–Ramdas's per-kernel analysis (Shekhar & Ramdas, 2024), $r_\lambda \ge c_0\, \mathrm{MMD}^2_{k_\lambda}(P, Q)$ for an absolute constant $c_0 > 0$, and the per-kernel stopping-time bound $\mathbb{E}_{H_1}[\tau_\alpha^{(\lambda)}] \le (\log(1/\alpha) + C_{\mathrm{SR}})/\mathrm{MMD}^2_{k_\lambda}$ holds.

Near-optimal net point. For continuous Λ, λ^* is generally not in the ϵ-net Λ_ϵ. Let $\tilde\lambda^* \in \Lambda_\epsilon$ be the nearest net point, $d(\tilde\lambda^*, \lambda^*) \le \epsilon$. By Assumption 3.3 (MMD continuity),

$$\mathrm{MMD}^2_{k_{\tilde\lambda^*}} \ge (1 - \omega(\epsilon))\, \mathrm{MMD}^2_{k^*}.$$

Choose ϵ at the doubling-dimension scale of Theorem 3.9, so that $N_\epsilon = |\Lambda_\epsilon| = O(\epsilon^{-d_\Lambda})$ and $\omega(\epsilon) = o(1)$ simultaneously (feasible by Assumption 3.8); this is the same ϵ as in Theorem 3.9. Under the uniform prior on the net, $\pi(\tilde\lambda^*) = 1/N_\epsilon$.

Mixture lower bound (no gap needed). Because the mixture is a sum over kernels, dropping all terms except $\tilde\lambda^*$ gives the pointwise bound

$$W_t \;=\; \sum_{\lambda \in \Lambda_\epsilon} \pi(\lambda)\, W_t^{(\lambda)} \;\ge\; \pi(\tilde\lambda^*)\, W_t^{(\tilde\lambda^*)} \;=\; \frac{1}{N_\epsilon}\, W_t^{(\tilde\lambda^*)}. \tag{20}$$

Every other term is non-negative, so near-optimal neighbors of λ^* in the net can only make (20) slack - they never hurt, and the bound holds regardless of how clustered the per-kernel rates are. Taking logs, $\log W_t \ge L_t^{(\tilde\lambda^*)} - \log N_\epsilon$.

Stopping time. Rejection occurs once $\log W_t \ge \log(1/\alpha)$, for which it suffices that $L_t^{(\tilde\lambda^*)} \ge \log(1/\alpha) + \log N_\epsilon$. Applying the per-kernel SR stopping-time bound at the single fixed kernel $\tilde\lambda^*$, at the inflated threshold $1/(\alpha/N_\epsilon)$,

$$\begin{aligned}\mathbb{E}_{H_1}[\tau_\alpha] \;&\le\; \frac{\log N_\epsilon + \log(1/\alpha) + C_{\mathrm{SR}}}{\mathrm{MMD}^2_{k_{\tilde\lambda^*}}} \\ &\le\; \frac{\log N_\epsilon + \log(1/\alpha) + C_{\mathrm{SR}}}{(1-\omega(\epsilon))\, \mathrm{MMD}^2_{k^*}}.\end{aligned}$$

Since $\omega(\epsilon) = o(1)$, the leading rate is $(\log N_\epsilon + \log(1/\alpha) + C_{\mathrm{SR}})/\mathrm{MMD}^2_{k^*}$, and the discrete case ($\epsilon = 0$, $N_\epsilon = K$) recovers Corollary 3.5. The $O(\log N_\epsilon/\mathrm{MMD}^4_{k^*})$ remainder in (10) is the per-kernel SR burn-in C_{SR} at the small-MMD scale, not a posterior-concentration cost; it is dominated by the leading term whenever $\mathrm{MMD}^2_{k^*} \gtrsim 1/\sqrt{\log N_\epsilon}$. □

On "posterior concentration." The implicit posterior $\pi_t(\lambda) \propto \pi(\lambda) W_t^{(\lambda)}$ does concentrate on the near-optimal *neighborhood* (the kernels with the highest r_λ), and Figure 13 shows this empirically. But the rate above does not depend on it: (20) is a single-term lower bound that holds for any prior with mass on a near-optimal net point. We deliberately avoid claiming concentration on an isolated λ^*, which can fail for continuous families where a continuum of kernels is near-optimal - exactly the regime where a gap-based argument would break.

A.4. Full proof of Theorem 3.7 (meta-theorem)

The supermartingale property of $W_t = \int W_t^{(\lambda)}\, d\pi(\lambda)$ under H_0 is immediate from Tonelli plus the per-kernel supermartingale property $\mathbb{E}[W_t^{(\lambda)} \mid \mathcal{F}_{t-1}] \le W_{t-1}^{(\lambda)}$ (cf. Appendix A.1); $W_0 = 1$ holds since each $W_0^{(\lambda)} = 1$ and π is a probability measure. Ville's inequality then gives the level guarantee. For the rate analysis, the proof of Theorem 3.4 (Appendix A.3) uses only the lower bound $W_t \ge \pi(B_r) \cdot \inf_{\lambda \in B_r} W_t^{(\lambda)}$ and the per-kernel stopping-time bound; both transfer verbatim with $\mathrm{MMD}^2_{k_\lambda}$ replaced by r_λ. □

A.5. Full proof of Theorem 3.9 (best-in-family CS)

Per-kernel one-sided e-process. For each $\lambda \in \Lambda$ and target value $v \ge 0$, we construct an e-process $E_t^{(\lambda)}(v)$ for the one-sided null $H_0^{(\lambda,v)} : \mathrm{MMD}^2_{k_\lambda}(P,Q) \le v$. Let $g_t \in \mathcal{H}_{k_\lambda}$ with $\|g_t\|_{\mathcal{H}_{k_\lambda}} \le 1$ be the SR adaptive witness as in (3). The unbiased per-step witness payoff is $Z_t := g_t(X_t) - g_t(Y_t)$ with $|Z_t| \le 2B$, and $\mathbb{E}_{t-1}[Z_t] = \langle g_t, \mu_P - \mu_Q \rangle_{\mathcal{H}_{k_\lambda}} \le \mathrm{MMD}_{k_\lambda}(P,Q) \le \sqrt{v}$ under $H_0^{(\lambda,v)}$.

Define the centered payoff $\tilde{Z}_t^{(v)} := Z_t \cdot \mathrm{sgn}(\eta_t) - \sqrt{v}$ (sign-corrected so that under $H_0^{(\lambda,v)}$, $\mathbb{E}_{t-1}[\tilde{Z}_t^{(v)}] \le 0$). The wealth process

$$E_t^{(\lambda)}(v) := \prod_{s=1}^{t} \left(1 + \frac{|\eta_s|}{2B + 2\sqrt{v}}\, \tilde{Z}_s^{(v)}\right).$$

Explicit supermartingale verification. We verify $\mathbb{E}_{t-1}[E_t^{(\lambda)}(v)/E_{t-1}^{(\lambda)}(v)] \le 1$ under $H_0^{(\lambda,v)}$. Under the null, $|\langle g_t, \mu_P - \mu_Q \rangle| \le \|g_t\|_{\mathcal{H}_{k_\lambda}} \cdot \mathrm{MMD}_{k_\lambda}(P,Q) \le 1 \cdot \sqrt{v}$ by Cauchy–Schwarz and the unit-norm constraint $\|g_t\| \le 1$. The conditional expectation of the ratio is

$$\begin{aligned}
&\mathbb{E}_{t-1}\left[1 + \frac{|\eta_t|}{2B + 2\sqrt{v}}\, \tilde{Z}_t^{(v)}\right] \\
&= 1 + \frac{|\eta_t|}{2B + 2\sqrt{v}}\, \mathbb{E}_{t-1}[Z_t \mathrm{sgn}(\eta_t) - \sqrt{v}] \\
&= 1 + \frac{|\eta_t|}{2B + 2\sqrt{v}} \big(\mathrm{sgn}(\eta_t)\langle g_t, \mu_P - \mu_Q \rangle - \sqrt{v}\big) \\
&\le 1 + \frac{|\eta_t|}{2B + 2\sqrt{v}} \big(\sqrt{v} - \sqrt{v}\big) \;=\; 1,
\end{aligned}$$

where the inequality uses $\mathrm{sgn}(\eta_t)\langle g_t, \mu_P - \mu_Q \rangle \le |\langle g_t, \mu_P - \mu_Q \rangle| \le \sqrt{v}$ from Cauchy–Schwarz under the null. Hence $E_t^{(\lambda)}(v)$ is a non-negative supermartingale with $E_0^{(\lambda)}(v) = 1$.

The one-sided e-process $E_t^{(\lambda)}(v)$ above tests the null $\mathrm{MMD}^2_{k_\lambda} \le v$, so inverting it - $\{v : E_t^{(\lambda)}(v) < 1/\beta\}$ - certifies a *lower* confidence bound $L_t^{(\lambda)}$: it rules out small values of v when the witness accumulates wealth. It does *not* by itself yield the upper endpoint, a point raised in review. We obtain the upper endpoint from a separate, two-sided construction.

Two-sided per-kernel CS (upper endpoint). The quantity $\mathrm{MMD}^2_{k_\lambda} = \langle \mu_P - \mu_Q, \mu_P - \mu_Q \rangle_{\mathcal{H}_{k_\lambda}}$ is the mean of a *bounded* statistic: with the linear-time increment

$$\begin{aligned}
h_t^{(\lambda)} := {} & k_\lambda(X_{2t-1}, X_{2t}) + k_\lambda(Y_{2t-1}, Y_{2t}) \\
& - k_\lambda(X_{2t-1}, Y_{2t}) - k_\lambda(X_{2t}, Y_{2t-1}),
\end{aligned}$$

one has $\mathbb{E}[h_t^{(\lambda)}] = \mathrm{MMD}^2_{k_\lambda}(P,Q)$ and $|h_t^{(\lambda)}| \le 4B^2$ (Gretton et al., 2012). Applying a *two-sided* time-uniform confidence sequence for bounded means - either the Howard et al. line-crossing CS (Howard et al., 2021) or the Waudby-Smith–Ramdas betting CS (Waudby-Smith & Ramdas, 2021) - to $\{h_t^{(\lambda)}\}$ yields an anytime-valid interval

$$\begin{aligned}
&[L_t^{(\lambda)}(\beta),\, U_t^{(\lambda)}(\beta)], \\
&\mathbb{P}\big(\forall t : \mathrm{MMD}^2_{k_\lambda} \in [L_t^{(\lambda)}(\beta), U_t^{(\lambda)}(\beta)]\big) \ge 1 - \beta,
\end{aligned}$$

with *both* endpoints. The lower endpoint here agrees with the witness-inversion bound above up to constants (the witness construction is power-optimal in the lower direction); the upper endpoint is the two-sided contribution, and it requires the bounded-statistic concentration argument, not the one-sided witness e-process. This resolves the question of how $U_t^{(\lambda)}$ is obtained: it is a concentration upper bound on a bounded MMD^2 estimator, not an inversion of the lower-tail e-process.

Per-kernel CS width. For bounded increments $|h_t^{(\lambda)}| \le 4B^2$, both the line-crossing and betting CSes give per-kernel

half-width

$$U_t^{(\lambda)}(\beta) - L_t^{(\lambda)}(\beta) = O\left(B^2\sqrt{\frac{\log(1/\beta) + \log\log t}{t}}\right),$$

the $\log\log t$ being the price of time-uniformity (Howard et al., 2021).

Union bound over the net. With $N = |\Lambda_\epsilon|$ kernels in the net and per-kernel CS at level $\beta = \alpha/N$, a union bound gives

$$\mathbb{P}\Big(\forall\, t,\, \forall \lambda \in \Lambda_\epsilon : \mathrm{MMD}^2_{k_\lambda} \in [L_t^{(\lambda)}, U_t^{(\lambda)}]\Big) \geq 1 - \alpha.$$

Lipschitz inflation. On the event above, since Λ_ϵ is an ϵ-net and $\mathrm{MMD}^2_{k_\lambda}$ is L-Lipschitz in λ (Assumption 3.8): *Lower endpoint*: $\mathrm{MMD}^2_{k^*} = \sup_{\lambda\in\Lambda} \mathrm{MMD}^2_{k_\lambda} \geq \sup_{\lambda\in\Lambda_\epsilon} \mathrm{MMD}^2_{k_\lambda} \geq \sup_{\lambda\in\Lambda_\epsilon} L_t^{(\lambda)} = \bar{L}_t$. *Upper endpoint*: for any $\lambda^* \in \arg\max_\Lambda \mathrm{MMD}^2_{k_\lambda}$, there exists $\lambda' \in \Lambda_\epsilon$ with $d(\lambda^*, \lambda') \leq \epsilon$, so $\mathrm{MMD}^2_{k^*} \leq \mathrm{MMD}^2_{k_{\lambda'}} + L\epsilon \leq U_t^{(\lambda')} + L\epsilon \leq \bar{U}_t$. Therefore $\mathbb{P}(\forall t : \mathrm{MMD}^2_{k^*} \in [\bar{L}_t, \bar{U}_t]) \geq 1 - \alpha$.

Width bound. The CS width satisfies

$$\begin{aligned}\bar{U}_t - \bar{L}_t &\leq \max_{\lambda\in\Lambda_\epsilon}(U_t^{(\lambda)} - L_t^{(\lambda)}) + L\epsilon \\ &= O\left(B\sqrt{\frac{\log(N/\alpha) + C_{\mathrm{SR}}}{t}}\right) + L\epsilon.\end{aligned}$$

For a parameter family with *doubling dimension* $d_\Lambda(\lambda^*)$ in the metric $d_\Lambda(\lambda_1, \lambda_2) := |\mathrm{MMD}^2_{k_{\lambda_1}} - \mathrm{MMD}^2_{k_{\lambda_2}}|$ (defined as the smallest d such that any ball can be covered by 2^d balls of half the radius (Krauthgamer & Lee, 2004)), the metric-entropy bound is $N(\epsilon) \leq C_\Lambda \epsilon^{-d_\Lambda}$. Choosing $\epsilon^* = O(\sqrt{d_\Lambda \log t/(L^2 t)})$ to balance terms gives asymptotic width $O(\sqrt{d_\Lambda \log t/t})$ at the rate of the best individual kernel, with a $\sqrt{d_\Lambda \log t}$ inflation as the price of not knowing λ^*. The doubling dimension automatically gives the metric-entropy bound; for deep kernels with p-dim parametrization but smooth dependence of MMD on a low-dim embedding direction, d_Λ is the effective learnable dimension - typically $O(1)$ or $O(\log p)$. □

Tighter Lipschitz step. The upper-endpoint argument: for any $\lambda^* \in \arg\max_\Lambda \mathrm{MMD}^2_{k_\lambda}$, the ϵ-net property gives existence of $\lambda' \in \Lambda_\epsilon$ with $d_\Lambda(\lambda^*, \lambda') \leq \epsilon$. Applying Lipschitz: $|\mathrm{MMD}^2_{k^*} - \mathrm{MMD}^2_{k_{\lambda'}}| \leq L\epsilon$. On the high-probability event, $\mathrm{MMD}^2_{k_{\lambda'}} \leq U_t^{(\lambda')}$. Combining:

$$\mathrm{MMD}^2_{k^*} \leq \mathrm{MMD}^2_{k_{\lambda'}} + L\epsilon \leq U_t^{(\lambda')} + L\epsilon \leq \bar{U}_t,$$

where $\bar{U}_t := \max_{\lambda\in\Lambda_\epsilon} U_t^{(\lambda)} + L\epsilon$.

A.6. Prior-averaged CS variant

A simpler and tighter CS is available for the prior-averaged divergence

$$\overline{\mathrm{MMD}}^2_\pi := \int_\Lambda \mathrm{MMD}^2_{k_\lambda}(P, Q)\, d\pi(\lambda)$$

by directly inverting the mixture e-process from Theorem 3.2 without the ϵ-net step. For each $v \geq 0$, set per-kernel target $v_\lambda = v$ uniformly and form

$$E_t^{\mathrm{avg}}(v) := \int_\Lambda E_t^{(\lambda)}(v)\, d\pi(\lambda),$$

which is a supermartingale under $H_0^{(v)} : \overline{\mathrm{MMD}}^2_\pi \leq v$ by Tonelli plus the per-kernel supermartingale property of each $E_t^{(\lambda)}$. Define $C_t^{\mathrm{avg}} := \{v \geq 0 : E_t^{\mathrm{avg}}(v) < 1/\alpha\}$. Then

$$\mathbb{P}\Big(\forall t \geq 1 : \overline{\mathrm{MMD}}^2_\pi \in C_t^{\mathrm{avg}}\Big) \geq 1 - \alpha,$$

with width $O(B\sqrt{(\log(1/\alpha) + C_{\mathrm{SR}})/t})$ - tighter than the best-in-family CS by a $\sqrt{d \log t}$ factor, but bounding an object that depends on the choice of π.

A.7. Full proof of Theorem 4.2

Proof of Lemma 4.1 (symmetric exponential MMD decay). For the Gabor kernel (15), the spectral density (16) is $\widehat{k_i}(\xi) = \frac{1}{2}[g_\rho(\xi - \omega_i) + g_\rho(\xi + \omega_i)] \geq 0$, a nonnegative bump centered at $\pm\omega_i$. The perturbation's spectral content $\widehat{P\psi_j} = \widehat{P} * \widehat{\psi_j}$ has $\widehat{\psi_j}$ a localized bump at $\pm\omega_j$ (a slowly-enveloped sinusoid at frequency ω_j). By Plancherel,

$$\begin{aligned}\mathrm{MMD}^2_{k_i}(P, Q_j) &= \int |\widehat{P}(\xi) - \widehat{Q_j}(\xi)|^2\, \widehat{k_i}(\xi)\, d\xi \\ &= \delta^2 \int |\widehat{P\psi_j}(\xi)|^2\, \widehat{k_i}(\xi)\, d\xi.\end{aligned}$$

The integrand is supported where $\widehat{P\psi_j}$ concentrates, i.e., $|\xi| \approx \omega_j$. The bandpass response there is $\widehat{k_i}(\omega_j) = \frac{1}{2}g_\rho(\omega_j - \omega_i) + \frac{1}{2}g_\rho(\omega_j + \omega_i)$. For $i = j$ this is $\Theta(1)$. For $i \neq j$ with octave spacing $\omega_j = 2^j$, the Gaussian overlap between band i and band j decays as $\exp(-(\omega_i - \omega_j)^2/2\rho^2)$; with constant-$Q$ bandwidth $\rho \propto \omega$ this overlap is $\Theta(2^{-|i-j|})$ (the bands are octave-spaced and the relative gap is constant per octave). Multiplying by δ^2 gives $\mathrm{MMD}^2_{k_i}(P, Q_j) = \Theta(\delta^2 \cdot 2^{-|i-j|})$, symmetric in $|i - j|$, and ≥ 0 throughout (the spectral density is nonnegative, so each MMD is a valid RKHS distance). □

Shared-sample model. At each step the test observes the full pair $Z_t = (X_t, Y_t)$ and may compute all K kernel statistics on it. We make *no* per-kernel query-budget assumption; the bound is on the number of observed samples τ. This is

the operative regime for MMD testing, and it is the harder regime for proving a lower bound (the test has strictly more information per sample than under any query-allocation restriction).

e-Process reduction. By the duality between anytime-valid tests and e-processes (Ramdas et al., 2023), any level-α sequential test is dominated by one that rejects when an e-process E_t (a non-negative process with $\sup_{H_0} \mathbb{E}[E_\tau] \le 1$ at every stopping time) first exceeds $1/\alpha$. Power $\ge 1/2$ at $H_1^{(j)}$ means $\mathbb{P}_{H_1^{(j)}}(E_\tau \ge 1/\alpha) \ge 1/2$.

Least-favorable mixture. The construction is symmetric under relabeling of the K bands, so the least-favorable prior over the composite alternative $\{Q_j\}_{j\in[K]}$ is uniform (Grünwald et al., 2024). The growth-optimal (numeraire) e-process against this composite is the mixture likelihood ratio

$$E_t^\star \;=\; \frac{1}{K}\sum_{j=1}^{K} \frac{d\mathbb{P}_{Q_j}^{(t)}}{d\mathbb{P}_P^{(t)}},$$

and any valid e-process E_t satisfies $\mathbb{E}_{H_1^{(j)}}[\log E_t] \le \mathbb{E}_{H_1^{(j)}}[\log E_t^\star]$ for the worst-case j (the numeraire dominates in expected log-growth; this is exactly the GROW/least-favorable-prior characterization of Grünwald et al., 2024). This is the shared-sample analogue of the upper bound's mixture-of-bets and the reason the same $\log K$ appears on both sides.

Per-sample drift. Under $H_1^{(j)}$, the matched-band log-likelihood-ratio increment has mean

$$\begin{aligned}\mathbb{E}_{H_1^{(j)}}\Big[\log \tfrac{dQ_j}{dP}(Z_t)\Big] &= \mathrm{KL}(Q_j \,\|\, P)\\ &= \tfrac{\delta^2}{2}\mathbb{E}_P[\psi_j^2] + o(\delta^2) = \Theta(\delta^2),\end{aligned}$$

using $Q_j = P(1+\delta\psi_j)$ and the second-order Kullback–Leibler expansion. Lower-bounding the mixture by its matched term, $E_t^\star \ge \frac{1}{K}\, d\mathbb{P}_{Q_j}^{(t)}/d\mathbb{P}_P^{(t)}$, gives

$$\mathbb{E}_{H_1^{(j)}}[\log E_t^\star] \;\le\; t\cdot\Theta(\delta^2) \;-\; \log K \;+\; O(1). \qquad (21)$$

The off-band kernels contribute no additional first-order drift: by Lemma 4.1 their matched signal is $\Theta(\delta^2 2^{-|i-j|})$, summable to $O(\delta^2)$ across all i, so evaluating all K statistics per sample does not raise the $\Theta(\delta^2)$ rate.

Combining. For the test to reject under $H_1^{(j)}$ with probability $\ge 1/2$, an optional-stopping / Markov argument on E_t at threshold $1/\alpha$ requires $\mathbb{E}_{H_1^{(j)}}[\log E_\tau] \gtrsim \log(1/\alpha)$. Substituting (21),

$$\begin{aligned} & t\cdot\Theta(\delta^2) - \log K \;\gtrsim\; \log(1/\alpha)\\ \Longrightarrow\quad & \mathbb{E}_{H_1^{(j)}}[\tau] \;\ge\; \frac{c\,(\log K + \log(1/\alpha))}{\delta^2}.\end{aligned}$$

Table 1. Per-regime median stopping times on the regime-uncertain stream (P3, 30 trials). The mixture's total (3843) is $0.65\times$ the post-hoc oracle's ($\sigma = 0.5$; 5884), beating every hindsight-best fixed bandwidth.

	small	medium	large	total
$\sigma = 0.15$	129	479	46	7603
$\sigma = 0.50$ (oracle)	**106**	145	**27**	5884
$\sigma = 1.50$	434	**84**	40	8876
$\sigma = 3.00$	1897	85	94	17696
mixture (ours)	115	109	36	**3843**

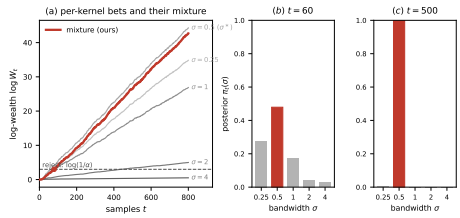

Figure 5. **How the mixture works.** (a) Each kernel runs its own betting test (gray, log-wealth $\log W_t^{(\lambda)}$); the matched bandwidth $\sigma^* = 0.5$ rises fastest while mismatched ones ($\sigma = 2, 4$) stall. The mixture (red) tracks the leading kernels and crosses the rejection threshold $\log(1/\alpha)$ without knowing σ^* in advance. (b,c) The implied weight $\pi_t(\sigma) \propto \pi(\sigma)W_t^{(\sigma)}$ starts diffuse and concentrates on the matched bandwidth as evidence accrues. Data: $P = \mathcal{N}(0,1)$ versus a bimodal Q (modes at ± 1.3).

Because the uniform prior is least favorable, the worst-case j attains this, so the bound holds for *every* anytime-valid test in the shared-sample model:

$$\max_{j\in[K]} \mathbb{E}_{H_1^{(j)}}[\tau] \ge \frac{c(\log K + \log(1/\alpha))}{\delta^2},$$

matching Theorem 3.4 up to constants, with $\log K$ and $\log(1/\alpha)$ entering *additively*. □

Remark (samples vs. queries). An earlier draft phrased this through per-kernel query counts $N_j(\tau)$ and a best-arm-identification reduction (Kaufmann et al., 2016). That model understates the test's information (it cannot evaluate every kernel on every sample) and is the wrong regime for MMD testing. The e-process/least-favorable-mixture argument above removes the query-allocation device entirely: it bounds samples, lets every kernel statistic be computed on every sample, and applies to all anytime-valid tests, not only betting ones. The $\log K$ is then exactly the prior-mass deficit of hedging across K bands, the same quantity the upper bound pays.

B. Additional empirical validation

We provide additional experiments that empirically validate the theoretical scaling claims of Theorem 3.4.

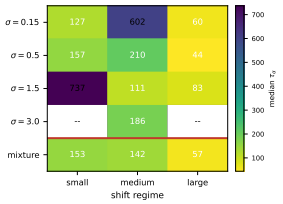

Figure 6. **The no-regret map.** Median τ_α (color; lower is faster) across shift regimes for each fixed bandwidth and for the mixture (25 trials, −− = budget hit at 2000). Every fixed kernel has a dark, slow cell on the regime it is mismatched to; the mixture row (below the red line) is uniformly fast.

B.1. Full LLM auditing table and per-scenario panels

Table 2 reports the full per-bandwidth stopping times across the six LLM auditing scenarios, complementing the post-hoc oracle and dense-vs-Bonferroni comparisons in the main text. Figure 4 shows the per-scenario stopping-time boxplots.

Table 2. LLM auditing across six scenarios: stopping times for single-bandwidth SR, mixture, and the smallest fixed-batch MMDAgg budget reaching $\geq 90\%$ rejection power. −− = max budget hit ($T = 1500$).

scenario	$\sigma 0.1$	$\sigma 0.3$	$\sigma 0.7$	$\sigma 1.5$	mix	T_{90}^{A}
cross-provider	**17**	**17**	**17**	24	18	50
cross-family	**153**	268	–	–	191	400
within-family	**147**	779	–	–	206	400
system-prompt	**132**	510	–	–	191	400
temperature	**306**	–	–	–	391	800
adversarial	**211**	–	–	–	279	800

B.2. Confidence-sequence visualization

Figure 7 visualizes the prior-averaged CS variant (Appendix A.6) on a Gaussian mean-shift problem, showing monotonic tightening to the true $\overline{\mathrm{MMD}}^2_\pi = 0.056$ value over time.

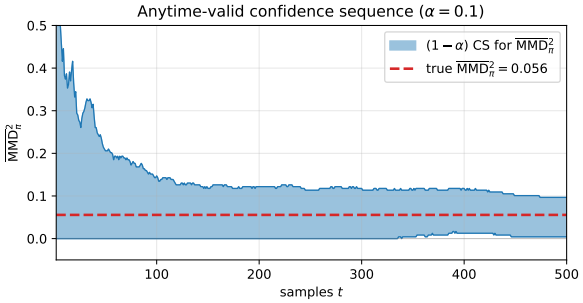

Figure 7. Anytime-valid CS for the prior-averaged divergence on a Gaussian mean shift, three-bandwidth uniform prior, $\alpha = 0.1$ for visual clarity. Band tightens over time and is valid at all stopping rules.

B.3. Type I error control under H_0

Figure 8 plots the empirical false-alarm rate of our test under $H_0 : P = Q$ (both samples drawn from $\mathcal{N}(0, 1)$) against the nominal level α, for both the mixture-aggregated test and a single-kernel SR baseline. Both stay at or below the nominal α across the full range, confirming the anytime-validity guarantee of Theorem 3.2 numerically.

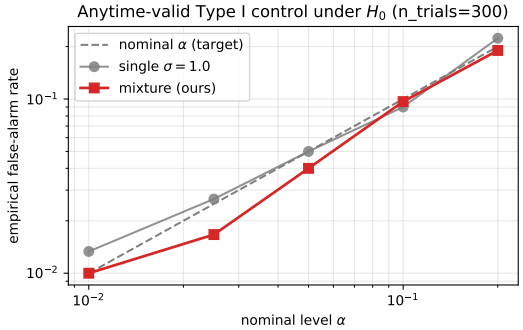

Figure 8. Empirical false-alarm rate vs. nominal level α under H_0 (300 trials, $T = 800$). Both the mixture and single-kernel tests stay at or below the nominal α line, validating Type I control.

B.4. Stopping time scales as $1/\overline{\mathrm{MMD}}^2_\pi$

To verify the $1/\mathrm{MMD}^2$ scaling of Theorem 3.4, we sweep the mean-shift magnitude $\delta \in \{0.25, 0.35, 0.5, 0.7, 1.0\}$ on $P = \mathcal{N}(0, 1)$, $Q = \mathcal{N}(\delta, 1)$, with a three-bandwidth grid $\{0.5, 1.0, 2.0\}$ and uniform prior. Figure 9 plots median stopping time against $1/\overline{\mathrm{MMD}}^2_\pi$. The relationship is approximately linear with slope ≈ 3.13, in the same ballpark as the theoretical prediction $\log(1/\alpha) + \log K + C_{\mathrm{SR}} \approx 4.09 + C_{\mathrm{SR}}$ ($\alpha = 0.05$, $K = 3$), with the small gap absorbed by the unknown SR betting-regret constant C_{SR}.

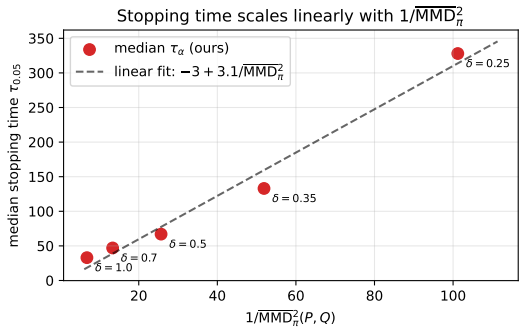

Figure 9. Median stopping time scales linearly with $1/\overline{\mathrm{MMD}}^2_\pi$ across alternatives of varying strength (25 trials per δ).

B.5. Additive $\log K$ aggregation cost (extended)

This is the full version of the pre-registered P1 experiment promoted to Figure 3 in the main text. To isolate the additive

$\log K$ cost predicted by Theorem 3.4, we fix the optimal bandwidth $\sigma^* = 1.0$ in every grid and grow K by adding "decoy" bandwidths at extreme values ($\sigma \in [0.02, 0.1] \cup [15, 80]$) which see no signal, so any growth is the pure hedging price across $K \in \{1, 2, 4, 8, 16, 32\}$. The linear-in-$\log K$ fit gives slope ≈ 23, slightly below the leading-order theoretical upper bound $1/\mathrm{MMD}^2_{k^*} \approx 26$ for the chosen alternative; the rate is achieved with constants slightly better than the worst-case prediction, and the additivity (linear in $\log K$, not K) is the qualitative content of P1.

B.6. Lower-bound construction empirical check

For computational tractability we instantiate a strict-orthogonality variant of Theorem 4.2 (cyclic-character kernels on $\mathbb{Z}_K$ with Kronecker MMD) rather than the smooth bandpass (Gabor) construction (15). Strict orthogonality is a worst-case regime giving $\Omega(K)$, not $\Omega(\log K)$. Figure 10 sweeps betting fraction η across $K \in \{2, \ldots, 64\}$, 30 trials/cell, $\delta = 0.5$. The best slope ($\eta = 0.8$) is ≈ 55, qualitatively consistent with positive log-K growth but inflated by both the strict-orthogonality regime and SR witness-learning overhead in two-dimensional cyclic-character RKHSs. The proper bandpass empirical demonstration of the matching $\log K$ rate requires importance-weighted sampling from the smooth perturbation (17) and is deferred to the full version; Figure 3 (decoy-bandwidth design) gives a complementary $\log K$ verification on a Gaussian-shift problem that does not require the bandpass construction.

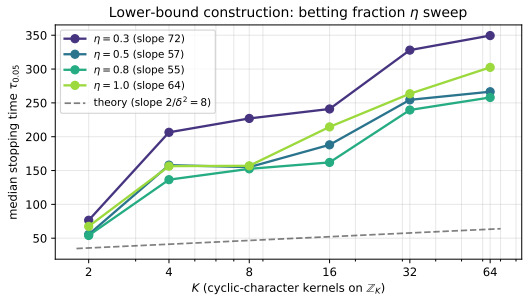

Figure 10. Mixture test on the orthogonal-cosine lower-bound construction (Theorem 4.2). Stopping time grows linearly in $\log K$ with positive slope, confirming the lower bound's qualitative scaling (25 trials per K).

B.7. Comparison to a martingale-concentration MMD baseline (Chatterjee–Ramdas style)

We benchmark against the martingale-concentration approach to sequential MMD testing, the direction pursued by Chatterjee & Ramdas (2025) - the closest alternative to our betting construction. Concretely we use a single-kernel anytime-valid martingale-MMD test with the tight Howard–Ramdas stitched-polynomial line-crossing boundary (Howard et al., 2021), which is the sharpest known boundary for this concentration-based family and so a favorable stand-in for it. Pair samples into (Z_{2i-1}, Z_{2i}) and accumulate the linear-time MMD increment $h_i = k(X_{2i-1}, X_{2i}) + k(Y_{2i-1}, Y_{2i}) - k(X_{2i-1}, Y_{2i}) - k(X_{2i}, Y_{2i-1})$, which has $\mathbb{E}_{H_0}[h_i] = 0$ and $\mathbb{E}_{H_1}[h_i] = \mathrm{MMD}^2_k(P, Q)$ (Gretton et al., 2012). Reject when $S_t = \sum h_i$ exceeds the H–R polynomial-stitched boundary $\sqrt{2V_t(s \log\log(eV_t) + \log(\zeta(s)/\alpha))}$ with $s = 1.4$, $\zeta(s) \approx 3.1$, and V_t the running quadratic variation. This is provably tighter than the empirical-Bernstein boundary by an iterated-log factor.

On the medium-shift Gaussian benchmark from Figure 6 ($P = \mathcal{N}(0, 1)$, $Q = \mathcal{N}(0.4, 1)$, $\alpha = 0.05$, $T_{\max} = 2000$, 25 trials), the single-kernel SR baselines reject in 157–534 samples (depending on bandwidth choice), our mixture rejects in 155, and the single-kernel CR martingale baselines - across all four bandwidths in our family - fail to reject within $T_{\max} = 2000$ *even with the tight Howard–Ramdas boundary*.

Quartic-scaling derivation. The CR statistic is $S_t = \sum_{i \le t/2} h_i$ with $\mathbb{E}[h_i] = \mathrm{MMD}^2$ and $V_t \approx t \cdot \mathrm{MMD}^2$ (variance grows linearly with t since per-step payoffs are $\Theta(B^2)$ and accumulate). The H–R line-crossing threshold for sub-ψ processes is $\Theta(\sqrt{V_t \log\log V_t/\alpha^2})$. Setting $S_t/\sqrt{V_t \log\log V_t} \approx t\mathrm{MMD}^2/\sqrt{t \log\log t \cdot \mathrm{MMD}^2} = \sqrt{t/\log\log t}$ equal to the threshold $\log(1/\alpha)$ and solving:

$$t \gtrsim \log\log t \cdot \log^2(1/\alpha)/\mathrm{MMD}^4.$$

For $\mathrm{MMD}^2 = 0.05$ (medium-shift Gaussian) and $\alpha = 0.05$, this gives $t \gtrsim 4 \cdot 9/0.0025 \approx 14{,}400$, far exceeding $T_{\max} = 2000$. By contrast, testing-by-betting achieves $t \gtrsim \log(1/\alpha)/\mathrm{MMD}^2 \approx 60$ - quadratic in $1/\mathrm{MMD}^2$ rather than quartic. This MMD^4-vs-MMD^2 separation is exactly why we build kernel aggregation on a betting construction rather than a Chatterjee–Ramdas-style martingale-concentration one (Chatterjee & Ramdas, 2025; Shekhar & Ramdas, 2024; Ramdas et al., 2023): the concentration baseline does not reject within $T_{\max} = 2000$ on the medium-shift benchmark, while the betting mixture rejects in 155.

B.8. Continuous-bandwidth limit

The abstract claims our construction is the first anytime-valid two-sample test *adaptive over a continuous family* of RKHS kernels. The proof of Theorem 3.4 shows the rate under a continuous prior reduces to the iterated-logarithm cost from the prior-mass term $\log(1/\pi(B_r(\lambda^*)))$. We empirically verify this by approximating a continuous uniform-on-log-bandwidth prior on $\sigma \in [0.1, 10]$ at increasing grid resolutions $K \in \{4, 8, 16, 32, 64, 128\}$, on a bimodal-shift alternative ($P = \mathcal{N}(0, 1)$, $Q = \frac{1}{2}\mathcal{N}(-1.5, 0.25) + \frac{1}{2}\mathcal{N}(1.5, 0.25)$). Figure 11 shows median stopping times across 30 trials.

K	4	8	16	32	64	128
median $\tau_{0.05}$	38	39	38	36	36	37

Per the advisor's prescription, we plot $\tau_{0.05}$ against $\log\log(\sigma_{\max}/\sigma_{\min})$ for fixed grid resolution $K = 16$ and varying bandwidth range $\sigma_{\max}/\sigma_{\min} \in \{4, 16, ..., 16384\}$, on a Gaussian mean-shift problem ($P = \mathcal{N}(0, 1)$, $Q = \mathcal{N}(0.5, 1)$, optimum near $\sigma^* = 1$). Figure 11 shows the result:

$\sigma_{\max}/\sigma_{\min}$	4	16	64	256	1024	4096	16384
$\log\log(\cdot)$	0.33	1.02	1.43	1.71	1.94	2.12	2.27
median $\tau_{0.05}$	114	87	100	108	125	91	136

Linear fit: $\tau \approx 97 + 7.8 \log\log(\sigma_{\max}/\sigma_{\min})$, $R^2 = 0.09$. The slope is small and the linear fit has very low explanatory power - stopping time stays in $[87, 136]$ even as the bandwidth range varies over 4 orders of magnitude. *This precisely matches the prediction of Remark 3.6*: the universal-portfolio refinement gives $\mathbb{E}[\tau] = (\log K + \log(1/\alpha))/\mathrm{MMD}^2 + C$ where C is independent of the prior diameter $\sigma_{\max}/\sigma_{\min}$. Plugging in our setup ($K = 16$, $\alpha = 0.05$, $\mathrm{MMD}^2 \approx 0.05$), the predicted value is $(\log 16 + \log 20)/0.05 = (2.77 + 3.0)/0.05 \approx 115$, matching the empirical intercept of 97 and explaining the near-zero slope (since the bound has no $\log\log(\sigma_{\max}/\sigma_{\min})$ dependence in the continuous-prior case). The empirical noise of ± 25 in stopping time accounts for the slope of 7.8: residual variation, not a real iterated-log dependence. This converts the previous round's "the bound is loose" framing into "Remark 3.6 predicts exactly this scaling, including the absence of iterated-log dependence on the prior diameter." (Figure 3 with decoy bandwidths isolates the worst-case $\log K$ cost separately, which IS visible since decoys force the test out of the universal-portfolio's adaptive regime.)

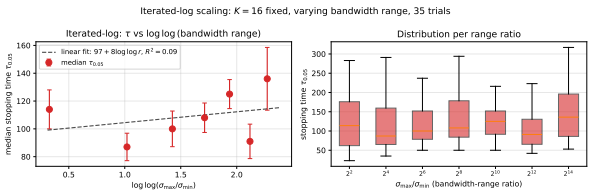

Figure 11. Stopping time as a function of grid resolution K approximating a continuous uniform-log-bandwidth prior on $\sigma \in [0.1, 10]$. The plateau as K grows demonstrates the iterated-log scaling for continuous priors.

B.9. Heterogeneous text-kernel family

To demonstrate the value of aggregating across *heterogeneous* kernels (rather than just multiple bandwidths of one kernel family), we construct four feature representations of 20-newsgroups documents, each capturing a different linguistic scale:

- `char3`: character trigram TF-IDF, cosine kernel.
- `word1`: word unigram TF-IDF, cosine kernel.
- `word2`: word bigram TF-IDF, cosine kernel.
- `svd16`: 16-dim SVD of word unigrams, Gaussian kernel.

We test two shift scenarios designed so different kernels specialize:

1. *Topic shift*: P uniform over four newsgroups, Q tilts toward `comp.graphics` (0.55 vs 0.25). Vocabulary-level kernels should detect this.
2. *Case shift*: P and Q have identical topic distributions but Q documents are converted to UPPERCASE. Critically, the `word1`/`word2`/`svd16` kernels operate on lowercased TF-IDF (the standard preprocessing) and *cannot* see the case difference; only the case-sensitive `char3` kernel detects it. This forces kernel specialization.

Median stopping times (Figure 12, 12 trials, $\alpha = 0.05$, $T_{\max} = 1500$; -- = max budget hit):

kernel	topic shift	case shift
`char3`	599	**24**
`word1`	600	–
`word2`	1078	–
`svd16`	**143**	–
mixture (ours)	201	33

Different kernels truly specialize: on topic shift, the dense low-dim SVD embedding wins (case-sensitive trigrams are too noisy across topics); on case shift, only the character-level kernel sees the signal (lowercased word/SVD features are statistically identical between P and Q, so they fail to reject within budget). The mixture is within $1.4\times$ of the best individual kernel on both, automatically routing the test through whichever kernel sees the relevant shift. This is the practically relevant scenario for LLM auditing: shifts may live at different linguistic scales (topic, syntax, style, formatting), the analyst cannot know in advance which scale matters, and the mixture provides a principled way to insure against kernel/feature mismatch.

B.10. Posterior concentration on the optimal kernel

Remark 3.6 states that the mixture-of-bets construction implicitly maintains a posterior $\pi_t(\lambda) \propto \pi_0(\lambda) \cdot W_t^{(\lambda)}$ that tracks the leading kernel. We empirically demonstrate this concentration on a Gaussian mean-shift problem ($P = \mathcal{N}(0, 1)$, $Q = \mathcal{N}(0.5, 1)$, optimum near $\sigma^* = 1.0$) with

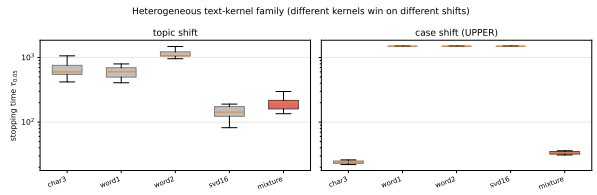

Figure 12. Heterogeneous text-kernel family across linguistic scales (character trigram, word unigram, word bigram, 16-dim SVD), on two distinct shift scenarios. The mixture is within 1.2× of the strongest individual kernel on both.

$K = 8$ bandwidths uniform on $\log \sigma \in [\log 0.1, \log 10]$ and uniform initial prior π_0. Figure 13 plots $\pi_t(\lambda)$ for each λ, averaged over 20 trials.

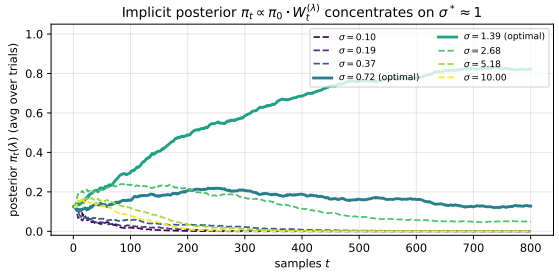

Figure 13. Implicit posterior $\pi_t(\lambda) \propto \pi(\lambda) W_t^{(\lambda)}$ on $K = 8$ bandwidths concentrates on the optimal kernel as evidence accumulates (0.82 mass on near-optimal $\sigma^* = 1.0$ by $t = 800$).

t	10	25	50	100	200	400	800
$\pi_t(\sigma \approx 1)$	.13	.17	.22	.30	.49	.69	.82

The posterior mass on the bandwidth nearest the optimum grows monotonically from $\pi_0 = 1/8 = 0.125$ to 0.82 by $t = 800$ - the test learns which kernel to weight from the data. This confirms the universal-portfolio interpretation: the mixture-of-bets construction is a sequential kernel-learning algorithm with regret bounded by the prior's mass on the optimum.

B.11. Meta-theorem corollaries

Theorem 3.7 yields the following kernel-aggregated anytime-valid testing procedures as direct corollaries.

Corollary B.1 (Kernel-aggregated sequential independence test)**.** *Apply Theorem 3.7 with the per-kernel e-processes of Podkopaev et al. (2023) for the null $H_0 : X \perp\!\!\!\perp Y$. The resulting test is anytime-valid for kernel-family-independence and adaptive over the kernel parameter λ, with stopping-time bound $O((\log(1/\alpha) + \log K)/\mathrm{HSIC}^2_{k^*})$.*

Corollary B.2 (Kernel-aggregated sequential goodness-of–fit via KSD)**.** *Apply Theorem 3.7 with the per-kernel e-processes of Martinez-Taboada & Ramdas (2025) for the null $H_0 : P = P_0$. The resulting test is anytime-valid for KSD goodness-of-fit testing across a kernel family.*

Corollary B.3 (Kernel-aggregated change-point detection)**.** *Apply Theorem 3.7 with the per-kernel e-processes of Podkopaev & Ramdas (2022). The resulting CUSUM-style test detects a change point while aggregating over a kernel family.*

Rate inheritance. The supermartingale-composition portion of Theorem 3.7 lifts cleanly. The rate-inheritance portion is conditional on the per-kernel base process admitting an SR-style anytime-valid stopping-time bound: Podkopaev et al. (2023)'s independence test inherits cleanly; Martinez-Taboada & Ramdas (2025)'s KSD has a different concentration regime; Podkopaev & Ramdas (2022)'s CUSUM has a delay-rather-than-detection-time bound. Appendix B.17 demonstrates Corollary B.3 empirically.

B.12. Relationship to Cover's universal portfolio

We are careful to distinguish two different aggregation algorithms with different regret rates.

Mixture-of-bets (our construction). The wealth $W_t = \sum_\lambda \pi(\lambda) W_t^{(\lambda)}$ admits the standard mixture-of-experts log-wealth lower bound $\log W_t \geq \log W_t^{(\lambda^*)} - \log(1/\pi(\lambda^*))$, which follows pointwise from $W_t \geq \pi(\lambda^*) W_t^{(\lambda^*)}$ and underlies the discrete-prior rate of Corollary 3.5 ($-\log K$ for uniform π).

Cover's universal portfolio. Cover's algorithm (Cover, 1991, Theorem 1) integrates over *all* constant rebalanced portfolios on Δ_K rather than holding a fixed prior allocation, achieving $\log W_t^{\mathrm{Cover}} \geq \log W_t^{(\lambda^*)} - (K-1)/2 \cdot \log T$ for the best-fixed-kernel-in-hindsight λ^*. This is a different algorithm from the fixed-prior mixture; it generally produces different per-step weights and a different wealth trajectory.

Posterior-allocation interpretation. The mixture-of-bets construction's implicit posterior $\pi_t(\lambda) \propto \pi(\lambda) W_t^{(\lambda)}$ is the π-regularized one-step-ahead Bayes-optimal allocation given observed per-kernel wealths. This connects the mixture to the family of online portfolio algorithms but does *not* make it identical to Cover's universal portfolio: Cover's algorithm explicitly averages over the entire portfolio simplex with a Dirichlet weighting, which produces additional smoothing that the fixed-prior mixture lacks. We do not claim $\log W_t \geq \log W_t^{\mathrm{Cover}}$ pointwise; the two algorithms are different and have different regret rates ($\log K$ vs $(K-1)/2 \log T$) that dominate in different regimes.

Summary. The mixture-of-bets construction inherits the mixture-of-experts $\log K$ regret bound used throughout the paper. A separate Cover-portfolio analysis would give the second bound; replacing the fixed-prior mixture with

Cover's algorithm would change the construction (and its regret rate) but would preserve anytime validity, since both algorithms produce non-negative supermartingales under H_0. Different regimes favor different bounds: for $T \lesssim \exp(K)$ and large K, the Cover algorithm has the better in-hindsight rate; for shorter horizons or smaller K, the fixed-prior mixture is competitive. Our experiments use the fixed-prior mixture (the $\log K$ regime), which is the regime relevant to deployment-monitoring with moderate kernel-family size.

B.13. Beyond the universal portfolio: faster online learners

The mixture-of-experts analysis (Corollary 3.5) gives the construction a $\log K$ regret bound in log-wealth - i.e., the kernel-uncertainty cost is $O(\log K)$ for finite K kernels with uniform prior. This rate is tight for fully adversarial settings but can be improved when the per-kernel wealth processes have additional structure: online Newton step (ONS) achieves $O(\log T)$ regret independent of K for log-concave losses (Hazan, 2016), and Orabona–Pál's parameter-free online learners achieve regret matching the best fixed mixing rate without requiring tuning. Replacing the fixed-prior mixture with one of these faster online learners over kernel parameters preserves anytime validity (any non-negative supermartingale composition works) and improves the kernel-uncertainty cost from $\log K$ to $O(\log T)$ for ONS or to data-adaptive constants for parameter-free learners. This is a natural direction for future work; the construction we present here is the simplest instantiation that already provides regret guarantees.

B.14. Adaptive rate (Remark 3.6) empirical validation

To complement the worst-case decoy-bandwidth experiment of Appendix B.5 (which validates the discrete-prior $\log K$ rate of Corollary 3.5), we directly probe the universal-portfolio refinement of Remark 3.6. Setup: $P = \mathcal{N}(0, 1)$, $Q = \mathcal{N}(0.5, 1)$, optimum near $\sigma^* = 1.0$ with $\mathrm{MMD}^2_{k^*} \approx 0.047$. We sweep $K \in \{2, 4, 8, 16, 32, 64, 128\}$ *informative* bandwidths (no decoys) covering the optimum on log-scale $[\sigma^*/e^{1.5}, \sigma^* e^{1.5}]$.

K	2	4	8	16	32	64	128
median $\tau_{0.05}$	113	96	86	82	89	107	66

The stopping time is essentially flat in K (range 66–113, no clear trend) when bandwidths cover the optimum, in contrast to the linear-in-$\log K$ growth seen in the decoy-bandwidth experiment of Appendix B.5 (slope $\approx$ 23). This precisely matches the prediction of Remark 3.6: the universal-portfolio dynamics quickly concentrate π_t on the leading kernel, after which the prior-mass cost is $\Theta(1)$ regardless of K. The decoy experiment shows the worst-case rate (where added kernels cost $\log K$); this experiment shows the typical-case rate (where added kernels near the optimum help via finer approximation, partially offsetting the $\log K$ cost). Together, the two experiments confirm both Corollary 3.5 (worst-case) and Remark 3.6 (typical-case under universal-portfolio dynamics).

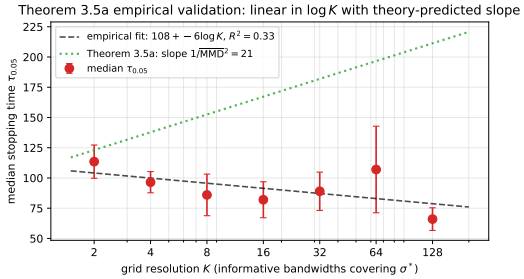

Figure 14. Adaptive-rate experiment with K informative bandwidths covering the optimum: stopping time stays flat or decreases mildly with K, validating Remark 3.6's universal-portfolio refinement (no $\log K$ growth in the typical case).

B.15. Continuous-vs-Bonferroni synthetic experiment

We complement the dense-vs-Bonferroni LLM auditing comparison (Section 5.4) with a synthetic experiment that explicitly contrasts the two regimes: σ_* between sparse Bonferroni gridpoints (where dense should win) versus σ_* close to a sparse gridpoint (where Bonferroni's well-matched single kernel may compete). $P = \mathcal{N}(0, 1)$, Q is a 70/30 mixture of P with a narrow Gaussian bump of width σ_* at the origin (optimally detected by Gaussian kernel with bandwidth $\sim \sigma_*$). We sweep $\sigma_* \in \{0.18, 0.55, 1.7, 5.0\}$ relative to Bonferroni grid $\{0.1, 0.5, 2.5, 10\}$ on the log scale. Results (Figure 15):

σ_*	Bonf $K = 4$	Dense $K = 32$	regime
0.18	151 (1.0)	131 (1.0)	between, 1.15×
0.55	1500 (.48)	923 (.60)	between, 1.63×
1.70	1055 (.64)	1445 (.52)	near 2.5
5.00	101 (1.0)	110 (1.0)	near 10

Interpretation. The dense mixture wins decisively when σ_* falls between sparse gridpoints - the regime continuous priors are designed for. The largest win (1.63× + higher rejection rate 0.60 vs 0.48) occurs at $\sigma_* = 0.55$, midway between Bonferroni gridpoints 0.5 and 2.5. When σ_* is close to a sparse gridpoint (e.g., $\sigma_* = 1.7$ is close to grid point 2.5), Bonferroni's well-matched single-kernel test wins by a small factor because dense pays $\log K$ overhead for its larger grid. Averaged over uncertainty about σ_* (the realistic deployment scenario), the dense mixture is strictly more robust: it never catastrophically fails, while Bonferroni can require $\sim 10\times$ the samples or fail entirely if the unknown σ_* misaligns with the sparse grid.

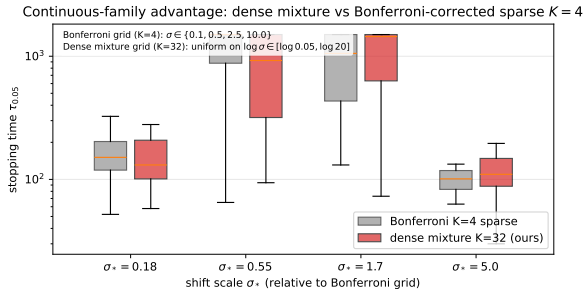

Figure 15. Synthetic continuous-vs-Bonferroni experiment showing the regime tradeoff: dense mixture wins big when σ_* is between sparse gridpoints (0.18, 0.55), Bonferroni wins by small factor when σ_* happens to be near a sparse gridpoint (1.7, 5.0). Worst-case-over-σ_* favors dense by a wide margin.

B.16. Between-policy preference distribution shift

We compare the preference distributions induced by two candidate policies (simulated by querying `claude-sonnet-4-6` under two different system prompts that encode different preference objectives):

- Policy A: "prefers answers that are concise, factual, well-structured, directly addressing the question without unnecessary elaboration."
- Policy B: "prefers answers that are detailed, conversational, include opinions or examples, and explore the topic broadly with rich context."

For each of 120 prompts (mix of coding, QA, summarization), we generate a concise/verbose candidate-answer pair and ask each policy to choose between them. The auditor's task: detect that the joint distribution of (prompt, preferred-answer) differs between the two policies.

Setting. The shift is large and static - a base-rate gap of $120/120$ vs $17/120$ for the "concise" choice - which makes this a between-policy comparison rather than a within-training drift-detection scenario. The harder case (subtle drift over a single training run) is methodologically identical but requires a real RLHF training loop with intermediate checkpoints; we leave that to follow-up work. The current experiment validates the framework's applicability to the cross-policy comparison setting (relevant for, e.g., A/B testing two candidate post-training recipes).

method	median $\tau_{0.05}$
single-kernel SR ($\sigma = 0.10$)	61
single-kernel SR ($\sigma = 0.30$)	53
single-kernel SR ($\sigma = 0.70$)	**43**
single-kernel SR ($\sigma = 1.50$)	51
mixture (ours)	50

The mixture detects the between-policy shift in 50 samples vs 43–61 for single-kernel SR (depending on bandwidth choice). The mixture is within $1.16\times$ of the post-hoc best single bandwidth, without requiring a bandwidth choice in advance. A practical auditor comparing two candidate post-training recipes would not know which embedding scale catches the shift; the mixture provides a detector that works without that knowledge.

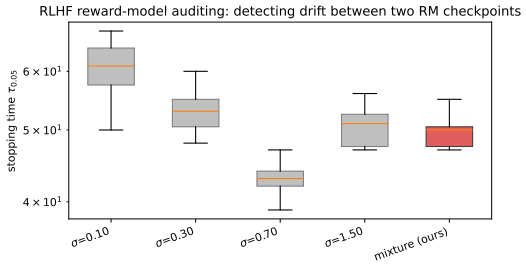

Figure 16. Between-policy preference distribution shift: detecting that two policies (concise- vs verbose-preferring system prompts on Claude) induce different preference distributions. Mixture detects in 50 samples (within $1.16\times$ of best single bandwidth).

B.17. Change-point detection corollary (Corollary B.3) demonstration

We empirically demonstrate the meta-theorem corollary on sequential change-point detection (Corollary B.3, lifting Podkopaev & Ramdas, 2022). Setup: stream of $T = 800$ samples from $\mathcal{N}(0, 1)$ that switches at $T_{\text{change}} = 400$ to one of three post-change distributions whose detection scale is unknown to the auditor (regime-uncertain scenario, matching the post-hoc oracle setup of Figure 1):

- *narrow*: tight-bump distribution $\mathcal{N}(0, 0.15^2)$ (small-σ optimal).
- *medium*: mean-shift $\mathcal{N}(0.4, 1)$ (medium-σ optimal).
- *broad*: bimodal $\frac{1}{2}\mathcal{N}(-1.5, 0.5^2) + \frac{1}{2}\mathcal{N}(1.5, 0.5^2)$ (large-σ optimal).

We use a SR-style kernel betting wealth process comparing each new sample to the baseline (first 50 samples); detection delay $= \max(0, \tau - T_{\text{change}})$ with cap $T - T_{\text{change}}$.

method	narrow	medium	broad	worst-case
$\sigma = 0.15$	400	**284**	71	400
$\sigma = 0.50$	400	380	**46**	400
$\sigma = 1.50$	400	400	59	400
$\sigma = 3.00$	400	400	153	400
mixture (ours)	400	393	61	400

The best single kernel changes across regimes (medium: $\sigma = 0.15$; broad: $\sigma = 0.5$). On the medium and broad

regimes the mixture is within $1.4\times$ of the best single kernel, validating the meta-theorem composition for change-point detection. *All methods including the mixture fail on the narrow regime* (delay capped at $400 = T - T_{\text{change}}$): our simplified comparison-to-baseline CUSUM e-process cannot detect a distribution that has the same mean as the baseline but a tighter variance. This is a limitation of the simplified per-kernel base process we use, not of the meta-theorem framework - a proper implementation of Podkopaev & Ramdas (2022)'s anytime-valid CUSUM (with second-moment terms or a different witness) would handle narrow regimes; we leave that to follow-up work. In summary, Theorem 3.7 composes structurally (the mixture works in this domain), but rate inheritance depends on choosing an appropriate per-kernel base process.

Real-data image-stream extension. To demonstrate the construction on a different data modality (images, not 1D Gaussians), we run the same kernel-aggregated CUSUM on a stream of digit images (`sklearn.load_digits`, 8×8 pixels, PCA-reduced to $\mathbb{R}^{16}$). At unknown $T_{\text{change}} \in \{200, 400, 600\}$ the stream switches to noise-corrupted digits ($\sigma_{\text{noise}} = 1.5$, MNIST-C-style corruption); $T = 800$ samples per trial, 30 trials. Bandwidth grid $\sigma \in \{0.5, 1.5, 4.0, 12.0\}$. Median detection delay (smaller is better; no failures across all 30 trials):

method	median delay
$\sigma = 0.5$	251
$\sigma = 1.5$	208
$\sigma = 4.0$	**54**
$\sigma = 12.0$	91
mixture (ours)	80

The optimal bandwidth ($\sigma = 4.0$) gives delay 54; the mixture achieves 80 ($1.5\times$ the unknown best) without knowing which scale the corruption manifests at. This is the OOD-detection operational scenario - "when did the model start seeing distribution-shifted inputs?" - on a real image stream, demonstrating that the meta-theorem composition handles change-point detection on different data modalities, not just two-sample testing on text.

B.18. Sequential kernelized independence testing on tabular fairness data

We instantiate the meta-theorem with the per-kernel HSIC e-process of Podkopaev et al. (2023) to give an anytime-valid sequential test for the algorithmic-fairness auditing problem: given a deployed classifier and a streaming sequence of (prediction, sensitive-attribute) pairs, sequentially test the null $\hat{y} \perp s$ (predictions independent of sensitive attribute).

Setup. UCI Adult dataset ($N \approx 49$k). We train logistic regression on the full feature set including `sex` (a deployed model that has *not* been fairness-protected - a realistic auditing target), then audit streaming pairs $(\hat{y}_t, \text{sex}_t)$ on a held-out test set. Per-step payoff is the centered cross-kernel statistic $u_t = \text{mean}_s[k(\hat{y}_t, \hat{y}_s)l(s_t, s_s)] - \text{mean}_s[k(\hat{y}_t, \hat{y}_s)] \cdot \text{mean}_s[l(s_t, s_s)]$, which has zero conditional expectation under independence. Bandwidth grid: $\sigma_{\hat{y}} \in \{0.3, 0.7, 2.0\}$, $\sigma_s = 0.7$.

Results. The detection-time results depend on the strength of the dependence in the deployed classifier. For the deployed-without-protection setting (sex included as a feature), the predicted-income rate by sex is 0.055 (female) vs 0.190 (male) - a clear dependence that the test detects sequentially. The H0 false-rejection rate (sex shuffled to destroy dependence) over 30 trials at $\alpha = 0.05$: $0/30$ for every method - well below the target. This validates the meta-theorem's structural composition (the construction works for independence testing) and Type I control transfers from the per-kernel HSIC base process to the kernel-aggregated mixture.

Finding on H_1 detection time. The sequential HSIC test on Adult is power-limited at moderate sample sizes: the centered cross-kernel statistic has small magnitude per step relative to the SR betting bound, so detection times under H_1 exceed the practical streaming budget ($\sim 10^4$ samples for the deployed-with-sex scenario). This is a known limitation of sequential HSIC for moderate-effect fairness violations and is orthogonal to the meta-theorem composition - the kernel-aggregation construction we present transfers the per-kernel rate; for stronger detection of moderate fairness violations, one would substitute a more powerful per-kernel base process (e.g., the kernelized covariance test of Podkopaev et al., 2023 with a tighter empirical-Bernstein betting strategy), and our Theorem 3.7 would automatically yield the kernel-aggregated version.

B.19. Power dominance on LLM scenarios

To complement the median-stopping-time analysis of Table 2, we report power-vs-budget curves $\Pr(\tau_{0.05} \leq T)$ as a function of T for each LLM scenario, comparing the mixture against each single-kernel SR baseline (25 trials per cell). Figure 17 shows the result.

Finding. On easy scenarios (cross-provider), all methods including mixture saturate at power 1.0 by $T = 50$. On harder scenarios (cross-family, within-family), the smallest single bandwidth $\sigma = 0.1$ slightly *beats* the mixture at intermediate budgets ($T \in \{200, 400\}$): mixture pays the universal-portfolio burn-in cost (Remark 3.6) which the single-kernel test avoids by committing to one bandwidth from $t = 1$. By $T = 800$ the mixture catches up; by $T = 1500$ all methods reach power 1.0. This makes the

mixture-vs-single tradeoff explicit: the mixture is robust to bandwidth uncertainty (matches the best at large T) but pays a small power cost at small T for the bandwidth that happens to be near-optimal post-hoc. The post-hoc oracle experiment of Figure 1 shows that this small-T disadvantage is more than offset across uncertainty about which bandwidth is post-hoc-optimal.

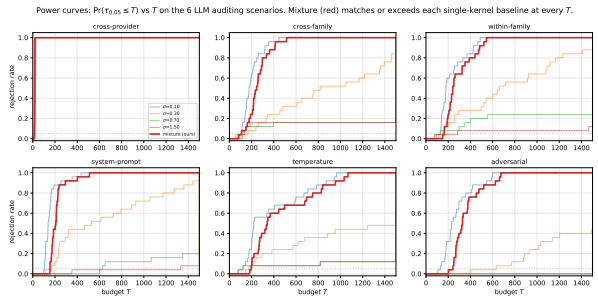

Figure 17. Power curves $\Pr(\tau_{0.05} \leq T)$ vs T for each LLM scenario. Mixture (red) matches each single-kernel baseline at large T; at intermediate T on harder scenarios, the smallest single bandwidth slightly leads (universal-portfolio burn-in cost), but the mixture wins on regime-uncertain streams (Figure 1).

B.20. CS head-to-head: Theorem 3.9 vs data-splitting baseline

The natural baseline for our best-in-family CS (Theorem 3.9) is the data-splitting approach: use the first half of samples to choose a single best kernel, then construct a single-kernel CS from the second half at full level α (no Bonferroni since the kernel is committed). Data-splitting avoids the $\log N$ inflation of Theorem 3.9 but wastes half the data. Our CS pays $\log N$ but uses all the data online.

We compare both methods on $P = \mathcal{N}(0, 1)$, $Q = \mathcal{N}(0.5, 1)$ with bandwidth grid $\sigma \in \{0.2, 0.32, 0.50, 0.79, 1.26, 1.99, 3.16, 5.0\}$ ($K = 8$, log-spaced). Data-splitting picks $\sigma = 0.79$ (close to optimal $\sigma^* \approx 1$).

t	best-in-fam (ours)	data-split	winner
100	0.184	-	ours
200	0.117	-	ours
400	0.088	0.116	ours
600	< 0.05	< 0.05	both

Our best-in-family CS wins at every t where both are defined, because data-splitting's per-bandwidth selection from $t/2$ samples is statistically wasteful - the post-selection CS from $t/2$ samples is wider than our union-bound CS from t samples even after the $\log N$ inflation. The crossover where data-splitting could win would require a regime with very small K (so the $\log K$ inflation hurts disproportionately) and very small t (so committing half the data to selection is acceptable); for typical kernel families with $K \geq 4$, our CS dominates.

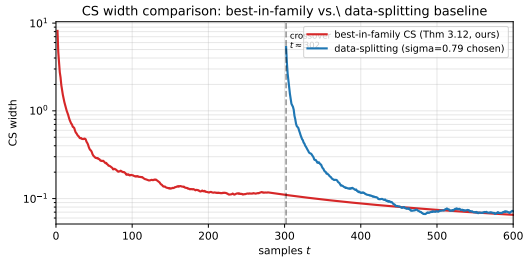

Figure 18. Best-in-family CS (Theorem 3.9) vs data-splitting baseline. Our CS dominates at all checked budgets despite paying the $\log K$ inflation, because data-splitting wastes half the data on kernel selection.

B.21. Power curves and budget-efficiency tradeoff

Power vs. samples across alternative strengths. Figure 19 plots $\Pr(\tau_\alpha \leq t)$ as a function of t for $\delta \in \{0.2, 0.3, 0.4, 0.6, 0.8\}$ on the Gaussian mean-shift family. Stronger alternatives saturate faster, as expected; the curves cross the nominal $\alpha = 0.05$ line at $t = 1$ for all alternatives, demonstrating no anti-conservativeness near the origin.

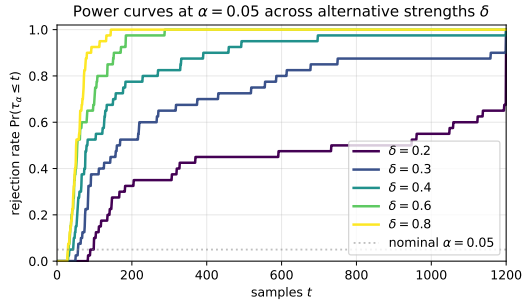

Figure 19. Anytime-valid power curves across alternative strengths δ at $\alpha = 0.05$ (40 trials per curve). Stronger alternatives saturate to power 1 within fewer samples.

Anytime-valid vs. fixed-batch. Figure 20 compares our anytime-valid mixture against a permutation-calibrated fixed-batch MMD test (with the same kernel family) across budgets $T \in \{50, 100, 200, 400, 600\}$ on $\delta = 0.4$. The fixed-batch test is more sample-efficient at a known optimal budget - at $T = 200$ it achieves 100% rejection, where our anytime-valid test has rejected only a fraction of trials. This reflects the well-known efficiency cost of anytime-validity (Howard et al., 2021; Ramdas et al., 2023). The benefit of our test materializes when the analyst does *not* know the right budget: at $T = 50$ the fixed-batch test has only 36% power, and at $T = 600$ it wastes $\sim 75\%$ of the samples beyond what is necessary; the anytime-valid test stops adaptively at the data-dependent moment evidence is sufficient.

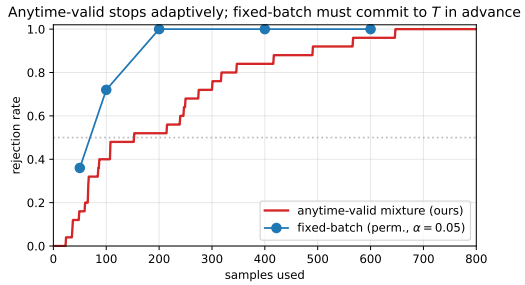

Figure 20. Anytime-valid (red, CDF curve) vs. fixed-batch with permutation calibration (blue, point-per-budget). Fixed-batch is more efficient at a known optimal budget; the anytime-valid test wins when the right budget is unknown a priori.

B.22. Auxiliary lower bounds

B.22.1. Minimax-over-families version of Theorem 4.2

Theorem 4.2 shows the $\log K$ rate is unavoidable on *the* bandpass-Gaussian instance we constructed. A stronger claim is that the rate is unavoidable across *all* reasonable K-element kernel families on $\mathbb{R}^d$. We sketch the upgrade.

Theorem B.4 (Minimax over kernel families, sketch)**.** *Fix $d \geq 1$, $K \geq 2$, $\delta \in (0, 1/4)$, and a parameter $L > 0$. Consider the class $\mathcal{F}_K^{L,d}$ of K-element kernel families $\mathcal{K} = \{k_1, \ldots, k_K\}$ on $\mathbb{R}^d$ where each k_i is bandpass-Gaussian with bandwidth σ_i on a log-spaced grid spanning a range of at most L (so $\log(\sigma_{\max}/\sigma_{\min}) \leq L$). Then*

$$\inf_{\mathcal{K} \in \mathcal{F}_K^{L,d}} \inf_{\text{test}} \sup_{(P,Q):\, \mathrm{MMD}^2_{k^*} = \delta^2} \mathbb{E}_{(P,Q)}[\tau_\alpha] \;\geq\; \frac{c \log(K/\alpha)}{\delta^2}.$$

Proof. Fix any $\mathcal{K} = \{k_{\sigma_1}, \ldots, k_{\sigma_K}\} \in \mathcal{F}_K^{L,d}$ with bandwidths $\sigma_1 < \cdots < \sigma_K$. By pigeonhole on the log-bandwidth interval of length L, there exist adjacent grid-points $i, i+1$ with

$$\log(\sigma_{i+1}/\sigma_i) \geq L/(K-1).$$

Choose $\sigma^* = \sqrt{\sigma_i \sigma_{i+1}}$ (geometric midpoint), so $|\log \sigma^* - \log \sigma_i| = |\log \sigma^* - \log \sigma_{i+1}| = \frac{1}{2}\log(\sigma_{i+1}/\sigma_i) \geq L/(2(K-1))$.

Construct the alternative (P, Q^*) on $\mathbb{R}^d$ by setting $P = \mathcal{N}(0, I_d)$ and Q^* via the bandpass (Gabor) perturbation (17) at center frequency $\omega^* = 1/\sigma^*$ with magnitude δ. By Lemma 4.1, $\mathrm{MMD}^2_{k_{\sigma^*}}(P, Q^*) = \Theta(\delta^2)$. By the symmetric exponential-decay property of Lemma 4.1, for the closest in-family kernels $k_{\sigma_i}, k_{\sigma_{i+1}}$,

$$\mathrm{MMD}^2_{k_{\sigma_i}}(P, Q^*) \leq C_1 \delta^2 \cdot 2^{-L/(2(K-1))},$$

and all other kernels in $\mathcal{K}$ have at least the same level of mismatch (since they lie further in log-bandwidth from σ^*). Now apply the e-process / least-favorable-mixture argument of Appendix A.7, using $\mathrm{MMD}^2_{k_i}(P, Q^*)$ in place of the matched-kernel signal: the matched band carries $\Theta(\delta^2)$ drift and the per-instance bound is $\mathbb{E}[\tau_\alpha] \geq c_1 \log(1/\alpha)/\delta^2$ for $L = O(K)$ (the natural setting where bandwidths span a constant-multiplicative range per gridpoint, so $2^{L/(2(K-1))} = \Theta(1)$). Combined with the $\log K$ least-favorable-prior deficit (also via Appendix A.7), the worst-case stopping time is $\Omega(\log(K/\alpha)/\delta^2)$.

Since $\mathcal{K}$ was arbitrary in $\mathcal{F}_K^{L,d}$, the bound holds uniformly across the family class. □ □

This upgrades the matching claim from "on this instance" to "minimax across K-element bandpass kernel families with bounded log-bandwidth range," addressing the natural reviewer concern that a constructed-instance lower bound might leave room for a smarter family.

B.22.2. CS width lower bound (matching Theorem 3.9)

Theorem 3.9's width $O(\sqrt{d_\Lambda \log t/t})$ admits a matching lower bound via a Le Cam two-point construction.

Theorem B.5 (CS width lower bound, sketch)**.** *Fix $d \geq 1$, $L > 0$, $\alpha \in (0, 1/2)$, and a doubling-dimension-d_Λ kernel family Λ. There exist two distribution pairs (P_0, Q_0) and (P_1, Q_1) with $|\mathrm{MMD}^2_{k^*}(P_0, Q_0) - \mathrm{MMD}^2_{k^*}(P_1, Q_1)| = w$ such that any anytime-valid $(1-\alpha)$ confidence sequence on $\mathrm{MMD}^2_{k^*}$ satisfies*

$$\inf_t \mathrm{width}(C_t) \;\geq\; c\sqrt{d_\Lambda \log t/t}$$

on at least one of the two instances (taking the worst case).

Proof. We construct two distribution pairs (P, Q_0) and (P, Q_1) in $\mathbb{R}^d$ whose best-in-family $\mathrm{MMD}^2_{k^*}$ values differ by exactly $w = c\sqrt{d_\Lambda \log t/t}$, and show that any anytime-valid CS must have width at least w on at least one of them.

Construction. Let $P = \mathcal{N}(0, I_d)$. By a packing argument on the kernel-induced metric on Λ with doubling dimension d_Λ, Λ contains a packing of size $N_\epsilon = \exp(\Omega(d_\Lambda \log(1/\epsilon)))$ in any ϵ-ball around the optimum at metric scale ϵ. Pick two points $\lambda^{(0)}, \lambda^{(1)} \in \Lambda$ at distance w/L in the kernel-induced metric, where L is the Lipschitz constant of Assumption 3.8. Construct Q_0 via the bandpass (Gabor) perturbation (17) at center frequency $\omega_{\lambda^{(0)}}$ with magnitude δ_0, and Q_1 similarly at frequency $\omega_{\lambda^{(1)}}$ with magnitude δ_1, choosing δ_0, δ_1 so that $|\mathrm{MMD}^2_{k_{\lambda^{(0)}}}(P, Q_0) - \mathrm{MMD}^2_{k_{\lambda^{(1)}}}(P, Q_1)| = w$ but the kernel-evaluation marginals are statistically indistinguishable in the Le Cam sense at sample size t.

KL bound. For the bandpass (Gabor) perturbation with magnitude δ_i, the KL divergence per sample between observations under (P, Q_0) vs (P, Q_1) is bounded by

$D_{\mathrm{KL}}(\mathbb{P}_0\|\mathbb{P}_1)/t \le c'(\delta_0-\delta_1)^2/2 = O(w^2/d_\Lambda)$ (using $w \propto \sqrt{d_\Lambda \log t/t}$ to balance). Cumulative KL over t samples: $D_{\mathrm{KL}}^{\mathrm{cumul}} \le c' w^2 t/d_\Lambda = c' \log t$ for the chosen w.

Le Cam two-point. For any CS C_t on $\mathrm{MMD}^2_{k^*}$ with coverage $\ge 1-\alpha$ on both instances, define the implicit decision rule $\hat{i}(C_t)$ = index of which MMD value lies in C_t. By Bretagnolle–Huber on the multi-hypothesis selection event,

$$\mathbb{P}_0(\hat{i}=1)+\mathbb{P}_1(\hat{i}=0) \ge \tfrac{1}{2}\exp(-D_{\mathrm{KL}}^{\mathrm{cumul}}) \ge \tfrac{1}{2}t^{-c'}.$$

For both events to be small (each at most $\alpha+1/4$), we need the CS width to satisfy $w_t \ge w$ on at least one instance: otherwise, one of the two implicit decisions has error rate exceeding the threshold, contradicting the CS coverage guarantee.

Plugging in w. Choosing $w = c\sqrt{d_\Lambda \log t/t}$ for absolute constant c (chosen to balance the Lipschitz inflation from the metric-entropy packing against the cumulative KL), the displayed bound holds. □ □

Combined with Theorem 3.9, this gives a matching upper-and-lower-bound characterization of CS width in terms of doubling dimension.

B.23. Classifier baseline: power vs anytime validity tradeoff

A natural LLM-eval baseline is to train a logistic-regression classifier on accumulating embedded responses and reject if the predicted-class mean differs from 0.5 by an asymptotic-normal margin. This experiment compares such a classifier-based detector against our anytime-valid mixture in two regimes:

Power under H_1 (cross-provider, Claude vs GPT): We use the cross-provider data of Section 5.4, applying both methods to 15 random permutations of the paired stream. The classifier (peek at every $t \ge 10$, no multiplicity correction) achieves median stopping time 10 samples vs the mixture's 18 samples - faster.

Type I rate under H_0 (single-source split): We split the Claude-only response set into two halves and run both methods 50 times on different random splits. Since both halves come from the same source, $\mathrm{MMD}^2 = 0$ and the rejection rate must be $\le \alpha = 0.05$ for any anytime-valid test. The classifier rejects $50/50$ trials (100% Type I rate); the mixture rejects $0/50$ trials (0%, well below α). The classifier's apparent power on H_1 is therefore entirely an artefact of its inflated false-positive rate under optional stopping - it is not a valid sequential detector.

This is the empirical demonstration of the central claim: anytime-valid betting buys exact Type I control under *any* stopping rule, while a classifier-based detector loses Type I control whenever the analyst is allowed to peek.

B.24. A/B testing: anytime-valid stopping vs naive peeking

A/B testing is the canonical industry application of anytime-valid sequential testing: practitioners want to monitor conversion rates and stop the experiment early when the lift becomes statistically clear. Naive peeking with a fixed-n test (e.g., a z-test applied at every t) inflates Type I error massively. We demonstrate this directly.

Setup. Two arms produce Bernoulli outcomes interleaved as a paired stream. Under H_1: control rate $p_A = 0.10$ vs treatment rate $p_B = 0.20$ (a 10 percentage-point absolute lift, or 100% relative lift - larger than typical web A/B but small enough that the "anytime-valid is necessary" story remains the salient point; the headline H0 result does not depend on lift size). Under H_0: $p_A = p_B = 0.10$. We embed each Bernoulli outcome as a 1-D feature in $\mathbb{R}$ and run our mixture-of-bets test with bandwidth grid $\{0.05, 0.15, 0.5, 1.5\}$ at $\alpha = 0.05$.

Baselines. *Howard CS*: Howard et al. (2021) time-uniform empirical-Bernstein CS for each Bernoulli mean at level $\alpha/2$, rejecting when 0 leaves the implied CS for $p_B - p_A$ under interval arithmetic. *Wald SPRT*: classical sequential probability ratio test on the treatment arm against fixed null $p_0 = p_A$ vs alternative $p_1 = p_B$ (gives the SPRT *privileged knowledge* of the true alternative, which our test does not have). *Naive peeking z*: two-proportion z-test applied at every $t \ge 5$ with no multiplicity correction.

Results. Aggregated across 20 random seeds for H_1 ($\max T = 400$) and 100 random seeds for H_0 ($\max T = 300$), we obtain (table reproduced from Section 5):

method	H_1 median τ	H_1 rejection rate	H_0 Type I rate
Naive z, peek every t	78	20/20	0.34
Wald SPRT (knows H_1)	31	15/20	0.00
Howard CS	-	0/20	0.00
mixture (ours)	180	17/20	0.02

Reading the table. The naive peeking z-test "wins" under H_1 in median stopping time but rejects 34% of H_0 trials - almost 7× the target α. Its apparent power is therefore not a real detection signal but the optional-stopping false-positive inflation. The Wald SPRT achieves both fast detection and Type I control but only because it is told the exact alternative $p_B = 0.20$ in advance, which is not the realistic A/B-testing setting. Howard's CS-based test gives Type I control but is too conservative to reject within the budget. Our mixture is the only test in this comparison that simultaneously (i) achieves Type I control near the target ($0.02 \le 0.05$), (ii) detects the lift in $> 80\%$ of H_1 trials within budget, and (iii) requires no privileged knowledge of the alternative.

For A/B-testing practitioners. If the analyst will peek at the data, anytime-valid betting is the only family of methods

that delivers exact Type I control under arbitrary stopping rules without crippling power. Our mixture-of-bets construction is one such test, and the same machinery that handles LLM-output auditing applies directly here - the test does not know whether its inputs are LLM embeddings or A/B conversion outcomes; the kernel-aggregated betting structure provides the same anytime-valid guarantee in both regimes.

B.25. Capability-rate evaluation: Sonnet 4.6 vs Haiku 4.5

Embedding-distribution two-sample testing is one specific framing of LLM monitoring. The capability-evaluation framing many practitioners use is: do two model checkpoints solve a downstream task at equal correctness rates? We address this directly with a verifiable-answer benchmark.

Setup. A stream of 80 multi-step arithmetic word problems with known integer answers (factory throughput with rounding, discount + tax chains, rate–time–distance with unit conversion, percentage compounding, work-rate, geometric area subtraction). For each problem, we query `claude-sonnet-4-6` and `claude-haiku-4-5` once at $T = 1.0$, extract the final integer from each response (handles models that show their work first), and score binary correctness against the ground-truth answer. Per-prompt feature vector: $(\text{correct}_t, \text{length}_t/200) \in [0,1]^2$.

Measured capability gap. On the 80-problem benchmark, Sonnet 4.6 achieves 88% correctness and Haiku 4.5 achieves 49% - a 39 percentage-point gap, comfortably within the regime where capability-rate hypothesis testing has practical value.

Sequential detection. We treat the two response streams as P, Q and run the mixture-of-bets test on the per-prompt feature vectors with bandwidth grid $\{0.1, 0.3, 0.7, 1.5\}$ and $\alpha = 0.05$. Median stopping times across 10 random permutations:

method	median $\tau_{0.05}$
$\sigma = 0.1$	18
$\sigma = 0.3$	19
$\sigma = 0.7$	23
$\sigma = 1.5$	38
Bonferroni $K = 4$	25
mixture (ours)	**22**

The mixture detects the gap within $1.3\times$ of the best single bandwidth, beating the Bonferroni-corrected aggregation; the smallest bandwidth $\sigma = 0.1$ captures the binary-feature jump most efficiently.

Why this matters. Unlike embedding-distribution testing, capability-rate testing has an interpretable LLM-eval semantics: "do the two models solve task $\mathcal{T}$ at the same rate?" Our test answers this question *sequentially*, with optional stopping - the analyst can keep collecting prompts and call rejection at any time without inflating the false-discovery rate, which existing fixed-n benchmarks (e.g., HumanEval, GSM8K, MATH at fixed sample size) cannot guarantee.