ONE SHOT IS A GRADIENT: RAW-MEASUREMENT OPTIMIZATION FOR QUANTUM KERNEL MACHINES

Anonymous authors
Paper under double-blind review

ABSTRACT

A quantum kernel is observable only through shots: destructive circuit runs that each return one bit whose mean is the kernel entry. Standard training averages many shots per kernel entry, typically estimating the whole Gram matrix before optimizing, so measurement grows with the N^2 pairs of training examples. Yet a stochastic optimizer never needs the matrix. It needs an unbiased gradient, which a single shot already provides, or a yes/no hinge decision, which sequential shots can certify. We therefore train the exact kernel SVM directly from raw shots. Raw-Shot Coordinate Ascent (RSCA) turns each shot into an unbiased coordinate gradient of the SVM dual, with a low-rank control variate that reduces noise without changing the solution. Aggregate-Margin-Oracle (AMO) Pegasos samples an example's margin until an anytime-valid test certifies its hinge decision. At fixed regularization, decision error ϵ costs $\widetilde{O}(\min\{N\epsilon^{-4}, \epsilon^{-6}\})$ shots, linear in N or independent of it. AMO becomes cheaper when few margins lie near the hinge, and a lower bound for each decision shows that no sequential test avoids this margin dependence. In shot-level simulations on MNIST with up to $N = 1024$ examples, RSCA's shot cost grows at most linearly in N. At $N = 1024$ it reaches a fixed error target with 2.5–$3\times$ fewer shots than our strongest entry-estimating baseline, fewer than one shot per pair, and its lead survives 1% depolarizing and readout noise. Strict AMO certifies every hinge decision at the per-decision cost that the lower bound predicts.

1 INTRODUCTION

Quantum feature maps turn a datum x into a quantum state, a unit vector $|\psi(x)\rangle$, and define the fidelity kernel, their squared overlap $K(x, x') = |\langle\psi(x) \mid \psi(x')\rangle|^2$. Kernel machines thereby gain features that may be intractable to compute classically (Schuld & Killoran, 2019; Havlíček et al., 2019). Hardware cannot read out K_{ij} directly: each run of an overlap circuit, a *shot*, returns one bit that equals one with probability K_{ij} and destroys the state. The usual interface treats the device as a noisy kernel calculator: it repeats the circuit about δ^{-2} times per entry for precision δ, fills the Gram matrix, and only then optimizes. With $\Theta(N^2)$ pairs (billions of runs for $N = 1000$ at $\delta = 0.01$) and entry errors that training can amplify, measurement becomes the dominant cost (Gentinetta et al., 2024; Shastry et al., 2023).

Existing approaches reduce this cost at different stages. Robust SVM formulations tolerate few-shot entries (Shastry et al., 2023), adaptive schemes allocate shots across entries (Xu et al., 2026; Miroszewski, 2026), and Nyström approximations and centroid methods measure fewer entries via restricted representations or a different predictor class (Williams & Seeger, 2001; Coelho et al., 2025; Tscharke et al., 2024). All of them keep the kernel entry as the unit of measurement. We ask instead: what must a measurement resolve to make progress on the fixed-kernel SVM objective?

It must resolve far less than an entry. The dual gradient of the SVM is a weighted sum of kernel entries, and each entry is the mean of one bit. Drawing one term by its weight and measuring it once therefore gives an unbiased gradient coordinate, just as one training example yields an unbiased stochastic gradient. A stochastic optimizer averages its noise across iterations anyway; estimating every entry first pays for that averaging twice. A primal hinge step, in contrast, needs to know

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

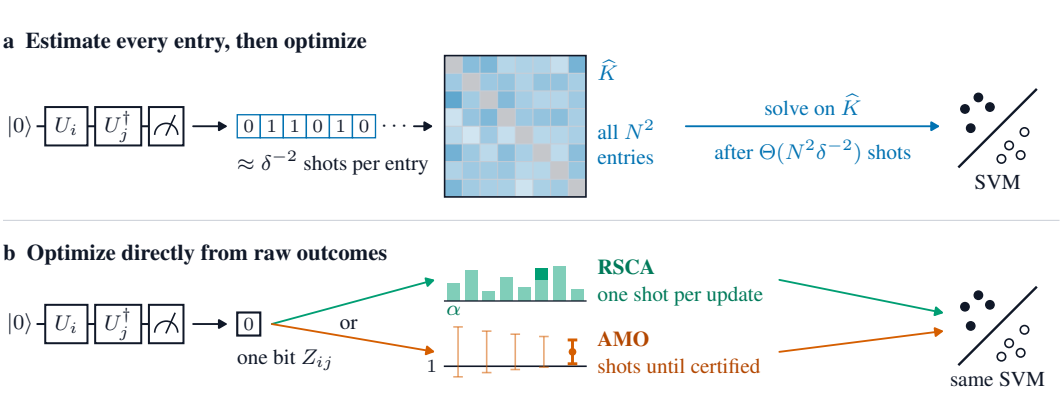

Figure 1: **Removing the estimation layer.** A shot is one bit Z_{ij} with mean K_{ij}. (a) The standard pipeline estimates every entry of K to precision δ (gray: known diagonal $K_{ii} = 1$) before training. (b) RSCA and AMO never form $\widehat{K}$. In RSCA each bit moves one coordinate of the dual vector α by an unbiased gradient step; AMO samples until its confidence interval for a margin clears the hinge at one, which takes few shots unless the margin is close to one.

only whether one margin lies below one. A threshold of a mean is not a mean, so the decision's shots feed a sequential test that certifies it. At fixed accuracy and regularization, the optimizer's shot count grows at most linearly in N, whereas entry estimation pays for every pair the optimizer touches, which in our experiments is nearly all N^2 of them. Figure 1 contrasts the two interfaces.

We turn these observations into two algorithms. *Raw-Shot Coordinate Ascent* (RSCA) scales one shot into an unbiased coordinate gradient of the exact SVM dual. Because a single bit is a noisy estimate, RSCA uses a frozen low-rank (Nyström) approximation of the kernel as a control variate: its part of the gradient is computed exactly, and each shot corrects only the remainder. *Aggregate-Margin-Oracle Pegasos* (AMO-Pegasos) treats each hinge decision as a sequential test. It keeps sampling an example's margin until a confidence interval, valid at any stopping time, lies on one side of one, up to a tolerance τ. Easy decisions cost a single shot; hard ones cost more. Both methods keep the original SVM objective. What changes is what the measurement budget buys: progress in training rather than kernel entries.

In optimization terms, RSCA is a stochastic form of dual coordinate ascent (Shalev-Shwartz & Zhang, 2013) driven by one pair-specific bit. Its frozen control takes the place of the exact snapshot gradient of SVRG (Johnson & Zhang, 2013), which would need accurate kernel sums. AMO combines Pegasos updates (Shalev-Shwartz et al., 2011) with a sequential test for each hinge decision (Howard et al., 2021). Term-sampled single-shot gradients of variational circuits (Sweke et al., 2020), doubly stochastic kernel gradients (Dai et al., 2014), and sampled weighted-sum inference (Gil-Fuster et al., 2026) use related primitives (Appendix C). Because every RSCA step is an unbiased gradient and every strict AMO action a certified τ-subgradient of the exact objective, standard stochastic-optimization theory counts the shots a target accuracy needs (Section 4).

Our contributions are:

- **A shot is a gradient.** Suitably weighted, one destructive measurement outcome is an unbiased stochastic gradient of the exact SVM dual, and a sequence of outcomes certifies a hinge decision. Training never needs the Gram matrix.
- **Variance control without changing the target.** A frozen low-rank control reduces gradient variance while the objective stays exact. Each shot corrects both ends of its pair and is shared by several step-size schedules. Centering at least halves the range of each AMO margin sample.
- **Shot complexity.** The analysis prices every shot and tells when to spend shots on directions (RSCA) and when on decisions (AMO). At fixed regularization, decision error ϵ costs $\widetilde{O}(N\epsilon^{-4})$ shots with RSCA and $\widetilde{O}(\epsilon^{-6})$ with AMO, whose rate is independent of N and

approaches $\widetilde{O}(\epsilon^{-2})$ when few margins lie near the hinge. A lower bound shows that AMO's Δ^{-2} cost for a margin at distance Δ from one is unavoidable (Section 4, Table 1).

- **Linear rather than quadratic measurement.** On three MNIST digit pairs with up to 1024 examples, RSCA's shot cost grows at most linearly in N, while our strongest entry-estimating baseline pays for nearly every pair. At $N = 1024$ RSCA needs 2.5–3$\times$ fewer shots, fewer than one per pair, and the lead persists under 1% depolarizing and readout noise. Strict AMO certifies every hinge decision at the cost the lower bound predicts (Section 5).

2 Setting and oracle model

Let $\mathcal{S} = \{(x_i, y_i)\}_{i=1}^N$, $y_i \in \{-1, +1\}$, and let K be a fixed positive-semidefinite kernel with $0 \le K(x, x') \le 1$ and $K(x, x) \le 1$, written $K(x, x') = \langle \phi(x), \phi(x') \rangle$ for a feature map ϕ. We study the regularized hinge objective without intercept

$$P(w) = \frac{\lambda}{2}\|w\|^2 + \frac{1}{N}\sum_{i=1}^{N}(1 - y_i\langle w, \phi_i\rangle)_+, \qquad \lambda > 0. \tag{1}$$

With $Y = \mathrm{diag}(y)$, its box-constrained dual and predictor are

$$D(\alpha) = \mathbf{1}^\top\alpha - \frac{1}{2\lambda}\alpha^\top YKY\alpha, \quad \alpha \in \mathcal{A} = [0, 1/N]^N, \qquad w(\alpha) = \lambda^{-1}\sum_j \alpha_j y_j \phi_j. \tag{2}$$

We report the decision error $\max_i |f(x_i) - f^\star(x_i)|$ of $f = \langle w(\alpha), \phi(\cdot)\rangle$, where $f^\star$ is the exact solution of Eq. 1 for the same kernel and regularization; the theory bounds the stronger error $\sup_{K(x,x)\le 1} |f(x) - f^\star(x)|$. Unlike test accuracy, this error asks whether training recovered the exact-kernel SVM itself rather than a different model of similar accuracy. We omit the intercept because an unregularized one would couple the dual coordinates; a regularized one reduces to a constant feature (Appendix D.1).

Raw-shot oracle. A quantum feature map is a circuit $U(x)$ that prepares $|\psi(x)\rangle = U(x)|0\rangle^{\otimes q}$ on q qubits. The inversion test runs $U(x_i)$, then the inverse circuit $U(x_j)^\dagger$, and measures every qubit. All qubits read zero with probability $|\langle\psi(x_j) \mid \psi(x_i)\rangle|^2 = K_{ij}$, so recording this event gives one bit

$$Z_{ij} \sim \mathrm{Bernoulli}(K_{ij}). \tag{3}$$

Each shot is a fresh run, and pairs may be chosen adaptively: conditional on the history $\mathcal{F}_{t-1}$ and the selected pair, we require $\mathbb{E}[Z_t \mid \mathcal{F}_{t-1}, I_t, J_t] = K_{I_tJ_t}$. The ideal fidelity kernel has $K_{ii} = 1$, so diagonal entries cost no shot; the noise study of Section 5 keeps this identity and charges the resulting bias to the methods.

Costs and access regimes. We count every circuit execution on the quantum processing unit (QPU), including those that build the control, and report distinct circuits, classical work, and memory separately. We study destructive pairwise measurement, the standard interface for fidelity kernels on current devices; randomized measurements (Huang et al., 2020; Haug et al., 2023) and amplitude estimation (Brassard et al., 2002; Montanaro, 2015) assume different access and costs (Appendix Table 3).

3 Raw outcomes as optimization primitives

Figure 2 summarizes the update flow of shared-shot RSCA and AMO-Pegasos; Algorithms 1 and 2 in Appendix E give pseudocode.

3.1 Raw-Shot Coordinate Ascent

The ith dual gradient is $g_i(\alpha) = 1 - \lambda^{-1}y_i\sum_j \alpha_j y_j K_{ij}$. Let $A = \sum_j \alpha_j \le 1$ be the total dual weight. If $A > 0$, draw J with probability α_j/A and obtain one Z_{iJ}; otherwise $g_i = 1$ exactly. Then

$$\widehat{g}_i = 1 - \frac{A}{\lambda}y_i y_J Z_{iJ}, \qquad \mathbb{E}[\widehat{g}_i \mid \alpha, i] = g_i(\alpha), \quad |\widehat{g}_i| \le B := 1 + \lambda^{-1}. \tag{4}$$

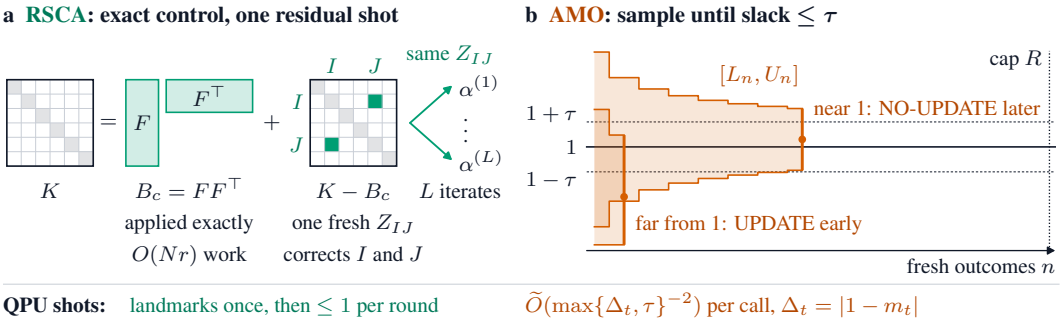

Figure 2: **From measurement to update.** (a) RSCA applies $B_c = FF^\top$ exactly; one fresh Z_{IJ} corrects the residual at both endpoints and is shared by all L trajectories. (b) AMO narrows $[L_n, U_n]$ until one action has hinge slack at most τ; budgeted AMO acts empirically at cap R. Curves are schematic; rates suppress λ and log factors.

RSCA draws I uniformly and, with step γ, performs $\alpha_I \leftarrow \mathrm{clip}_{[0,1/N]}(\alpha_I + \gamma \widehat{g}_I)$; drawing J takes $O(\log N)$ time (Appendix D.2).

Equation 4 is the interface change in its simplest form: since $\mathbb{E}[A y_J Z_{iJ} \mid \alpha, i] = \sum_j \alpha_j y_j K_{ij}$, the random partner and its bit together average to the exact kernel sum in g_i. The importance-weighted bit is then used once and discarded; no entry estimate is ever formed.

Control-variate direction. To reduce variance, RSCA measures the kernel against r landmark examples once, fits a Nyström factor $F \in \mathbb{R}^{N \times r}$, and freezes the control $B_c = FF^\top \approx K$. At the current iterate maintain $v = F^\top(\alpha \odot y)$ and the shot-free control gradient $g_i^{B_c} = 1 - \lambda^{-1} y_i F_i^\top v$. Unlike Eq. 4, draw I and J uniformly and independently, with one fresh Z_{IJ}, and define

$$\widehat{r}_I = -\frac{y_I}{\lambda} N \alpha_J y_J (Z_{IJ} - B_{c,IJ}) \quad \text{and symmetrically } \widehat{r}_J, \tag{5}$$

and use

$$h_t = \frac{1}{N} g^{B_c} + \frac{1}{2}(e_I \widehat{r}_I + e_J \widehat{r}_J). \tag{6}$$

Because the fresh residual corrects $K - B_c$,

$$\mathbb{E}[h_t \mid F, \mathcal{F}_{t-1}] = N^{-1} \nabla D_K(\alpha_t). \tag{7}$$

Equation 6 applies the control component exactly, averaging it over the coordinate draw (Rao–Blackwellization), and spends the raw shot only on a symmetric two-coordinate residual. A poor control therefore costs variance, never bias.

A shot's noise comes from which partner is drawn and from the coin flip; the control keeps the coin flip but shrinks the first part to the residual $K - B_c$, which is small when the landmarks are predictive. Because $K_{IJ} = K_{JI}$, the same bit corrects both endpoints, nearly two coordinate corrections per non-diagonal shot.

Shared-shot ensemble. Step sizes are costly to tune when every trial consumes shots. Neither the uniform pair nor the law of Z_{IJ} depends on the iterate, so step-size schedules fixed in advance can share every outcome. Each of L trajectories keeps its own iterate, control vector v, and AdaGrad state; Eq. 7 holds for each conditional on the joint history, and a preset convex average stays dual-feasible. The pair schedule can even be drawn and measured before training, with no device round trips. Each round needs at most one shot whatever L, so optimization uses at most T shots plus the control construction; the price is classical work, $O(LNr)$ per round.

3.2 Aggregate-Margin-Oracle Pegasos

Pegasos, stochastic subgradient descent on Eq. 1, draws an index I_t uniformly and, with margin $m_t = y_{I_t} \langle w_t, \phi_{I_t} \rangle$, uses

$$w_{t+1} = \left(1 - \frac{1}{t}\right) w_t + \frac{a_t}{\lambda t} y_{I_t} \phi_{I_t}, \qquad a_t = \mathbf{1}\{m_t < 1\}. \tag{8}$$

AMO replaces the ideal indicator by its oracle-selected action $a_t \in \{0, 1\}$. For either action rule, expanding w_t over past rounds gives

$$m_t = \frac{1}{\lambda(t-1)} \sum_{s<t} a_s y_{I_t} y_{I_s} K_{I_t I_s}. \tag{9}$$

At $t = 1$, set $w_1 = 0$ and update without a shot; the history formulas apply for $t \geq 2$. Let $A_t = \sum_{s<t} a_s$ count the past update rounds (the active occurrences) and $b_t = A_t/[\lambda(t-1)]$. Sampling one past update round S uniformly yields

$$X_t = b_t y_{I_t} y_{I_S} Z_{I_t I_S}, \qquad \mathbb{E}[X_t \mid \mathcal{F}_{t-1}, I_t] = m_t, \quad |X_t| \leq b_t \leq \lambda^{-1}. \tag{10}$$

As in RSCA, a sum of up to $t-1$ kernel terms becomes the mean of one scaled bit; subsampled Metropolis–Hastings tests similarly replace a data sum by a subsampled mean in a sequential test (Bardenet et al., 2014). If $A_t = 0$ or $b_t < 1$, $m_t < 1$ exactly and AMO updates with no shot.

Centering and exact diagonal stripping. Remove the d active occurrences equal to I_t, whose kernel value is known, let $k = A_t - d$ and $b_k = k/[\lambda(t-1)]$; if $k = 0$ the margin is known exactly. Otherwise, for each sample q, draw an independent remaining occurrence J_q and one fresh outcome $Z_{I_t J_q}$. With $s_{J_q} = y_{I_t} y_{J_q}$ and $\bar{s}$ the exact mean of $y_{I_t} y_j$ over the k remaining occurrences, define

$$B_q = \tfrac{1}{2} + s_{J_q}(Z_{I_t J_q} - \tfrac{1}{2}) \in \{0, 1\}, \qquad X_q = \frac{d}{\lambda(t-1)} + b_k \left(\tfrac{1}{2}\bar{s} - \tfrac{1}{2} + B_q\right). \tag{11}$$

Then $\mathbb{E}[X_q \mid \mathcal{F}_{t-1}, I_t] = m_t$, with stochastic range $b_k \leq b_t$, at most half the uncentered $2b_t$; conditional on the history the B_q are i.i.d. Bernoulli, so an exact beta-mixture confidence sequence can be inspected after every outcome.

Strict AMO. The samples feed a confidence sequence $[L_n, U_n]$ (Howard et al., 2021; Waudby-Smith & Ramdas, 2024) that covers m_t at all sample counts n simultaneously (anytime-valid), so AMO may stop as soon as the answer is clear. We use the beta-mixture sequence; any time-uniform sequence for a bounded mean keeps the certificate. UPDATE has hinge slack at most τ once $U_n \leq 1 + \tau$ and NO-UPDATE once $L_n \geq 1 - \tau$; AMO stops at the first such n (taking the smaller slack if both hold):

$$U_n \leq 1 + \tau \quad \text{or} \quad L_n \geq 1 - \tau. \tag{12}$$

A margin far from one is settled by a few bits; one within Δ of one needs order Δ^{-2} samples at fixed λ, as for telling coins of bias $\frac{1}{2} \pm \Delta$ apart, and τ caps this cost. Splitting the risk ρ across oracle calls makes all intervals hold at once along the adaptive optimizer path.

Budgeted AMO. At a cap of R samples, unresolved calls act on the sample mean. The resulting expected subgradient slack has an additional $R^{-1/2}$ term (Section 4); early stops retain the confidence certificate. A margin sample costs one shot however long the history grows.

4 Guarantees and information limits

Expectations are over adaptive index and measurement randomness under the fixed-kernel assumptions of Section 2. Proofs appear in Appendix D.

Theorem 1 (RSCA convergence)**.** *Initialize $\alpha_1 = 0$, use Eq. 4, choose $\gamma = (B\sqrt{NT})^{-1}$, and output $\bar{\alpha}_T = T^{-1} \sum_t \alpha_t$. Then*

$$\mathbb{E}[D(\alpha^\star) - D(\bar{\alpha}_T)] \leq B\sqrt{N/T}, \tag{13}$$

$$D(\alpha^\star) - D(\bar{\alpha}_T) \leq B\sqrt{N/T} + 2B\sqrt{2\log(1/\rho)/T} \tag{14}$$

with probability at least $1 - \rho$. Moreover, for any feasible α, $D(\alpha^\star) - D(\alpha) \geq (\lambda/2)\|w(\alpha) - w^\star\|^2$. Hence, with probability $\geq 1 - \rho$, decision error ϵ needs only $O(B^2[N + \log(1/\rho)]/(\lambda^2\epsilon^4))$ shots.

The controlled and shared-shot directions, with fixed or AdaGrad steps, satisfy the same $\sqrt{N/T}$ bound in expectation, with B replaced by a control-dependent constant (Appendix D.3).

The factor N arises because each shot moves one of N coordinates. The rate is the $1/\sqrt{T}$ of projected stochastic ascent, because shot noise persists and the dual is not strongly concave when K is rank-deficient. Primal strong convexity turns a dual gap of $O(\lambda\epsilon^2)$ into decision error ϵ, hence $T \propto N\epsilon^{-4}$ shots.

Theorem 2 (Strict and budgeted AMO). *Initialize $w_1 = 0$, let $\tau > 0$, and set $\bar{w}_T = T^{-1}\sum_t w_t$. Fix per-call failure probabilities before sampling, summing to at most $\rho \in (0,1)$ along every path. Strict AMO satisfies*

$$\mathbb{E}[P(\bar{w}_T) - P(w^\star)] \le \frac{2(1+\log T)}{\lambda T} + \tau + \frac{(1+\lambda^{-1})\rho}{T}. \tag{15}$$

For centered budgeted AMO with cap R, the bound increases by at most $1/(2\lambda\sqrt{R})$.

The strict result follows because Eq. 12 returns a τ-subgradient when all intervals hold; otherwise either action is still a bounded approximate subgradient. Strong convexity transfers objective gap to decision error. For $0 < \lambda, \epsilon \le 1$, setting $\tau = \Theta(\lambda\epsilon^2)$ and $T = \widetilde{O}((\lambda\epsilon)^{-2})$ rounds, each costing at most $\widetilde{O}((\lambda\tau)^{-2})$ shots, gives

$$Q_{\mathrm{AMO}} = \widetilde{O}(\lambda^{-6}\epsilon^{-6}). \tag{16}$$

These rates bound root-mean-square uniform decision error; Markov's inequality gives a fixed success probability (Appendix D.7).

Margin adaptation. AMO pays heavily only for margins near one, so its total cost depends on how often the optimizer meets them. Let $\Delta_t = |1 - m_t|$ and condition on the history just before drawing the uniform training index at round t. Assume constants C_m, u_0, β, uniform over histories and rounds, with

$$\mathbb{P}(\Delta_t \le u \mid \mathcal{F}_{t-1}) \le C_m u^\beta \quad (0 < u \le u_0). \tag{17}$$

Larger β means fewer rounds with a margin near one; Appendix A discusses the assumption at finite N. When all intervals hold, a strict call uses $\widetilde{O}(1 + \lambda^{-2}\max\{\Delta_t, \tau\}^{-2})$ shots, and the worst-case cap controls the rest. Taking $\rho = \Theta(\tau^2)$ and $\tau = \Theta(\lambda\epsilon^2)$, integrating this cost against Eq. 17 yields, at fixed λ,

$$\mathbb{E}Q_{\mathrm{AMO}} = \begin{cases} \widetilde{O}(\epsilon^{-(6-2\beta)}), & 0 \le \beta < 2, \\ \widetilde{O}(\epsilon^{-2}\log(1/\epsilon)), & \beta = 2, \\ \widetilde{O}(\epsilon^{-2}), & \beta > 2. \end{cases} \tag{18}$$

Optimization still needs roughly ϵ^{-2} rounds; β controls the average measurement cost of their hinge decisions. Thus $\beta > 2$ makes that second cost bounded up to logarithms, while $\beta = 0$ recovers ϵ^{-6}; Figure 4 tests both the per-decision cost and the distribution of gaps. On a finite dataset the gaps are discrete: Eq. 18 describes the regime in which τ exceeds the smallest nonzero gap $\Delta_{\min}$, and below it the per-call cost stops growing, at $\widetilde{O}(\lambda^{-2}\Delta_{\min}^{-2})$ (Appendix D.6).

Two regimes and a local lower bound. With constant success probability, combining branches gives

$$Q = \widetilde{O}\big(\min\{N\epsilon^{-4}, \epsilon^{-6}\}\big) \quad \text{at fixed } \lambda. \tag{19}$$

RSCA is preferable at small N or high accuracy, and worst-case AMO removes the explicit N dependence (Table 1). With $T_\epsilon = (\lambda\epsilon)^{-2}$ Pegasos rounds and $\lambda \le 1$, the RSCA branch is $O(NT_\epsilon^2)$ and the AMO branch $\widetilde{O}(T_\epsilon^3)$: each spends T_ϵ^2 shots per dual coordinate or per primal round, so the two bounds cross at $N^\star = \widetilde{\Theta}(T_\epsilon)$, which grows as regularization weakens.

No sequential test avoids the per-decision Δ^{-2} cost: deciding between margins $m = 1 \pm \Delta$ with $\Delta > \tau$ is telling two coins apart. For interior Bernoulli instances and ρ bounded away from $1/2$, sequential change of measure shows that any procedure returning a τ-valid action with probability at least $1 - \rho$ on both instances needs $\mathbb{E}R = \Omega(\sigma^2\Delta^{-2}\log(1/\rho))$ samples, where σ^2 is the variance of one margin sample. AMO's variance-sensitive bound matches this dependence (Appendices D.8 and D.6). Optimizer-directed measurement changes what a shot is used for, not the information it carries. If kernel values concentrate as the qubit count q grows (Thanasilp et al., 2024), so that instances needing opposite decisions differ by only $e^{-\Omega(q)}$ information per query, any method in this access model needs $e^{\Omega(q)}$ shots.

Table 1: Shots to decision error ϵ, hiding logarithms and λ; prior rates are for their own objectives.

Method	Target	Training shots	Note
Full QKE dual (Gentinetta et al., 2024)	squared slack	$N^{4.67}\epsilon^{-2}$	noisy prediction entries
Noisy Pegasos (Gentinetta et al., 2024)	hinge	$\min\{N^2\epsilon^{-6}, \epsilon^{-10}\}$	empirical assumption
Nyström QKE (Coelho et al., 2025)	surrogate	rank dependent	changes the model
RSCA	exact hinge	$N\epsilon^{-4}$	unbiased, no pair cache
AMO, worst case	exact hinge	ϵ^{-6}	no explicit N
AMO, $\beta > 2$	exact hinge	ϵ^{-2}	margin law, $\tau \gg \Delta_{\min}$

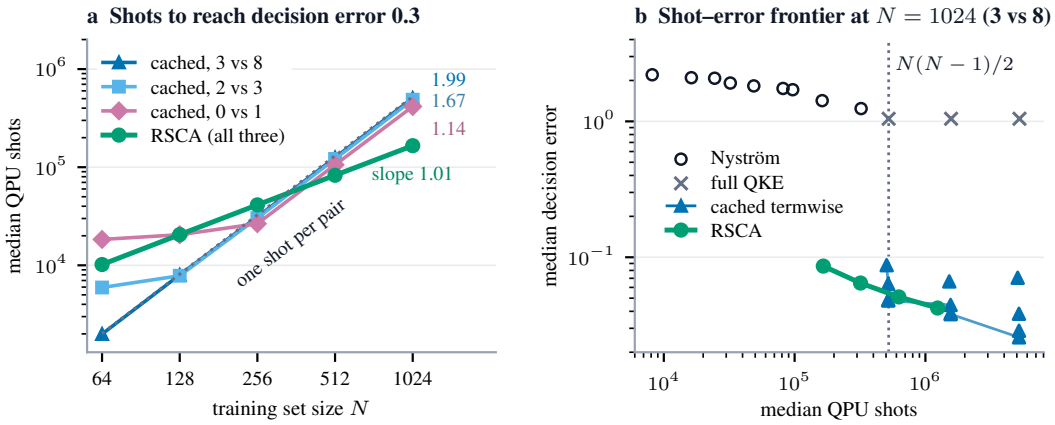

Figure 3: **Shot cost versus sample size on MNIST** ($\lambda = 0.05$, twenty evaluation seeds per cell). **(a)** Median shots to reach decision error 0.3; RSCA meets it at its shortest horizon on all three pairs (upper bounds), and the dotted line marks one shot per pair. **(b)** Median shot–error frontier at $N = 1024$ on 3 vs 8: RSCA reaches error 0.086 below the cheapest cached termwise configuration ($\approx N(N-1)/2$ shots); Nyström points are surrogates with a different target.

5 Experiments

We measure how closely each method recovers the exact-kernel SVM. In shot-level simulation the exact kernel is available only to the evaluator, and every training query returns a fresh Bernoulli outcome from Eq. 3. A device enters training only through such bits, so this is exact for any ideal circuit realizing K; noise and drift are studied below. Appendix E gives every configuration.

Linear rather than quadratic shot cost on MNIST. We train on MNIST digit pairs 3 vs 8, 2 vs 3, and 0 vs 1 with $N \in \{64, \dots, 1024\}$ balanced examples. Each image is reduced to four principal components and encoded by the four-qubit product map $K(x, z) = \prod_{k=1}^{4} \cos^2((x_k - z_k)/2)$; training never exploits its product structure, and its rank of up to $3^4 = 81$ is far above that of our four-landmark control, so the residual shots set the scaling. We use $\lambda = 0.05$. RSCA runs the shared-shot control variant with the step schedule of Appendix E. Its one constant, fixed once on development seeds of 3 vs 8, is reused unchanged for every pair and size. The strongest entry-estimating baseline we tested, cached termwise Pegasos, measures each kernel entry with a fixed number of shots the first time it is needed and reuses the estimate. Each budget-matched comparison takes the baseline's best cell on a grid of shots per pair and horizons, which favors the baseline.

Figure 3a shows the shots needed to reach median decision error 0.3. RSCA meets this target already at its shortest run length (horizon), about 160 shots per training example, for every pair and size. Its cost therefore grows linearly in N, and these counts are upper bounds that a finer grid could only lower. The cached baseline grows with slopes 1.14–1.99, and at $N = 1024$ RSCA needs 2.5–3.0$\times$ fewer shots (165.7k against 416k–506k).

The reason is pair coverage (Figure 3b). Pegasos touches almost every pair within five epochs, so even the cheapest cached configuration, one shot per pair for $5N$ rounds, costs 506k shots, close to

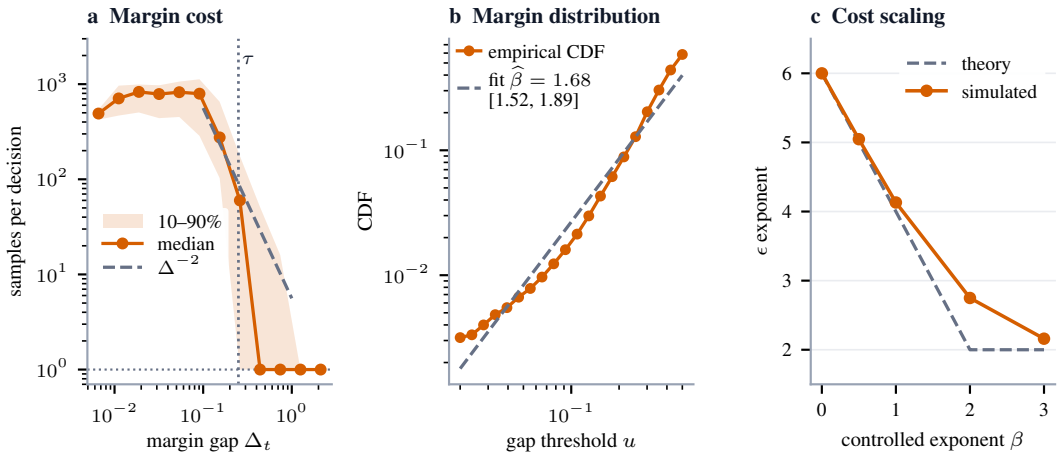

Figure 4: **AMO margin adaptation** (strict AMO, 20 trajectories at $\lambda = 0.5$). **(a)** Samples per decision against the exact gap: easy decisions take one sample, and cost grows roughly as Δ^{-2} toward the hinge (10th–90th percentile band; dashed slope -2; dotted lines mark $\tau = 0.25$ and the one-sample floor; the fitted slopes normalize samples by b_t^2). **(b)** Pooled near-margin CDF with power-law fit. **(c)** Exponents of ϵ in the summed per-call cost model under controlled margin laws against Eq. 18 (logarithms suppressed).

$N(N-1)/2$. RSCA reaches error 0.086 with a third of that budget, and with 63% of it beats that configuration in 16 of 20 paired trials ($p = 0.012$). RSCA's lead over the cached baseline grows with N, and the curves cross between $N = 256$ and 512. Demanding more accuracy or regularizing less shrinks the lead, consistent with RSCA's cost growing with accuracy and $1/\lambda$ (Theorem 1). At target 0.1 the cached baseline still leads at $N = 512$ on two pairs, and at $\lambda = 0.02$ it needs about 20% fewer shots at $N = 1024$ on two pairs while RSCA halves the cost on the third. Once the budget exceeds $N^2/2$ shots, cached entries are reused many times, the interfaces are not separable on 3 vs 8, and RSCA still leads on the other two pairs (Appendix E).

Full quantum kernel estimation (QKE: estimate every entry, then train) never reaches the target: its few-shot matrices are far from positive semidefinite, and projection onto the PSD cone pins every dual coordinate at its bound. Nyström surrogates train a different predictor (decision error 1.2–2.2). RSCA and cached termwise Pegasos both match the exact SVM's test accuracy (0.85 on 3 vs 8 at $N = 1024$) within 0.01: they train the same classifier and differ in what it costs.

Noise and drift do not reverse the ranking. We replace the ideal fidelity by the kernel of a noisy inversion test, in which each qubit is randomly scrambled (depolarized) with probability p and misread with probability p_{ro}. Every method keeps the unit diagonal, which adds a bias of at most 0.006 at $N = 1024$ and 0.083 at $N = 64$. At $p = p_{\mathrm{ro}} = 0.01$ the ranking is unchanged: RSCA wins 16–4 at $N = 256$ and matches the cached baseline's error with a third of its shots at $N = 1024$. Under strong noise ($p = 0.05$, $p_{\mathrm{ro}} = 0.03$), RSCA stays within 0.01 of the cached baseline at $N = 1024$ with a third of the shots, and the two interfaces reach parity at $N = 256$. At $N = 512$, a stress-test drift in p degrades long runs, which favors RSCA's short ones; a tenth of that drift gives parity. RSCA also needs nearly $4\times$ fewer distinct circuits, which matters when compiling and loading a new circuit costs more than repeating one (Appendix E).

Per-shot efficiency at matched budget. The synthetic study uses $N = 64$, $\lambda = 0.2$, balanced clusters of real one-qubit states with $K_{ij} = \cos^2(\theta_i - \theta_j)$, hyperparameters fixed on development trials, and 40 paired evaluation trials. At about 19.5k shots, shared-shot control RSCA reaches median error 0.0277 against 0.0356 for cached termwise Pegasos with ten shots per new pair. It has lower error in 27 of 40 paired trials (two-sided sign test $p = 0.038$) and uses no more shots in 39. With about ten shots per pair, cached entries are reused many times, so this gain reflects per-shot efficiency rather than pair coverage. Table 2 shows where it comes from: endpoint reuse without a control matches the cached baseline, and the shared-shot control supplies the gain. Below one pass over the pairs (about 2k shots), the aggregate-margin methods trace the frontier: budgeted AMO

Table 2: Synthetic family, 40 paired trials (medians and maxima taken separately). Symmetric RSCA reuses each bit at both endpoints without a control; control RSCA adds the frozen control without shot sharing.

Method	Median shots	Median error	Max. error
Shared-shot RSCA	19,443	0.0277	0.0748
Cached termwise (10)	19,880	0.0356	0.0947
Symmetric RSCA	19,666	0.0353	0.0929
Control RSCA	23,549	0.0266	0.0542
Cached termwise (12)	23,940	0.0258	0.0679
Full QKE (10/entry)	20,160	0.3149	0.3748

reaches 0.112 at 921 shots. At one pass, cached termwise Pegasos with one shot per pair reaches 0.051 at 1,980 shots (Appendix Figure 5).

AMO spends shots on hard margins. Strict AMO certifies every hinge decision and spends its shots where margins are close to one. Over 20 trajectories of 300 rounds at $N = 64$ it spends 16.9k shots per trajectory, and 10 of its 6,000 actions differ from the exact hinge action, all with hinge slack below τ. Its cost per decision falls roughly as Δ^{-2} above τ (pooled slope -2.08) and drops to a single sample for easy margins (Figure 4). Below gap $\tau/2$ it levels off, because the half-width stopping rule of the strict runs caps it (Appendix E). Centering and the beta-mixture sequence need $25\times$ fewer shots than Hoeffding bounds on the uncentered sample.

The overall rate also depends on how many margins are hard. The per-call cost model, summed over gaps from controlled margin laws with $\beta \in \{0, 0.5, 1, 2, 3\}$, reproduces the ordering of Eq. 18, and the visited gaps have fitted near-margin exponent $\widehat{\beta} = 1.68$ (95% CI $[1.52, 1.89]$), inside the adaptive range $0 < \beta < 2$ of Eq. 18 and far from its worst case $\beta = 0$, so margins close to one are rare along the optimizer's path. At the scales we test, RSCA is the method of choice; AMO suits budgets below one pass over the pairs and settings in which every decision must be certified.

6 Limitations

The guarantees cover a fixed normalized PSD kernel under conditionally unbiased measurement; device bias, drift, and error mitigation would add terms to them. Because the raw-shot interface keeps no memory, entry estimation becomes competitive again once the training horizon far exceeds N^2S for S shots per pair and cached entries amortize. The margin-adaptive rates rest on a history-conditional margin law. Our experiments simulate shots for one-qubit states and balanced MNIST pairs with a product feature map. Entangling maps, live hardware, imbalanced classes, inference cost (where the weighted-sum sampler of Gil-Fuster et al. (2026) applies), and head-to-head comparisons with adaptive shot allocation (Xu et al., 2026; Miroszewski, 2026) are natural next steps.

7 Conclusion

A quantum kernel outcome need not become a kernel estimate before it becomes useful. RSCA turns each outcome into an unbiased gradient step on the original SVM; AMO samples the complete primal margin, spends shots where hinge decisions are hard, and in its strict form certifies each decision at the cost the lower bound predicts. Theory prices both methods and identifies their complementary N–accuracy regimes. On MNIST, RSCA's shot cost grows at most linearly in N, so at large N it reaches the target error with fewer shots than measuring every pair once. At matched budget, its frozen control also makes each shot more informative, even where cached entries are reused many times (Table 2). Its pairs can be measured in one batch before training, with nearly $4\times$ fewer distinct circuits than the cached baseline. Measuring one weighted term per step is not specific to the SVM: any update that uses the kernel only through weighted sums of its entries, such as a gradient step on the kernel ridge regression system $(K + \lambda I)\alpha = y$, can run on raw shots. Whenever each query returns a destructive coin flip, measurements should estimate what the optimizer needs, not a matrix it never needed.

AI USE STATEMENT

In this work, we used generative AI tools for refining and debugging code for author-designed algorithms, preparing job-submission scripts, reformatting datasets, assisting in the writing of proofs, and translation, as well as for feedback on experimental methodology and on our interpretation of results. We have not used generative AI tools for developing the conceptual framework, designing experiments, proposing hypotheses, formulating the main theorems, providing the key ingredients of their proofs, or generating synthetic data; qualitative or thematic data analysis is not applicable to this work. Additionally, we used generative AI tools for writing plotting scripts for figures, polishing the writing of the paper, searching for and summarizing literature, and formatting references. We have reviewed all AI-assisted work: the code was tested with unit tests and an independent shot-count audit, all reported numbers were checked against the logged outputs, and the authors verified the proofs and references. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Appendix A states the assumptions and Appendix D gives complete proofs of all theoretical claims. Appendix E specifies every experiment: data generation and preprocessing, method configurations and grids, seeds, selection rules, and shot accounting; hyperparameters were fixed on development seeds disjoint from the evaluation seeds. An independent audit counts every shot, including landmark acquisition. The implementation, frozen protocol configurations, per-trial results, and the scripts that regenerate every table and figure will be released upon publication.

REFERENCES

Haim Avron, Michael Kapralov, Cameron Musco, Christopher Musco, Ameya Velingker, and Amir Zandieh. Random Fourier features for kernel ridge regression: Approximation bounds and statistical guarantees. In *International Conference on Machine Learning*, pp. 253–262, 2017.

Rémi Bardenet, Arnaud Doucet, and Chris Holmes. Towards scaling up Markov chain Monte Carlo: an adaptive subsampling approach. In *International Conference on Machine Learning*, pp. 405–413, 2014.

Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. In *Quantum Computation and Information*, volume 305 of *Contemporary Mathematics*, pp. 53–74. American Mathematical Society, 2002.

Rodrigo Coelho, Georg Kruse, and Andreas Rosskopf. Quantum-efficient kernel target alignment. In *Proceedings of the 17th International Conference on Agents and Artificial Intelligence (ICAART)*, pp. 763–772, 2025.

Bo Dai, Bo Xie, Niao He, Yingyu Liang, Anant Raj, Maria-Florina Balcan, and Le Song. Scalable kernel methods via doubly stochastic gradients. In *Advances in Neural Information Processing Systems*, volume 27, pp. 3041–3049, 2014.

Gian Gentinetta, Arne Thomsen, David Sutter, and Stefan Woerner. The complexity of quantum support vector machines. *Quantum*, 8:1225, 2024.

Elies Gil-Fuster, Seongwook Shin, Sofiene Jerbi, Jens Eisert, and Maximilian J. Kramer. Optimal algorithmic complexity of inference in quantum kernel methods. *arXiv preprint arXiv:2604.15214*, 2026.

Alex Gittens and Michael W. Mahoney. Revisiting the Nyström method for improved large-scale machine learning. *Journal of Machine Learning Research*, 17:1–65, 2016.

Jennifer R. Glick, Tanvi P. Gujarati, Antonio D. Córcoles, Youngseok Kim, Abhinav Kandala, Jay M. Gambetta, and Kristan Temme. Covariant quantum kernels for data with group structure. *Nature Physics*, 20(3):479–483, 2024.

Tobias Haug, Chris N. Self, and M. S. Kim. Quantum machine learning of large datasets using randomized measurements. *Machine Learning: Science and Technology*, 4(1):015005, 2023.

Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567:209–212, 2019.

Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *Annals of Statistics*, 49(2):1055–1080, 2021.

Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R. McClean. Power of data in quantum machine learning. *Nature Communications*, 12:2631, 2021.

Rie Johnson and Tong Zhang. Accelerating stochastic gradient descent using predictive variance reduction. In *Advances in Neural Information Processing Systems*, volume 26, pp. 315–323, 2013.

Jonas M. Kübler, Simon Buchholz, and Bernhard Schölkopf. The inductive bias of quantum kernels. In *Advances in Neural Information Processing Systems*, volume 34, pp. 12661–12673, 2021.

Artur Miroszewski. Adaptive measurement allocation for learning kernelized SVMs under noisy observations. *arXiv preprint arXiv:2605.22275*, 2026.

Ashley Montanaro. Quantum speedup of Monte Carlo methods. *Proceedings of the Royal Society A*, 471(2181):20150301, 2015.

Deanna Needell, Nathan Srebro, and Rachel Ward. Stochastic gradient descent, weighted sampling, and the randomized Kaczmarz algorithm. In *Advances in Neural Information Processing Systems*, volume 27, pp. 1017–1025, 2014.

Alessandro Rudi, Raffaello Camoriano, and Lorenzo Rosasco. Less is more: Nyström computational regularization. In *Advances in Neural Information Processing Systems*, volume 28, pp. 1657–1665, 2015.

Maria Schuld and Nathan Killoran. Quantum machine learning in feature Hilbert spaces. *Physical Review Letters*, 122(4):040504, 2019.

Shai Shalev-Shwartz and Tong Zhang. Stochastic dual coordinate ascent methods for regularized loss minimization. *Journal of Machine Learning Research*, 14:567–599, 2013.

Shai Shalev-Shwartz, Yoram Singer, Nathan Srebro, and Andrew Cotter. Pegasos: Primal estimated sub-gradient solver for SVM. *Mathematical Programming*, 127:3–30, 2011.

Abhay Shastry, Abhijith Jayakumar, Apoorva Patel, and Chiranjib Bhattacharyya. Shot-frugal and robust quantum kernel classifiers. *arXiv preprint arXiv:2210.06971*, 2023.

Ryan Sweke, Frederik Wilde, Johannes Jakob Meyer, Maria Schuld, Paul K. Faehrmann, Barthélémy Meynard-Piganeau, and Jens Eisert. Stochastic gradient descent for hybrid quantum-classical optimization. *Quantum*, 4:314, 2020.

Supanut Thanasilp, Samson Wang, M. Cerezo, and Zoë Holmes. Exponential concentration in quantum kernel methods. *Nature Communications*, 15:5200, 2024.

Kilian Tscharke and Pascal Debus. A multiclass quantum aligned centroid kernel. *arXiv preprint arXiv:2607.19782*, 2026.

Kilian Tscharke, Sebastian Issel, and Pascal Debus. QUACK: Quantum aligned centroid kernel. In *IEEE International Conference on Quantum Computing and Engineering (QCE)*, pp. 1425–1435, 2024.

Ian Waudby-Smith and Aaditya Ramdas. Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(1):1–27, 2024.

Christopher K. I. Williams and Matthias Seeger. Using the Nyström method to speed up kernel machines. In *Advances in Neural Information Processing Systems*, volume 13, pp. 682–688, 2001.

Jian Xu, Chao Li, Delu Zeng, John Paisley, and Qibin Zhao. AQKA: Active quantum kernel acquisition under a shot budget. *arXiv preprint arXiv:2605.14672*, 2026.

A Assumptions, notation, and access models

Let $Y = \mathrm{diag}(y)$, $\mathcal{A} = [0, 1/N]^N$, and let $\mathcal{F}_{t-1}$ contain all optimizer, index, and oracle history before round t's fresh pair/outcome. The main results assume: (A1) K is fixed, symmetric, PSD, and normalized; (A2) $\mathbb{E}[Z_t \mid \mathcal{F}_{t-1}, I_t, J_t] = K_{I_t J_t}$; (A3) samples within a margin call are conditionally independent, while across calls a martingale condition suffices; and (A4) classical bookkeeping is exact. A SWAP-test bit has mean $(1 + K)/2$ and must be affinely transformed, changing constants; the paper uses inversion-test semantics throughout. The norm bound $K(x, x) \leq 1$ suffices for the basic rate proofs. The zero-shot diagonal rule and AMO's exact diagonal stripping also use the fidelity identity $K_{ii} = 1$; bounded diagonals alone do not justify either shortcut.

For the conditional margin law, history is fixed before the uniform training index is drawn. Each of the finite N examples then has mass $1/N$. If $\beta > 0$, any gap $d \in (0, u_0]$ must satisfy $1/N \leq C_m d^\beta$, hence $d \geq (NC_m)^{-1/\beta}$; exact hinge ties are excluded, since a tie would contribute $1/N$ for every $u > 0$ while $C_m u^\beta \to 0$. Uniform C_m, u_0, β along histories and rounds, and along the precision sequence for an asymptotic rate, is thus a substantive condition; it differs from a test-distribution Tsybakov condition, and a pooled empirical CDF describes the visited gaps without certifying it.

Raw shots, coherent oracle queries, gate complexity, and classical work are different resources; amplitude estimation and classical shadows assume stronger access (Appendix C), and Eq. 19 prices raw shots only.

Table 3: Interfaces differ in objective, intermediate object, and access (QUACK and McQuack: Tscharke et al., 2024; Tscharke & Debus, 2026).

Family	Fixed SVM?	Intermediate	Access
Full/adaptive QKE	yes/robust	entries	shots
Nyström, fixed features	surrogate	features	entries/classical
QUACK/McQuack	no	centroid fidelities	shots
Classical shadows	conditional	shadow record	randomized meas.
Coherent weighted sum	conditional	aggregate estimate	coherent
RSCA / AMO	yes	optional control	raw shots

B Anytime confidence construction

Fix the history at the start of one call. Its centered bits are conditionally i.i.d. $B_1, B_2, \ldots \sim \mathrm{Bernoulli}(p)$; Appendix D.4 derives this representation from the sampled training history. Write $S_n = \sum_{q=1}^n B_q$. For $a, b > 0$, the beta-mixture likelihood ratio is

$$M_n(p) = \int_0^1 \frac{\theta^{S_n}(1-\theta)^{n-S_n}}{p^{S_n}(1-p)^{n-S_n}} \frac{\theta^{a-1}(1-\theta)^{b-1}}{\mathrm{B}(a,b)}\, d\theta = \frac{\mathrm{B}(S_n + a, n - S_n + b)}{\mathrm{B}(a,b)p^{S_n}(1-p)^{n-S_n}}. \tag{20}$$

Here $M_0 = 1$ and endpoint ratios use their one-sided limits. For call risk $0 < \delta < 1$, let C_n be the closed hull of $\{p \in [0,1] : M_n(p) < 1/\delta\}$. It is an interval: on $(0,1)$, $\log M_n(p)$ is a constant minus $S_n \log p + (n - S_n)\log(1-p)$, a convex function of p. It is nonempty because the likelihood is maximized at $\widehat{p}_n = S_n/n$, so $M_n(\widehat{p}_n) \leq 1$ (with the endpoint convention).

Proposition 3 (Conditional anytime coverage)**.** *For every fixed call history,*

$$\mathbb{P}\big(p \in C_n \text{ for all } n \geq 1 \mid \text{call history}\big) \geq 1 - \delta.$$

If round t chooses its risk δ_t before observing its fresh outcomes, and $\sum_{t=1}^T \delta_t \leq \rho$ on every path, then the intervals cover all queried margins at all inspected sample counts with probability at least $1 - \rho$.

Proof. For $0 < p < 1$, the likelihood ratio inside the integral in Eq. 20 has one-step conditional multiplier with mean

$$p\frac{\theta}{p} + (1-p)\frac{1-\theta}{1-p} = 1.$$

Nonnegativity allows conditional expectation to pass through the integral, so $M_n(p)$ is a nonnegative martingale. Under $p = 0$, all observations are zero almost surely and

$$M_n(0) = \frac{\mathrm{B}(a, n+b)}{\mathrm{B}(a,b)}, \qquad \frac{M_{n+1}(0)}{M_n(0)} = \frac{n+b}{n+a+b} \leq 1.$$

Under $p = 1$, the corresponding ratio is $(n+a)/(n+a+b) \leq 1$. Thus at either endpoint the process is a nonnegative supermartingale on the null's support; an incompatible observation excludes it.

Let $\nu = \inf\{n : M_n(p) \geq 1/\delta\}$. For any finite m, the stopped supermartingale has $\mathbb{E}M_{\nu\wedge m}(p) \leq 1$, by iterating the conditional supermartingale inequality. On $\{\nu \leq m\}$ its value is at least $1/\delta$, hence $\mathbb{P}(\nu \leq m) \leq \delta$. Letting $m \to \infty$ gives $\mathbb{P}(\nu < \infty) \leq \delta$, which proves the first assertion.

For an affine margin $m = c + wp$, $w \geq 0$, transform C_n into $[L_n, U_n] = c + wC_n$. Conditional on each call history the chance of any coverage failure is at most its selected δ_t. The tower property and union bound give

$$\mathbb{P}\left(\bigcup_{t=1}^{T} G_t^c\right) \leq \sum_{t=1}^{T} \mathbb{E}[\mathbb{P}(G_t^c \mid \text{call history})] \leq \mathbb{E}\sum_{t=1}^{T} \delta_t \leq \rho,$$

where G_t denotes coverage at every inspection in call t. This argument permits adaptive training histories and stopping times. □

C Expanded related-work comparison

Quantum-kernel learning depends on the feature map's inductive bias and on the data available to classical learners (Kübler et al., 2021; Huang et al., 2021; Glick et al., 2024). Kernel concentration can make polynomial-shot estimates uninformative (Thanasilp et al., 2024). The comparisons below separate measurement allocation, predictor approximation, and the optimization primitive.

Entry uncertainty versus raw-shot optimization. ShofaR protects an optimization program against a confidence region of noisy kernel entries (Shastry et al., 2023), and adaptive schemes allocate shots to entries judged task-sensitive (Xu et al., 2026; Miroszewski, 2026). Both retain a numerical-entry layer. RSCA instead decides which stochastic direction to draw and never requires the sampled entry to become accurate. Robust or adaptively allocated estimates remain attractive when they can be reused across multiple tasks; raw-shot optimization targets a single fixed-kernel training objective under a destructive-shot budget. Gentinetta et al. (2024, p. 4, fn. 3) use squared slack for their dual rather than the exact dual of their hinge primal. Their dual decision error also includes noisy prediction entries and relies on a noisy-halfspace assumption (their Assumption 1; pp. 9, 24–26), and their noisy Pegasos bound assumes stability to sufficiently accurate margin sums. Table 1 therefore lists each rate with its own objective.

Surrogates versus controls. Nyström features, fixed finite random-feature surrogates, and centroid fidelities can offer strong systems trade-offs (Gittens & Mahoney, 2016; Rudi et al., 2015; Avron et al., 2017; Coelho et al., 2025; Tscharke et al., 2024), but their approximation error is distinct from shot noise and they generally optimize a different predictor class. Our Nyström object plays a control-variate role closer to preconditioning: a poor frozen control increases residual variance but, because fresh shots correct $K - B_c$, does not change the limiting SVM. Doubly stochastic gradients (Dai et al., 2014) are an important exception to a blanket random-feature/surrogate classification: fresh random features drawn at every step give unbiased functional gradients that converge to the original RKHS optimum. RSCA's weight-proportional partner sampling is a form of the importance sampling studied for SGD by Needell et al. (2014).

Weighted sums before AMO. Gil-Fuster et al. (2026, App. C), in their catalogue of inference strategies, sample an index with probability $|c_i|/\|c\|_1$ and use one destructive fidelity outcome for its signed contribution, giving sample-and-average inference at cost $O(\|c\|_1^2/\epsilon^2)$ for fixed coefficients. AMO turns this primitive into a training oracle: the coefficients change at every round, the estimate is converted into a sequential hinge decision with an anytime-valid certificate, and the resulting shots are charged inside an optimization guarantee with explicit subgradient slack.

Other measurement interfaces. Classical shadows (Huang et al., 2020) can reuse randomized measurement records, for a fixed set of observables with controlled shadow norm; adaptively chosen requests need fresh records or a union bound. Coherent amplitude estimation (Brassard et al., 2002; Montanaro, 2015) can alter precision dependence if one can prepare the weighted active-history state and call the feature-map oracle and its inverse coherently. Their query and state-preparation costs are measured in the corresponding access models.

D PROOFS

Throughout, $N, T \geq 1$, $\lambda > 0$, and all vector norms are Euclidean or RKHS norms as appropriate. For a predictor w, write

$$e(w) = \sup_{x:K(x,x)\leq 1} |\langle w - w^\star, \phi(x)\rangle|.$$

The history before a round includes every previous outcome and iterate. All sampling expectations below condition on this history before drawing the current training coordinate or pair.

D.1 DUALITY AND CONVERSION TO DECISION ERROR

Lemma 4 (Primal–dual identities)**.** *The hinge objective has the dual in Eq. 2. A dual maximizer $\alpha^\star$ exists and $w^\star = w(\alpha^\star)$ is the unique primal minimizer. For every feasible α,*

$$e(w(\alpha))^2 \leq \|w(\alpha) - w^\star\|^2 \leq \frac{2}{\lambda}[D(\alpha^\star) - D(\alpha)]. \tag{21}$$

For every w, $e(w)^2 \leq 2[P(w) - P(w^\star)]/\lambda$.

Proof. The component of w orthogonal to $\operatorname{span}\{\phi_i\}_{i=1}^N$ changes no training prediction and only increases the regularizer. We may therefore work in this finite-dimensional span. Introduce slacks $\xi_i \geq 0$ and constraints $1 - y_i\langle w, \phi_i\rangle \leq \xi_i$. With nonnegative multipliers α_i, μ_i, the Lagrangian is

$$\mathcal{L} = \frac{\lambda}{2}\|w\|^2 + \frac{1}{N}\sum_i \xi_i + \sum_i \alpha_i(1 - y_i\langle w, \phi_i\rangle - \xi_i) - \sum_i \mu_i\xi_i.$$

Minimization in the now unconstrained slack variables requires $\alpha_i + \mu_i = 1/N$, or $0 \leq \alpha_i \leq 1/N$. Minimization in w gives $w = \lambda^{-1}\sum_i \alpha_i y_i \phi_i$. Substitution yields Eq. 2 and weak duality. Compactness of $\mathcal{A}$ gives a maximizer $\alpha^\star$. Define $w^\star = w(\alpha^\star)$ and $g_i = 1 - y_i\langle w^\star, \phi_i\rangle$. Coordinate optimality implies $g_i \leq 0$ at $\alpha_i^\star = 0$, $g_i \geq 0$ at $\alpha_i^\star = 1/N$, and $g_i = 0$ at interior coordinates. In every case $\alpha_i^\star g_i = (g_i)_+/N$. Therefore

$$P(w^\star) = \frac{\lambda}{2}\|w^\star\|^2 + \sum_i \alpha_i^\star g_i = \sum_i \alpha_i^\star - \frac{\lambda}{2}\|w^\star\|^2 = D(\alpha^\star).$$

Weak duality proves primal optimality and equality of the optima. The primal objective is λ-strongly convex, hence this minimizer is unique.

Put $\Delta = \alpha - \alpha^\star$. Optimality over the convex box gives $\nabla D(\alpha^\star)^\top \Delta \leq 0$, by differentiating along the segment from $\alpha^\star$ to α. Since D is quadratic,

$$\begin{aligned} D(\alpha^\star) - D(\alpha) &= -\nabla D(\alpha^\star)^\top \Delta + \frac{1}{2\lambda}\Delta^\top Y K Y \Delta \\ &\geq \frac{1}{2\lambda}\Big\|\sum_i \Delta_i y_i \phi_i\Big\|^2 = \frac{\lambda}{2}\|w(\alpha) - w^\star\|^2. \end{aligned}$$

Cauchy–Schwarz and $\|\phi(x)\| \leq 1$ imply the first inequality in Eq. 21. Finally, strong convexity at the unconstrained minimizer gives $P(w) - P(w^\star) \geq (\lambda/2)\|w - w^\star\|^2$. □

Regularized intercept. For an appended constant feature, let $\psi(x) = (\phi(x), \kappa_b)$ and $s = \sqrt{1 + \kappa_b^2}$. Set $\widetilde{\phi} = \psi/s$ and $\widetilde{w} = su$. Then $\langle \widetilde{w}, \widetilde{\phi}\rangle = \langle u, \psi\rangle$ and $(\lambda/s^2)\|\widetilde{w}\|^2/2 = \lambda\|u\|^2/2$. This proves equivalence of the objectives under $\widetilde{K} = (K + \kappa_b^2)/s^2$ and $\widetilde{\lambda} = \lambda/s^2$. Outputting one with probability κ_b^2/s^2, and otherwise using an original Bernoulli shot, has mean $\widetilde{K}$. Thus the normalized oracle and all preceding identities apply to this regularized intercept.

D.2 ONE-SHOT RSCA: EXPECTATION AND HIGH PROBABILITY

Proof of Theorem 1. Fix a feasible α and coordinate i, and put $A = \sum_j \alpha_j \le 1$. For $A > 0$, averaging the sampled partner and its outcome gives

$$\mathbb{E}[\widehat{g}_i \mid \alpha, i] = 1 - \frac{A}{\lambda} y_i \sum_j \frac{\alpha_j}{A} y_j K_{ij} = \nabla_i D(\alpha).$$

If $A = 0$, feasibility implies $\alpha = 0$, and the returned value one is exactly the gradient. In either case $|\widehat{g}_i| \le B = 1 + \lambda^{-1}$. Let H_t denote the pre-round history, and define

$$R_t = \|\alpha_t - \alpha^\star\|^2, \qquad X_t = \widehat{g}_{I_t}(\alpha^\star_{I_t} - \alpha_{t,I_t}).$$

Coordinate projection is nonexpansive and fixes the feasible comparator, so

$$2\gamma X_t \le R_t - R_{t+1} + \gamma^2 B^2. \tag{22}$$

Uniform coordinate sampling and concavity of D imply

$$D(\alpha^\star) - D(\alpha_t) \le \langle \nabla D(\alpha_t), \alpha^\star - \alpha_t \rangle = N\mathbb{E}[X_t \mid H_t]. \tag{23}$$

Taking expectations in Eq. 22, summing, and using $R_1 = \|\alpha^\star\|^2 \le 1/N$ and $R_{T+1} \ge 0$ yields

$$\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[D^\star - D(\alpha_t)] \le \frac{1}{2\gamma T} + \frac{N\gamma B^2}{2}.$$

Concavity gives $D(\bar{\alpha}_T) \ge T^{-1}\sum_t D(\alpha_t)$. Substituting $\gamma = (B\sqrt{NT})^{-1}$ proves Eq. 13.

For the high-probability statement define $M_t = \mathbb{E}[X_t \mid H_t] - X_t$. Since $|\alpha^\star_i - \alpha_{t,i}| \le 1/N$, we have $|X_t| \le B/N$ and $|M_t| \le 2B/N$. These are martingale differences. For a conditionally centered variable $Y \in [-c, c]$, convexity bounds $e^{\theta Y}$ by the chord between its endpoint values, so $\mathbb{E}[e^{\theta Y} \mid H] \le \cosh(\theta c) \le e^{\theta^2 c^2/2}$. The last inequality follows termwise from $(2k)! \ge 2^k k!$ in the power series. Taking $c = 2B/N$ and iterating over rounds gives

$$\mathbb{E}\exp\left(\theta \sum_{t=1}^{T} M_t\right) \le \exp\left(\frac{2\theta^2 B^2 T}{N^2}\right).$$

Markov's inequality and minimization over $\theta > 0$ therefore imply

$$\mathbb{P}\left\{\sum_{t=1}^{T} M_t > \frac{2B}{N}\sqrt{2T\log(1/\rho)}\right\} \le \rho.$$

Before taking expectations, summing Eqs. 22 and 23 gives

$$D^\star - D(\bar{\alpha}_T) \le \frac{1}{2\gamma T} + \frac{N\gamma B^2}{2} + \frac{N}{T}\sum_{t=1}^{T} M_t.$$

This is Eq. 14. Lemma 4 now shows that $e(w(\bar{\alpha}_T)) \le \epsilon$ with probability at least $1 - \rho$ whenever

$$T \ge \frac{4B^2}{\lambda^2 \epsilon^4}\left(\sqrt{N} + 2\sqrt{2\log(1/\rho)}\right)^2.$$

Each round requests at most one physical shot, which proves the stated cost. □

Sampling and averaging work. A Fenwick tree stores nonnegative masses α_j and their binary-block sums. A uniform draw in $[0, A)$ selects the unique half-open cumulative interval of length α_j, hence partner j has probability α_j/A. Binary lifting and a point update each visit $O(\log N)$ tree nodes. For the average, store each coordinate's accumulated sum and the first unaccounted time ℓ_i. Before changing that coordinate at the end of round t, add $(t - \ell_i + 1)\alpha_i$ and set $\ell_i = t + 1$. A final pass adds each remaining constant segment through round T. These disjoint segments sum exactly to $\sum_{t=1}^{T} \alpha_t$, with constant averaging work per changed coordinate and one $O(N)$ final pass.

D.3 CONTROLLED DIRECTIONS, ADAPTIVE STEPS, AND SHARED OUTCOMES

Condition on the landmark observations so that F and $B_c = FF^\top$ are fixed. Write $L_F = \max_i \|F_i\|$ and let H_t include every replica's history. For independent uniform I, J and their shared fresh outcome,

$$\mathbb{E}[\widehat{r}_I \mid H_t, I = i] = -\frac{y_i}{\lambda} \sum_j \alpha_j y_j (K_{ij} - B_{c,ij}) = \nabla_i D_K(\alpha) - g_i^{B_c}.$$

The analogous identity holds for the other endpoint because K and B_c are symmetric. Multiplying by coordinate indicators and averaging gives

$$\mathbb{E}\left[\tfrac{1}{2}(e_I \widehat{r}_I + e_J \widehat{r}_J) \mid H_t\right] = \frac{1}{N}(\nabla D_K - g^{B_c}).$$

Adding the dense control proves Eq. 7. When $I = J$, the two half-contributions combine into one coordinate contribution; the argument does not assume independence of the two endpoints.

The following variance identity describes the role of the control. For fixed i, let $V_i = N\alpha_J y_J (Z_{iJ} - B_{c,iJ})$. Conditional variance decomposition gives

$$\mathrm{Var}(V_i \mid H_t) = N \sum_j \alpha_j^2 K_{ij}(1 - K_{ij}) + \mathrm{Var}_J(N\alpha_J y_J (K_{ij} - B_{c,ij}))\,, \quad (24)$$

$$\mathrm{Var}_J(N\alpha_J y_J (K_{ij} - B_{c,ij})) = N \sum_j \alpha_j^2 (K_{ij} - B_{c,ij})^2 - \left(\sum_j \alpha_j y_j (K_{ij} - B_{c,ij}) \right)^2 .$$

Indeed, conditional on $J = j$, the variance of V_i is $N^2 \alpha_j^2 K_{ij}(1 - K_{ij})$ and its mean is $N\alpha_j y_j (K_{ij} - B_{c,ij})$. This proves both lines. An exact control removes the between-partner term; an entrywise residual bound $|K_{ij} - B_{c,ij}| \le \zeta$ bounds that term by $N\zeta^2 \sum_j \alpha_j^2$. Applying g^{B_c}/N densely also removes the control component's coordinate-sampling noise: it equals the conditional expectation of $(e_I g_I^{B_c} + e_J g_J^{B_c})/2$. The full direction thus has the conditional variance of its sparse residual.

A uniform direction bound and projected ascent. The inequalities $\sum_j \alpha_j \le 1$, $N\alpha_j \le 1$ and $|F_i^\top F_j| \le L_F^2$ imply

$$|g_i^{B_c}| \le 1 + L_F^2/\lambda, \quad |\widehat{r}_I|, |\widehat{r}_J| \le (1 + L_F^2)/\lambda, \quad \|h_t\| \le H_F := \frac{1 + L_F^2/\lambda}{\sqrt{N}} + \frac{1 + L_F^2}{\lambda}.$$

For $\alpha_{t+1} = \Pi_{\mathcal{A}}(\alpha_t + \eta h_t)$, projection and $\mathrm{diam}(\mathcal{A})^2 = 1/N$ give

$$\sum_{t=1}^{T} \langle h_t, \alpha^\star - \alpha_t \rangle \le \frac{1}{2N\eta} + \frac{\eta}{2} \sum_{t=1}^{T} \|h_t\|^2.$$

Conditional unbiasedness and concavity then yield

$$\mathbb{E}[D^\star - D(\bar{\alpha}_T) \mid F] \le \frac{1}{2\eta T} + \frac{N\eta H_F^2}{2}. \quad (25)$$

In particular, $\eta = (H_F \sqrt{NT})^{-1}$ gives $H_F \sqrt{N/T}$. Taking expectations before bounding $\|h_t\|$ replaces H_F^2 by $T^{-1} \sum_t \mathbb{E}\|h_t\|^2$, and $\mathbb{E}\|h_t\|^2 = N^{-2}\mathbb{E}\|\nabla D_K(\alpha_t)\|^2 + \mathbb{E}\nu_t$ with ν_t the conditional variance of the sparse residual, which Eq. 24 and the spread of the residual mean over coordinates bound; the control enters only through ν_t. These bounds are conditional on the control; unconditional bounds follow by averaging whenever the displayed moments are integrable.

Observed-gradient AdaGrad and suffix averages. Let $a_0, v_0 > 0$ be fixed, put $v_{t,i} = (v_0^2 + \sum_{u=1}^{t} h_{u,i}^2)^{1/2}$, and update each coordinate by

$$\alpha_{t+1,i} = \mathrm{clip}_{[0,1/N]} \left(\alpha_{t,i} + \frac{a_0 h_{t,i}}{v_{t,i}} \right).$$

For any fixed suffix $s, \dots, T$ of length M, scalar projection gives

$$\sum_{t=s}^{T} h_{t,i}(\alpha_i^\star - \alpha_{t,i}) \le \frac{1}{2a_0}\sum_{t=s}^{T} v_{t,i}(d_{t,i}^2 - d_{t+1,i}^2) + \frac{a_0}{2}\sum_{t=s}^{T}\frac{h_{t,i}^2}{v_{t,i}}, \qquad d_{t,i} = \alpha_{t,i} - \alpha_i^\star.$$

Since $v_{t,i}$ is nondecreasing and $d_{t,i}^2 \le 1/N^2$, summation by parts bounds the first sum by $v_{T,i}/N^2$. Also

$$\frac{h_{t,i}^2}{v_{t,i}} = (v_{t,i} - v_{t-1,i})\left(1 + \frac{v_{t-1,i}}{v_{t,i}}\right) \le 2(v_{t,i} - v_{t-1,i}).$$

Summing coordinates and taking conditional expectations proves

$$\mathbb{E}[D^\star - D(\bar\alpha_{s:T}) \mid F] \le \frac{N}{M}\sum_i \mathbb{E}\left[\frac{v_{T,i}}{2a_0N^2} + a_0(v_{T,i} - v_{s-1,i})\,\middle|\, F\right]. \tag{26}$$

The projection argument is pathwise, so it remains valid even though the step denominator uses the current outcome. With $\|h_t\| \le H_F$, $\sum_i v_{T,i} \le Nv_0 + H_F\sqrt{NT}$ by Cauchy–Schwarz. Thus any suffix of length proportional to T has an inverse-square-root bound for fixed a_0, N, F. Setting a_0 proportional to $1/N$ recovers the $H_F\sqrt{N/T}$ leading order.

Fixed mixtures and physical cost. For feasible outputs $\bar\alpha^{(\ell)}$ and fixed nonnegative weights ω_ℓ summing to one, $\widetilde\alpha = \sum_\ell \omega_\ell \bar\alpha^{(\ell)}$ belongs to $\mathcal{A}$ and

$$D^\star - D(\widetilde\alpha) \le \sum_\ell \omega_\ell[D^\star - D(\bar\alpha^{(\ell)})], \qquad e(w(\widetilde\alpha)) \le \sum_\ell \omega_\ell e(w(\bar\alpha^{(\ell)})).$$

The first inequality is concavity and the second is linearity of $w(\alpha)$ and the triangle inequality. They cover both replica and suffix mixtures. For square-integrable replica outputs V_ℓ, convexity also gives $\mathbb{E}\|\sum_\ell \omega_\ell(V_\ell - \mathbb{E}V_\ell)\|^2 \le \sum_\ell \omega_\ell \mathbb{E}\|V_\ell - \mathbb{E}V_\ell\|^2$, without an independence assumption. Conditioned on the joint history, each replica receives the same pair law and a Bernoulli outcome of mean K_{IJ}, so its conditional-mean identity holds separately. Correlation across replicas does not change these bounds. Imagine a fresh outcome also on rounds where it is unused. Skipping it when all residual multipliers vanish, or replacing a known diagonal outcome by one, leaves every update unchanged. If χ_t indicates that any replica requires a physical shot, then $Q = Q_{\text{landmarks}} + \sum_{t=1}^{T}\chi_t \le Q_{\text{landmarks}} + T$. For complexity counts take $r \ge 1$, padding a zero control with a zero column if necessary. Storing F and L iterates and aggregates takes $O(LNr)$ space as an upper bound; computing each dense control and updating its aggregate takes $O(LNr)$ work per round. A change δ in coordinate i updates the aggregate by $v \leftarrow v + \delta y_i F_i$, directly from its definition.

D.4 The aggregate sample and approximate hinge action

Allow any actions $a_t \in \{0, 1\}$ in Eq. 8, including the oracle's approximate decisions. Multiplying the update by λt and induction from $w_1 = 0$ give, for $t \ge 2$,

$$w_t = \frac{1}{\lambda(t-1)}\sum_{s<t} a_s y_{I_s}\phi_{I_s}, \qquad \|w_t\| \le \frac{A_t}{\lambda(t-1)} \le \lambda^{-1}. \tag{27}$$

Fix this history and the chosen example $I_t = i$. If $A_t > 0$, uniform sampling of an active occurrence and the fresh-shot mean imply

$$\mathbb{E}X_t = \frac{A_t}{\lambda(t-1)}\frac{1}{A_t}\sum_{s<t:a_s=1} y_i y_{I_s} K_{iI_s} = m_t.$$

If $A_t = 0$, $m_t = 0$ exactly. In general $|m_t| \le b_t$, which justifies the zero-shot update when $b_t < 1$.

For diagonal stripping, let d be the number of active occurrences of i and $k = A_t - d$. The known diagonal contributes $d/[\lambda(t-1)]$. If $k = 0$, the entire margin is known. Otherwise put $w = b_k = k/[\lambda(t-1)]$ and let $s_J = y_i y_J$ for a uniformly sampled remaining occurrence. Since $Z \in \{0, 1\}$, $B_q = 1/2 + s_J(Z - 1/2)$ equals Z for $s_J = 1$ and $1 - Z$ for $s_J = -1$. It is therefore Bernoulli, and

$$\mathbb{E}B_q = \frac{1}{2} + \frac{1}{k}\sum_{J \text{ remaining}} s_J(K_{iJ} - 1/2).$$

Consequently $X_q = c + wB_q$, where $c = d/[\lambda(t-1)] + w(\bar{s}/2 - 1/2)$, has mean m_t and support width w. Fresh occurrence sampling and fresh outcomes give the same conditional Bernoulli law at each draw. Iterating these conditional probabilities proves conditional i.i.d. sampling within the frozen call. In particular,

$$\mathrm{Var}(X_q \mid \text{call history}) = w^2 p(1-p) \le w^2/4 \le 1/(4\lambda^2). \tag{28}$$

History sampling work. An array of active occurrences supports an unstripped uniform draw and an append in $O(1)$ time. For diagonal-stripped sampling, store occurrence counts in a Fenwick tree. Let c_i be the chosen example's count and C_i the cumulative count before i. For $k = A_t - c_i > 0$, draw an integer u uniformly from $\{0, \dots, k-1\}$ and select cumulative rank u if $u < C_i$, or $u + c_i$ otherwise. This skips precisely the diagonal block; all remaining occurrences have probability $1/k$. Prefix lookup, rank selection, and count updates take $O(\log N)$ time and $O(N)$ tree storage.

For $\ell_i(w) = (1 - y_i\langle w, \phi_i\rangle)_+$, define the action slack

$$s(a, m) = \begin{cases} (m-1)_+, & a = 1, \\ (1-m)_+, & a = 0. \end{cases}$$

For every comparator v,

$$\ell_i(v) \ge \ell_i(w) + \langle -a y_i \phi_i, v - w\rangle - s(a, m), \qquad m = y_i\langle w, \phi_i\rangle. \tag{29}$$

For $a = 0$ this follows from $\ell_i(v) \ge 0$; for $a = 1$ its right side equals $1 - y_i\langle v, \phi_i\rangle$, which is bounded by $\ell_i(v)$. If $m \in [L_n, U_n]$, the two action slacks are bounded respectively by $(U_n - 1)_+$ and $(1 - L_n)_+$. Thus Eq. 12 selects an action with slack at most τ. On any path, even without coverage, $s(a, m) \le 1 + |m| \le C_\lambda := 1 + \lambda^{-1}$.

D.5 AMO OPTIMIZATION AND FIXED-BUDGET DECISIONS

Proof of Theorem 2. Let $P_i(w) = \lambda\|w\|^2/2 + \ell_i(w)$, $g_t = \lambda w_t - a_t y_{I_t} \phi_{I_t}$, and $s_t = s(a_t, m_t)$. Adding the exact quadratic identity to Eq. 29 gives

$$P_{I_t}(w_t) - P_{I_t}(w^\star) \le \langle g_t, w_t - w^\star\rangle - \frac{\lambda}{2}\|w_t - w^\star\|^2 + s_t.$$

Equation 27 implies $\|g_t\| \le 2$. Write $R_t = \|w_t - w^\star\|^2$ and $\eta_t = 1/(\lambda t)$. Expanding $w_{t+1} = w_t - \eta_t g_t$ yields

$$P_{I_t}(w_t) - P_{I_t}(w^\star) \le \frac{\lambda(t-1)}{2} R_t - \frac{\lambda t}{2} R_{t+1} + \frac{2}{\lambda t} + s_t. \tag{30}$$

Before drawing I_t, w_t is fixed and I_t is uniform. Therefore the conditional expectation of the left side is $P(w_t) - P(w^\star)$. On the call-coverage event G_t, $s_t \le \tau$; otherwise $s_t \le C_\lambda$. Proposition 3 and the risk allocation give $\sum_t \mathbb{E}s_t \le T\tau + C_\lambda\rho$. Summing Eq. 30 telescopes all distances to $-\lambda T \mathbb{E}R_{T+1}/2 \le 0$. Thus, with $H_T = \sum_{t=1}^T 1/t \le 1 + \log T$, convexity at the average gives

$$\mathbb{E}[P(\bar{w}_T) - P(w^\star)] \le \frac{2H_T}{\lambda T} + \tau + \frac{C_\lambda \rho}{T},$$

which is Eq. 15.

For a call capped at $R \ge 1$, couple the sampling to an infinite conditional i.i.d. sequence and let $\bar{X}_R$ be its first-R empirical mean, whether or not the call stops early. At the cap the empirical action is $\mathbf{1}\{\bar{X}_R < 1\}$. If this action is wrong, 1 lies between m_t and $\bar{X}_R$, so its slack is at most $|\bar{X}_R - m_t|$. Early covered decisions have slack at most τ, and any early coverage failure has slack at most C_λ. In all cases,

$$s_t \le \tau + C_\lambda \mathbf{1}_{G_t^c} + |\bar{X}_R - m_t|.$$

Conditional i.i.d. sampling and Eq. 28 imply

$$\mathbb{E}[|\bar{X}_R - m_t| \mid \text{call history}] \le \sqrt{\mathbb{E}[(\bar{X}_R - m_t)^2 \mid \text{call history}]} = \sqrt{\mathrm{Var}(X_q)/R} \le \frac{1}{2\lambda\sqrt{R}}.$$

This expectation is taken before conditioning on reaching the cap. Inserting it in the same telescoping argument proves the additional term. □

D.6 FINITE STOPPING BOUNDS FOR THE BETA-MIXTURE ORACLE

For the quantitative shot bounds fix $0 < a, b \le 1$, including the uniform prior $(1, 1)$ and the Jeffreys prior $(1/2, 1/2)$. Coverage in Proposition 3 holds for all positive a, b. Put $\kappa = 1/\mathrm{B}(a, b) \in (0, 1]$ and $\Lambda_n = \log((n + 1)/(\kappa\delta))$. The beta density is at least κ on $(0, 1)$, so the mixture dominates κ times the uniform-prior mixture. For $q = S_n/n$, the latter is

$$M_n^{\mathrm{unif}}(p) = \frac{1}{(n+1)\binom{n}{S_n}p^{S_n}(1-p)^{n-S_n}} = \frac{\exp(n\,\mathrm{kl}(q,p))}{(n+1)\binom{n}{S_n}q^{S_n}(1-q)^{n-S_n}} \ge \frac{\exp(n\,\mathrm{kl}(q,p))}{n+1}.$$

The inequality uses that the denominator's binomial probability is at most one; $0^0 = 1$ covers $q = 0, 1$. Therefore every $p \in C_n$ satisfies

$$\mathrm{kl}(q, p) \le \frac{\Lambda_n}{n}, \qquad |p - q| \le \sqrt{\frac{\Lambda_n}{2n}}. \tag{31}$$

For completeness, the second inequality follows because, as a function of q, $\mathrm{kl}(q, p)$ is zero with zero derivative at $q = p$ and has second derivative $1/[q(1 - q)] \ge 4$. Twice integrating gives $\mathrm{kl}(q, p) \ge 2(q - p)^2$, with endpoints obtained by limits.

For the affine sample $X = c + wB$, set $r_n = w\sqrt{\Lambda_n/(2n)}$. Its confidence interval is contained in $[c + wq - r_n, c + wq + r_n]$, so its width is at most $2r_n$ on every path. The smaller hinge slack is at most half this width: if the interval crosses one, the two slacks sum to its width, and otherwise one slack is zero. Hence the rule stops whenever $r_n \le \tau$, even on a coverage-failure path. On coverage, write $\Delta = |1 - m|$ and $d = \max\{\Delta, \tau\}$. If $r_n \le d/4$, then either $d = \tau$ and the width criterion holds, or $d = \Delta$ and the covered interval, of width at most $\Delta/2$, lies strictly on the correct side of one. This also forces stopping.

Lemma 5 (An explicit logarithmic envelope)**.** *For $A \ge 1$ define*

$$\mathcal{N}(A, \delta) = \left\lceil 4A \log \frac{8A}{\kappa\delta} \right\rceil.$$

At $n = \mathcal{N}(A, \delta)$, $n \ge A \log((n + 1)/(\kappa\delta))$.

Proof. Put $L = \log(8A/(\kappa\delta)) > 2$. Then $n \ge 4AL$ and $n + 1 \le 4AL + 2 \le 6AL$. Consequently

$$\log \frac{n+1}{\kappa\delta} \le L + \log(3L/4) \le 2L \le n/A,$$

where $\log(3L/4) \le L$ follows from $3L/4 \le e^L$ for $L \ge 0$. □

Let R count the samples requested by a strict call, with $R = 0$ for an analytically resolved call, and assume $\tau > 0$. Applying the lemma with $A_d = 1 + 8w^2/d^2$ and $A_\tau = 1 + 8w^2/\tau^2$ gives

$$R \le \mathcal{N}(A_\tau, \delta) \quad \text{on every path}, \qquad R\mathbf{1}_G \le \mathcal{N}(A_d, \delta)\mathbf{1}_G. \tag{32}$$

Indeed $n \ge 8w^2\Lambda_n/d^2$ implies $r_n \le d/4$; the same argument with τ supplies the pathwise bound. No asymptotic logarithm is needed to cover the first requested sample, since $A_d, A_\tau \ge 1$. Physical shots are at most requested samples, and

$$\mathbb{E}[R \mid \text{call history}] \le \mathcal{N}(A_d, \delta) + \delta\mathcal{N}(A_\tau, \delta). \tag{33}$$

Here Δ, w are fixed by the call history and $\mathbb{P}(G^c \mid \text{call history}) \le \delta$. This proves both the instance-dependent bound and the failure-cost term.

Variance-sensitive refinement. The same confidence set adapts to the conditional variance $\sigma^2 = w^2p(1 - p)$. To see this without replacing the stopping rule, write $v_q = q(1 - q)$ and $u = |p' - q|$. Integrating the derivative of $\mathrm{kl}(q, p')$ in its second argument gives

$$\mathrm{kl}(q, p') \ge \int_0^u \frac{z}{v_q + z}\, dz \ge \frac{u^2}{2(v_q + u)},$$

since $p'(1-p') \le v_q + |p' - q|$ along either direction from q. Thus $p' \in C_n$ implies, with $\ell = \Lambda_n/n$, $|p' - q| \le \sqrt{2v_q\ell} + 2\ell$. On coverage this also holds for the true p. Since $|v_q - p(1 - p)| \le |q - p|$,

putting $x = \sqrt{v_q}$ gives $x^2 \le p(1-p) + \sqrt{2\ell}\, x + 2\ell$, and hence $x \le \sqrt{p(1-p)} + 3\sqrt{\ell}$ by completing the square. Every point of the margin confidence interval is consequently within

$$r_n^{\mathrm{var}} := w(\sqrt{2v_q\ell} + 2\ell) \le \sqrt{2\sigma^2\ell} + 7w\ell$$

of $c + wq$ on coverage. For $A_{\mathrm{var}} = 1 + 128\sigma^2/d^2 + 56w/d$, the time $n = \mathcal{N}(A_{\mathrm{var}}, \delta)$ makes each of the last two terms at most $d/8$. The preceding stopping argument therefore proves

$$\mathbb{E}[R \mid \text{call history}] \le \mathcal{N}(A_{\mathrm{var}}, \delta) + \delta\mathcal{N}(A_\tau, \delta). \tag{34}$$

In an interior Bernoulli family and for small d, the leading instance term is $\widetilde{O}(\sigma^2/d^2)$; the linear w/d term and the explicit failure contribution remain as displayed.

D.7 Accuracy, worst-case shots, and margin adaptation

Lemma 4 and Theorem 2 give

$$\mathbb{E}e(\bar{w}_T)^2 \le \frac{2}{\lambda}\left[\frac{2(1+\log T)}{\lambda T} + \tau + \frac{C_\lambda \rho}{T}\right]. \tag{35}$$

Thus the AMO accuracy statements are root-mean-square guarantees for the uniform decision error. For a concrete choice with $0 < \lambda, \epsilon \le 1$, set

$$\tau = \lambda\epsilon^2/8, \quad \rho = \tau^2/4, \quad A = 64/(\lambda^2\epsilon^2), \quad T = \lceil A \log A \rceil, \quad \delta_t = \rho/T.$$

Since $A \ge 64$, $T \le 2A\log A$, $\log T \le 2\log A$, and $1 + \log T \le 3\log A$. The optimization term is at most $6/(\lambda A) = 3\lambda\epsilon^2/32$; the confidence term is at most $\lambda\epsilon^2/128$, using $C_\lambda \le 2/\lambda$ and $T \ge 1$. Together with $\tau = \lambda\epsilon^2/8$, their sum is less than $\lambda\epsilon^2/2$, proving $\mathbb{E}e(\bar{w}_T)^2 \le \epsilon^2$. For any fixed $0 < \zeta < 1$, replacing ϵ by $\sqrt{\zeta}\,\epsilon$ and applying Markov's inequality gives error at most ϵ with probability at least $1 - \zeta$, with the same powers of ϵ.

The support width obeys $w \le \lambda^{-1}$. Hence Eq. 32 gives the deterministic training-cost bound

$$Q \le T\mathcal{N}(1 + 8\lambda^{-2}\tau^{-2}, \rho/T) = \widetilde{O}\left(T[1 + \lambda^{-2}\tau^{-2}]\right) = \widetilde{O}(\lambda^{-6}\epsilon^{-6}), \tag{36}$$

where the last simplification uses the stated parameter ranges. In these expressions the hidden factors are logarithms of positive quantities such as $8(1 + 8\lambda^{-2}\tau^{-2})T/(\kappa\rho)$.

Now assume Eq. 17, with constants uniform in rounds, histories and the accuracy sequence, and $0 < \tau \le u_0$. For every nonnegative gap Δ,

$$(\Delta \vee \tau)^{-2} \le u_0^{-2} + 2\int_\tau^{u_0} \mathbf{1}\{\Delta \le u\} u^{-3}\, du.$$

For $\Delta \le u_0$ this is equality, and otherwise the left side is at most u_0^{-2}. Conditional Tonelli integration and the margin law give

$$\mathbb{E}[(\Delta_t \vee \tau)^{-2} \mid H_t] \le u_0^{-2} + 2C_m I_\beta(\tau), \tag{37}$$

$$I_\beta(\tau) = \begin{cases} (\tau^{\beta-2} - u_0^{\beta-2})/(2-\beta), & 0 \le \beta < 2, \\ \log(u_0/\tau), & \beta = 2, \\ (u_0^{\beta-2} - \tau^{\beta-2})/(\beta-2), & \beta > 2. \end{cases}$$

Here H_t is the history before the uniform training index is drawn. Let $L = \log(8(1 + 8\lambda^{-2}\tau^{-2})T/(\kappa\rho))$. Since $A_d \le A_\tau \le 1 + 8\lambda^{-2}\tau^{-2}$, Eq. 32 implies, for a numerical constant C,

$$\mathbb{E}[R_t \mathbf{1}_{G_t} \mid H_t] \le CL\left(1 + \lambda^{-2}\mathbb{E}[(\Delta_t \vee \tau)^{-2} \mid H_t]\right).$$

This uses the original conditional law with an event indicator; it does not condition that law on G_t. The pathwise cap separately gives $\mathbb{E}[R_t \mathbf{1}_{G_t^c} \mid H_t] \le (\rho/T)\mathcal{N}(1 + 8\lambda^{-2}\tau^{-2}, \rho/T)$. Summing both contributions yields

$$\mathbb{E}Q \le CTL[1 + \lambda^{-2}(u_0^{-2} + 2C_m I_\beta(\tau))] + \rho\mathcal{N}(1 + 8\lambda^{-2}\tau^{-2}, \rho/T). \tag{38}$$

For fixed λ, C_m, u_0, a, b, the choices $T = \widetilde{O}(\epsilon^{-2})$, $\tau = \Theta(\epsilon^2)$ and $\rho = \Theta(\tau^2)$ give respectively $\widetilde{O}(\epsilon^{-(6-2\beta)})$, $\widetilde{O}(\epsilon^{-2}\log(1/\epsilon))$, and $\widetilde{O}(\epsilon^{-2})$ in the three cases of Eq. 37. The failure term is only logarithmic at fixed λ. This proves Eq. 18. Combining Eq. 36 with the RSCA bound, and choosing the branch in advance from N, ϵ, proves Eq. 19.

D.8 SEQUENTIAL LOWER BOUNDS AND LEGAL KERNEL INSTANCES

Lemma 6 (Sequential testing cost)**.** *Consider two oracle instances with the same adaptive decision rule and internal randomization. Suppose every paid query has conditional KL divergence at most d_{KL} from the minus to the plus instance, and zero-cost observations have the same law. If the procedure stops almost surely under both instances and identifies each with error at most $0 < \rho < 1/2$, its number R of paid queries satisfies*

$$d_{\mathrm{KL}}\mathbb{E}_- R \geq \mathrm{kl}(1-\rho, \rho). \tag{39}$$

Proof. Pad a stopped transcript with a fixed termination symbol. The probability of a finite transcript factors into the algorithm's conditional query and stopping probabilities and the oracle's conditional outcome probabilities. The algorithm factors are the same under the two instances and cancel in the likelihood ratio. Taking the minus-law expectation of its logarithm therefore gives, for the first n paid-query opportunities,

$$D_{\mathrm{KL}}(P_-^{(n)} \| P_+^{(n)}) = \sum_{t=1}^{n} \mathbb{E}_-[\mathbf{1}\{R \geq t\} D_{\mathrm{KL}}(P_-(Z_t \mid H_t) \| P_+(Z_t \mid H_t))] \leq d_{\mathrm{KL}}\mathbb{E}_-(R \wedge n).$$

This identity also permits common internal random seeds by conditioning on them and including them in the transcript. Let A_n be the event that the procedure has stopped by opportunity n and declared minus. Applying the log-sum inequality separately on A_n and its complement gives

$$D_{\mathrm{KL}}(P_-^{(n)} \| P_+^{(n)}) \geq \mathrm{kl}(P_-(A_n), P_+(A_n)).$$

Almost-sure stopping implies that these probabilities converge to $p \geq 1-\rho$ and $q \leq \rho$, respectively. Monotone convergence gives $\mathbb{E}(R \wedge n) \uparrow \mathbb{E}R$. Continuity, or lower semicontinuity at endpoints, of binary KL implies $d_{\mathrm{KL}}\mathbb{E}_- R \geq \mathrm{kl}(p, q)$. For $p > q$, binary KL increases with p and decreases with q, as follows by differentiation. It is therefore at least $\mathrm{kl}(1-\rho, \rho)$. Infinite expected cost already satisfies the asserted lower bound. □

Local margin testing. Fix $b > 0$ and $\Delta > \tau \geq 0$ such that $p_\pm = (1 \pm \Delta)/b \in [p_0, 1-p_0]$ for fixed $0 < p_0 < 1/2$. Let a shot be $X = bZ$, $Z \sim \mathrm{Bernoulli}(p_\pm)$, so the two margins are $1 \pm \Delta$. A wrong hinge action has slack $\Delta > \tau$; consequently an oracle returning a τ-valid action with probability $1-\rho$ under each instance is a test of these two hypotheses. The inequality $\log x \leq x - 1$ gives

$$\mathrm{kl}(p, q) \leq p(p/q - 1) + (1-p)((1-p)/(1-q) - 1) = \frac{(p-q)^2}{q(1-q)}.$$

Thus each Bernoulli observation has KL at most $4\Delta^2/[b^2 p_0(1-p_0)]$ in either direction. Lemma 6 implies

$$\mathbb{E}_\pm R \geq \frac{b^2 p_0(1-p_0)}{4\Delta^2}\,\mathrm{kl}(1-\rho, \rho). \tag{40}$$

For fixed $\rho_0 < 1/2$ and $0 < \rho \leq \rho_0$,

$$\begin{aligned}\mathrm{kl}(1-\rho, \rho) &= (1-2\rho)\log\frac{1-\rho}{\rho} \\ &\geq (1-2\rho_0)\left(1 - \frac{\log 2}{\log(1/\rho_0)}\right)\log(1/\rho) =: c_0 \log(1/\rho),\end{aligned}$$

where $c_0 > 0$. Since $\sigma_\pm^2 = b^2 p_\pm(1-p_\pm) \leq b^2/4$, Eq. 40 proves $\mathbb{E}_\pm R = \Omega(\sigma_\pm^2 \Delta^{-2} \log(1/\rho))$. Together with Eq. 34, this identifies the leading local margin dependence, with the upper bound retaining its displayed lower-order and failure-cost terms.

Realization by normalized quantum states. These oracle instances can be fidelity kernels rather than arbitrary entry arrays. For $N \geq 2$, take a Hilbert space with orthonormal basis $|0\rangle, |1\rangle, \ldots, |N\rangle$, define, for $0 < p < 1$,

$$|\psi_j(p)\rangle = p^{1/4}|0\rangle + \sqrt{1-\sqrt{p}}\,|j\rangle, \qquad j = 1, \ldots, N.$$

Each state has norm one. Distinct states have inner product $\sqrt{p}$, so their fidelity is p, and the Gram matrix is $K(p) = (1-p)I + p\mathbf{1}\mathbf{1}^\top \succeq 0$ with diagonal one. It embeds in $\lceil \log_2(N+1) \rceil$ qubits. For a fixed row, any nonnegative combination of off-diagonal fidelities with total weight b has margin bp. Every off-diagonal raw query has law $\mathrm{Bernoulli}(p)$, and all diagonal answers are the same known constant under the two instances. Adaptive pair choices therefore satisfy the testing lemma. A call of AMO itself realizes $m = bp$ by taking $\lambda = 1/b$, two positive labels, $t = 2$, a first active example j, and a current example $i \neq j$: Eq. 27 then gives $m = K_{ij}/\lambda$. This constructs a legal local margin family; the lower bound concerns the margin oracle.

Exponentially small per-query information. Suppose two task instances require opposite decisions and, for every available query and history, their outcome laws have conditional KL at most Ce^{-2cq}, where q is the number of qubits and $C, c > 0$ are fixed. The same transcript proof gives

$$\mathbb{E}_- R \geq C^{-1} e^{2cq} \,\mathrm{kl}(1-\rho, \rho).$$

Hence fixed error probability below $1/2$ requires exponentially many shots. If both laws are interior Bernoulli distributions with probabilities in $[p_0, 1-p_0]$ and their differences are at most $C_1 e^{-cq}$, the preceding Bernoulli KL inequality supplies the required uniform bound with $C = C_1^2/[p_0(1-p_0)]$. The uniform KL condition is the information assumption used in this conclusion.

E Experimental details

Pseudocode. Algorithm 1 is the shared-shot control RSCA used in both studies, and Algorithm 2 is AMO-Pegasos with the centered sample and the stopping rule of Eq. 12; the strict runs use the conservative variant described below.

Data and configurations. Balanced labels determine two angle clusters with means 0 and 0.9 radians and Gaussian standard deviation 0.18. Every measurement is a fresh Bernoulli draw from the corresponding fidelity. The methods share the data in each paired trial; hyperparameters are fixed using separate development trials. Exact kernels provide the reference SVM and decision errors.

Shared-shot control RSCA uses 19,000 ordered-pair rounds and four landmarks, with three shots per required landmark pair (738 acquisition shots). The control mixes a constant kernel with the Nyström fit, $B_c = 0.3\,\mathbf{1}\mathbf{1}^\top + 0.4\,B_{\mathrm{Nys}}$ (landmark-Gram eigenvalues at most 0.5 dropped), so F stacks $\sqrt{0.3}\,\mathbf{1}$ and $\sqrt{0.4}\,F_{\mathrm{Nys}}$; a residual request is skipped only when every trajectory can resolve it analytically. The retained control rank is 2–4, with median 3; the shared ensemble costs 3.65×10^7 classical operations per run at $N = 64$ and $L = 5$. Five trajectories with multipliers $m \in \{2, 3, 4, 5, 6\}$ share each residual shot; trajectory m uses observed-gradient AdaGrad (Appendix D.3) with $v_0 = 10^{-6}$ and base step

$$a_0 = \frac{c\,m}{(\lambda + 1 + 2L_F^2)N\sqrt{f}}, \qquad f = \tfrac{1}{2} + \tfrac{1}{2N},$$

where $L_F = \max_i \|F_i\|$ is the largest control-feature row norm and c a constant. The base step is proportional to $1/N$ and independent of T, the choice for which Eq. 26 gives the $\sqrt{N/T}$ order. The base step carries no factor B^{-1}: AdaGrad already normalizes by the gradient scale, and a B^{-1} factor would shrink the effective step with λ a second time. The synthetic study sets $c = \lambda = 0.2$; the MNIST study uses $c = 0.15$, selected on development seeds (below). Averages over the last half and last quarter of each trajectory receive equal weights. Cached termwise Pegasos follows Eq. 8 with step $1/(\lambda t)$, $\{1, 3, 10, 12, 14, 15\}$ shots per newly queried pair (ten in Figure 5b; ten and twelve in Table 2), and checkpoints at $\{300, 1000, 3000, 10000\}$ rounds, reporting the uniform average of its iterates; both interfaces therefore use iterate averaging, and RSCA adds suffix averages and five replicas. Its shot count is set by the pairs it visits. Full QKE uses $\{1, 3, 10, 30\}$ shots per entry, Nyström SVMs use $\{4, 8, 16\}$ landmarks with $\{1, 3, 10\}$ shots per pair, and fixed aggregate Pegasos takes $\{1, 3, 10\}$ samples per margin. The symmetric and control RSCA ablations run 20,000 and 23,200 rounds with AdaGrad multipliers 1 and 4.

AMO configurations. All AMO runs use $\rho = 0.05$. The strict margin-law trajectories use the centered sample of Eq. 11 with the beta-mixture sequence of Appendix B (Jeffreys prior $a = b = 1/2$; interval endpoints are computed in the log domain by safeguarded Newton steps inside an

Algorithm 1 Shared-shot control RSCA

Require: labels y, regularization λ, rounds T, control factor $F \in \mathbb{R}^{N\times r}$ from the landmark shots, multipliers $m_1, \dots, m_L$

for each ℓ: $\alpha^{(\ell)} \leftarrow 0$, $v^{(\ell)} \leftarrow 0 \in \mathbb{R}^r$, $s^{(\ell)} \leftarrow v_0^2\mathbf{1}$, base step a_ℓ (below, with $m = m_\ell$)
for $t = 1, \dots, T$ **do**
  draw I, J uniformly and independently from $\{1, \dots, N\}$
  for each ℓ: $h^{(\ell)} \leftarrow N^{-1}(\mathbf{1} - \lambda^{-1} y \odot F v^{(\ell)})$ ▷ exact control gradient
  if $\alpha_I^{(\ell)} > 0$ or $\alpha_J^{(\ell)} > 0$ for some ℓ **then**
    obtain Z_{IJ} (a shot unless $I = J$); $e \leftarrow Z_{IJ} - F_I^\top F_J$
    for each ℓ: $h_I^{(\ell)} \leftarrow h_I^{(\ell)} - \frac{N}{2\lambda} y_I y_J \alpha_J^{(\ell)} e$, $h_J^{(\ell)} \leftarrow h_J^{(\ell)} - \frac{N}{2\lambda} y_J y_I \alpha_I^{(\ell)} e$
  end if
  for each ℓ: $s^{(\ell)} \leftarrow s^{(\ell)} + h^{(\ell)} \odot h^{(\ell)}$, $\alpha' \leftarrow \mathrm{clip}_{[0,1/N]}\big(\alpha^{(\ell)} + a_\ell\, h^{(\ell)} \oslash \sqrt{s^{(\ell)}}\big)$,
    $v^{(\ell)} \leftarrow v^{(\ell)} + F^\top((\alpha' - \alpha^{(\ell)}) \odot y)$, $\alpha^{(\ell)} \leftarrow \alpha'$
end for
return the equal-weight mean over ℓ of the averages of $\alpha_t^{(\ell)}$ over the last half and the last quarter of the rounds

Algorithm 2 AMO-Pegasos with the centered sample

Require: labels y, λ, rounds T, tolerance τ, call risks δ_t, optional cap R

$\mathcal{H} \leftarrow \emptyset$ ▷ active multiset; $w_t = \frac{1}{\lambda(t-1)} \sum_{j\in\mathcal{H}} y_j \phi_j$
for $t = 1, \dots, T$ **do**
  draw I uniformly; $A \leftarrow |\mathcal{H}|$
  if $t = 1$ or $A < \lambda(t-1)$ **then** $a \leftarrow 1$ ▷ $m_t < 1$ exactly; no shot
  else
    $d \leftarrow$ multiplicity of I in $\mathcal{H}$; $k \leftarrow A - d$; $b \leftarrow k/[\lambda(t-1)]$
    if $k = 0$ **then** $a \leftarrow \mathbf{1}\{d/[\lambda(t-1)] < 1\}$
    else
      $\bar{s} \leftarrow$ mean of $y_I y_j$ over the other k occurrences j
      for $n = 1, 2, \dots$ **do**
        draw J uniformly from the other k occurrences; take a shot Z_{IJ}
        $B_n \leftarrow \frac{1}{2} + y_I y_J (Z_{IJ} - \frac{1}{2})$; update the interval for $\mathbb{E}B$ at risk δ_t
        map it to $[L_n, U_n]$ by $x \mapsto d/[\lambda(t-1)] + b(\bar{s}/2 - 1/2 + x)$
        if $U_n \le 1 + \tau$ or $L_n \ge 1 - \tau$ **then** $a \leftarrow$ the action with the smaller slack; **break**
        end if
        if $n = R$ **then** $a \leftarrow \mathbf{1}\{$sample mean of the margin $< 1\}$; **break**
        end if
      end for
    end if
  end if
  if $a = 1$ **then** add I to $\mathcal{H}$
  end if
end for
return an average of the iterates w_t

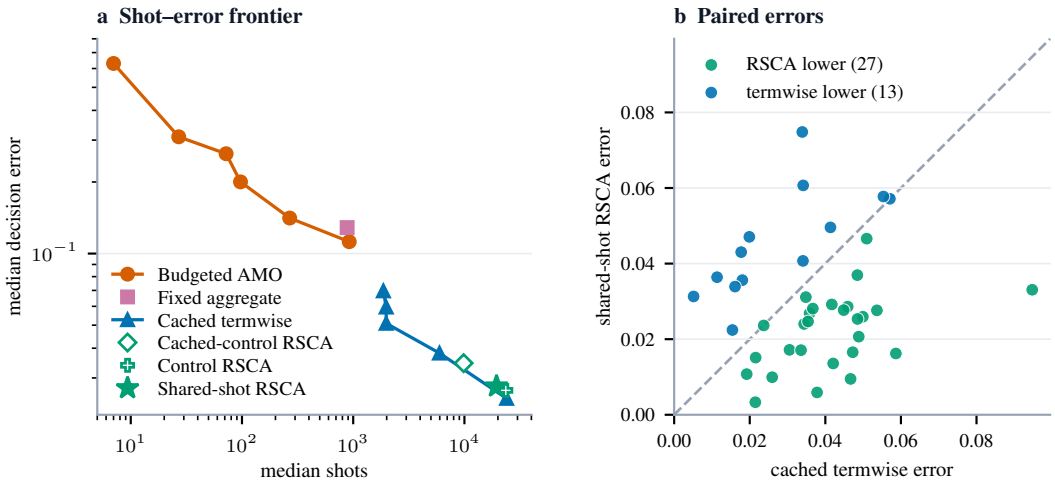

Figure 5: **Synthetic family.** **(a)** Median shot–error frontier; fixed aggregate Pegasos takes a fixed number of margin samples per call without a certificate; cached-control RSCA uses running means of earlier outcomes as its control. **(b)** Paired errors for shared-shot RSCA and cached termwise Pegasos (10 shots per new pair); points below the diagonal favor RSCA.

enclosing bracket, returning the outer end so that numerical error can only widen the interval), $\tau = 0.25$, $\lambda = 0.5$, 300 rounds, seeds 0–19, and call risks $6\rho/(\pi^2 t^2)$ that sum to at most ρ over any horizon; since $6\rho/(\pi^2 t^2) \geq 6\rho/(\pi^2 T^2)$ for $t \leq T$, the shot bound of Appendix D.6 holds with ρ/T replaced by this value, which changes only the logarithm. These strict runs use a conservative form of Eq. 12: a call stops once $[L_n, U_n]$ excludes one or extends at most $\tau/2$ on each side of the sample mean, and acts on the side of the mean. Equation 12 already holds whenever this rule stops, so the reported sample counts bound those of Eq. 12 on the same outcomes, and every covered action has slack at most $\tau/2$. The budgeted frontier points use the centered sample, $\tau = 0.3$, $T \in \{10, 30, 100, 300\}$ Pegasos rounds, caps $R \in \{1, 3, 10, 30\}$ (the 921-shot point has $T = 100$, $R = 10$), call risks ρ/T, and Hoeffding radii $b\sqrt{\log(2/\delta_k)/(2n)}$ inspected at sample counts $1, 2, 4, \ldots$ with per-inspection risks δ_k summing to the call's risk; this is anytime-valid by a union bound, so Eq. 12 keeps its certificate, with a larger constant than the beta-mixture bound of Appendix D.6. Budgeted calls also update without shots when the number of same-label active occurrences divided by $\lambda(t-1)$, an upper bound on m_t since $K \leq 1$, is below one. Nyström SVMs are reported as surrogates because they change the target.

Protocol run and audit. All synthetic numbers in Section 5, Table 2, and Figure 5 come from one run of the pre-specified protocol (seeds 3000–3039, 166 configurations per seed), and an independent shot audit checked all 6,640 run records. On the frontier of Figure 5a, budgeted AMO reaches 0.112 at 921 shots and the best Nyström surrogate 0.151 at 2,460.

Paired sign test. Across the 40 independent paired trials, 27 favor shared-shot RSCA and 13 favor cached termwise Pegasos, with no ties. Under the null of equal win probability, the count is $\mathrm{Binomial}(40, 1/2)$, so the exact two-sided probability is $2\sum_{k=0}^{13}\binom{40}{k}2^{-40} = 0.038$. The comparison uses the pre-specified 10,000-round checkpoint of the baseline; the same 27–13 split holds against its 1,000- and 3,000-round checkpoints (median errors 0.0336 and 0.0347), so the result does not depend on that choice. RSCA also used no more shots than the baseline in 39 of the 40 trials.

Margin analysis. The 20 strict-AMO trajectories ($\lambda = 0.5$) use 337,769 shots, 16,888 per trajectory; the same trajectories with Hoeffding radii on the uncentered sample of Eq. 10 use 8,347,562. Every call stops certified, and replaying the exact margins shows 10 of 6,000 actions differing from the exact hinge action, all with exact slack below τ. In the bins of Figure 4, median samples per decision are 491–829 for gaps below 0.12, 276 near 0.15, 60 near 0.26, and 1 above 0.34. Cost fits use rounds with $\Delta_t > \tau$, positive sample count, and positive scale b_t, excluding ambiguous and cap-exhausted calls; we fit logical samples divided by b_t^2 against the gap separately for each trajectory

(mean slope -2.16, standard deviation 0.76) and pooled over trajectories (-2.08, whole-trajectory bootstrap 95% interval $[-2.46, -1.74]$). Cost reaches the one-sample floor above gap about 0.35 and plateaus at several hundred samples below gap $0.12 \approx \tau/2$, where the half-width stopping rule caps it. The pooled CDF weights each visited round equally; its 95% interval uses 1,000 bootstrap resamples of whole trajectories.

Concentration check. For $q = 1, \ldots, 10$, a Hoeffding test with doubling inspections and risk $\rho = 0.05$ decides between Bernoulli means $\frac{1}{2} \pm 0.1e^{-0.45q}$, with 100 repetitions per q and seeds 0–9; the fitted rate is the slope of log median shots in q, and the predicted rate is 2×0.45; the fit gives 0.907 per qubit (empirical 95% interval $[0.872, 0.930]$).

Mechanism details and variants. The AMO panels test both links behind adaptivity: per-decision cost responds as Δ^{-2}, while the optimizer path supplies a distribution of gaps. Panel (c) draws gaps from each controlled law and sums the per-call cost $\max\{\Delta, \tau\}^{-2}$, with dyadic inspections and a fixed logarithmic factor, over $2000/\tau$ rounds; it isolates the integration behind Eq. 18. For controlled $\beta = 0, 0.5, 1, 2, 3$, the fitted exponents are 6.00, 5.05, 4.13, 2.75, and 2.16 versus 6, 5, 4, 2 with a logarithmic boundary term, and 2. The largest finite-range deviation occurs at the logarithmic boundary $\beta = 2$.

The RSCA variants identify the gain: symmetric endpoint reuse alone reaches 0.0353 median error; the frozen control reaches 0.0266 at higher cost; and the shared-control ensemble retains most of this reduction at 19,443 shots by preventing five schedules from multiplying QPU cost. Its classical cost scales with the number of trajectories and the control rank.

MNIST protocol. Pairs are drawn class-balanced from the 70,000-image MNIST pool; each cell uses N training and 1,000 test images. We use $\lambda = 0.05$ because at $\lambda = 0.2$ and $N = 64$ most exact dual coordinates sit at the upper bound. Standardization, PCA to four components, and clipping to $[-\pi/2, \pi/2]$ are fitted on the training split only. The feature map is the four-qubit product-RY map with exact fidelity $K(x, z) = \prod_k \cos^2((x_k - z_k)/2)$; the exact Gram matrix is computed only by the evaluator. The frozen RSCA configuration is that of the synthetic study (four landmarks, three shots per landmark pair, multipliers $\{2, \ldots, 6\}$, last-half and last-quarter suffix averages) with horizons $T \in \{150N, 300N, 600N, 1200N\}$. The step constant c was selected on development seeds 0–19 of 3 vs 8 by a sweep over $c \in [0.02, 0.6]$ at $\lambda \in \{0.02, 0.05\}$ and $N \leq 512$: the best value is roughly constant in N and λ, and, for $N \geq 128$, $c = 0.15$ is within 1% of the per-cell optimum in geometric mean. The rule was then frozen and every reported MNIST number comes from fresh seeds: 40–59 on 3 vs 8, 60–79 on 2 vs 3 and 0 vs 1, and 80–99 for the noise study. Baselines per cell: cached termwise Pegasos with $\{1, 3, 10\}$ shots per pair and checkpoints at $\{5N, 15N, 50N, 150N\}$ rounds; full QKE with $\{1, 3, 10\}$ shots per entry and PSD projection; Nyström SVMs with $\{8, 16, 32\}$ landmarks and $\{1, 3, 10\}$ shots. Shots-to-target is the smallest median shot count of any configuration whose median error is at most the target; slopes are least-squares fits in $\log N$. For 3 vs 8, 2 vs 3, and 0 vs 1, cached termwise Pegasos has slopes 1.99, 1.67, and 1.14 and needs 506k, 488k, and 416k shots at $N = 1024$, against 165.7k for RSCA (slope 1.01), so RSCA saves at least $3.0\times$, $2.9\times$, and $2.5\times$; at 319k shots RSCA reaches error 0.065. At target 0.1 the cached baseline still leads at $N = 512$ on 3 vs 8 and 2 vs 3; at $\lambda = 0.02$ it reaches error 0.3 at $N = 1024$ with about 20% fewer shots than RSCA on 3 vs 8 and 0 vs 1, while RSCA halves the cost on 2 vs 3. At $N = 512$ the one- and three-shot full-QKE matrices have smallest eigenvalues -19 and -11, and PSD projection pins every coordinate at the bound (error 1.23). Paired comparisons use the same twenty seeds and a two-sided sign test; above $N^2/2$ shots on 3 vs 8 the counts are 7–13 and 12–8 at 626k and 1.24M shots, neither significant, and on the other two pairs RSCA wins 13–7 to 20–0. Noise runs replace the evaluator kernel by the effective inversion-test kernel under per-qubit depolarizing p and readout error p_{ro} while every method keeps the unit-diagonal identity, so the diagonal bias is charged to the methods; it is at most 0.006 at $N = 1024$ and 0.083 at $N = 64$. At $p = p_{\mathrm{ro}} = 0.01$ the effective diagonal is 0.94, and at $N = 1024$ RSCA reaches 0.085 at 166k shots against 0.084 at 505k for cached termwise Pegasos; strong noise compresses kernel values toward $1/16$ and gives 12–8 at $N = 256$. Drift, run at $N = 512$ on top of the 1% noise, raises p linearly with the shots consumed; at 0.4 per million shots it degrades runs longer than about 300k shots and favors short RSCA runs, and a tenth of that rate gives parity (12–8). Distinct circuits count the pairs measured at least once: 136k for RSCA against 504k for cached termwise Pegasos at $N = 1024$.