Interaction Is Unnecessary for Order-Optimal One-Bit Mean Estimation

Pahan Dewasurendra
Johns Hopkins University

Abstract

We resolve the open question of whether interaction is necessary for order-optimal one-bit mean estimation under a finite central moment. Let $\mu \in [-\lambda, \lambda]$ and $\mathbb{E}|X-\mu|^k \le \sigma^k$ for fixed $k > 1$. We construct a fully nonadaptive protocol whose query list is fixed before any bit is observed and whose sample complexity matches the adaptive one-bit minimax rate in every moment regime. The refinement cost is $(\sigma/\epsilon)^{k/(k-1)} \log(1/\delta)$ for $1 < k < 2$, $(\sigma/\epsilon)^2 \log(\sigma/\epsilon) \log(1/\delta)$ for $k = 2$, and $(\sigma/\epsilon)^2 \log(1/\delta)$ for $k > 2$, plus the optimal $\log(\lambda/\sigma)$ localization cost.

The protocol first runs an existing nonadaptive codebook localizer. It then uses only decoding, not new queries, to choose a padded path through a prequeried dictionary of shifted modulo maps. Adjacent modulo remainders have finite-valued differences that are locally constant near the mean. Their variance is therefore charged only to samples that cross a scale-dependent boundary. A dyadic tail identity pays for all scales with one central-moment budget. This removes both the location dependence and the tail aliasing that obstruct global one-shot refinement.

1 Introduction

Suppose each of n agents observes one independent real sample and may send only one bit. Can a center recover the population mean as efficiently as if the bits were chosen interactively? This question separates communication from statistical information in its most elementary form.

For a distribution with mean $\mu \in [-\lambda, \lambda]$ and bounded k-th central moment, adaptive one-bit protocols are now understood sharply. Lau and Scarlett (2026b) give threshold-query estimators that match the one-bit minimax sample complexity in every regime $k > 1$. They also give a two-stage protocol using general one-bit queries. Its first, nonadaptive stage localizes μ to an interval of length $O(\sigma)$. Its second stage chooses refinement queries around that interval.

The remaining zero-adaptivity case was posed explicitly by Lau and Scarlett (2026a). Nonadaptive thresholds and intervals require a linear localization penalty in λ/σ, but that lower bound does not cover arbitrary measurable quantizers. General queries can encode many locations at once. The open problem asks whether they can simultaneously localize, refine, and control unbounded tails without an adaptive transition.

We give a positive answer. Our protocol fixes every query before observing any response and attains the adaptive minimax rate. The key is not a new use of periodic quantizers. Gray-code and modulo quantizers are established tools in distributed estimation and signal reconstruction (Cai and Wei, 2024; Boufounos, 2012; Mayekar et al., 2021). The key is a new statistical decomposition that makes periodic refinement robust to arbitrary distributions under only one finite moment.

For period P and shift s, let

$$R_{P,s}(x) = x - s - P \left\lfloor \frac{x-s}{P} \right\rfloor \in [0, P).$$

We estimate a remainder at a base scale and telescope its expectation through dyadically increasing periods. The difference of adjacent remainders is a finite-valued periodic function. After localization, the decoder selects one of three prequeried shifts whose discontinuities are far from μ. At scale P, the selected difference is then constant unless $|X - \mu| \gtrsim P$. Its estimator has variance proportional to

$$P^2 \mathbb{P}\{|X-\mu| \gtrsim P\},$$

not merely P^2. Summing this quantity over scales with an appropriate sample allocation costs only $\mathbb{E}|X-\mu|^k$.

A sufficiently large top period makes the remaining wraparound bias smaller than ϵ.

Our contributions are:

1. A fully nonadaptive one-bit protocol that matches the adaptive minimax rate over every finite-moment class $\mathcal{D}(k,\lambda,\sigma)$.

2. A multiscale modulo identity whose adjacent increments admit tail-sensitive one-bit variance bounds.

3. A decoder-side selection argument that turns a nonadaptive localization code into location-specific refinement without an additional query stage.

Relation to prior modulo methods. MODGAME uses Gray functions and known Gaussian convolution to attain optimal distributed Gaussian rates (Cai and Wei, 2024). Wyner–Ziv estimators use modulo residues when each datum is deterministically close to decoder side information (Mayekar et al., 2021). A scalar residue in $\{0,\dots,q-1\}$ requires $\log_2 q$ bits, and its no-alias condition is infeasible at $q=2$ for a positive uncertainty radius. Their one-bit distance-adaptive alternative assumes a bounded domain. We instead have side information near the population mean, while individual samples are unbounded. The adjacent-remainder construction controls the resulting aliases in aggregate.

2 Model and Main Result

For $k>1$ and $\lambda\ge\sigma>0$, define

$$\mathcal{D}(k,\lambda,\sigma)=\big\{D:\mu(D)\in[-\lambda,\lambda],\\ \mathbb{E}_D|X-\mu(D)|^k\le\sigma^k\big\}.$$

A one-bit protocol observes independent $X_t\sim D$ only through $Y_t=\mathbf{1}\{X_t\in A_t\}$. In a fully nonadaptive protocol, all measurable sets $A_1,\dots,A_n$ are selected before any Y_t is observed, possibly using protocol randomness. The final estimator may use the complete query list and bit transcript.

Write

$$\mathfrak{r}_k(\sigma,\epsilon,\delta)=\begin{cases}(\sigma/\epsilon)^2\log(1/\delta), & k>2,\\ (\sigma/\epsilon)^2\log(\sigma/\epsilon)\log(1/\delta), & k=2,\\ (\sigma/\epsilon)^{k/(k-1)}\log(1/\delta), & 1<k<2.\end{cases} \tag{1}$$

Theorem 1 (Interaction-free optimal rate)**.** *Fix $k>1$. There are constants $c_k,C_k>0$ such that for every $\lambda\ge\sigma>0$, $0<\epsilon\le c_k\sigma$, and $\delta\in(0,1/2)$, a fully nonadaptive one-bit protocol satisfies*

$$\sup_{D\in\mathcal{D}(k,\lambda,\sigma)}\mathbb{P}_D\{|\widehat{\mu}-\mu(D)|>\epsilon\}\le\delta$$

using

$$n\le C_k\left\{\log\frac{\lambda}{\sigma}+\mathfrak{r}_k(\sigma,\epsilon,\delta)\right\}$$

samples.

For δ below the universal constant in its statement, the matching lower bound of Lau and Scarlett (2026b, Theorem 9) applies to every one-bit protocol, hence also to fully nonadaptive protocols. Theorem 1 therefore closes the minimax rate in the high-confidence regime covered by that lower bound, while its upper bound holds for every $\delta\in(0,1/2)$.

3 Related Work and the Missing Step

Finite-moment one-bit estimation. The preliminary sequential estimator of Lau and Scarlett (2026c) covered finite variance with near-optimal logarithmic factors. The later full result (Lau and Scarlett, 2026b) handles every $k>1$, removes avoidable factors, and proves the one-bit penalty at $k=2$. Its nonadaptive codebook localizer is one of our ingredients. Its refinement partitions two tails around the localized interval and allocates queries geometrically. Those query sets depend on the first-stage output.

Abdalla and Chen (2026) connect uniformly dithered one-bit quantization with trimmed means. Their fully quantized estimator uses one global dithering range. Under a moment bound, making that range large enough to control truncation also inflates the Bernoulli variance. Their location-free variant uses a small unquantized pilot. Uniform dithering is also used for high-dimensional covariance, regression, and matrix completion (Chen et al., 2023). These methods do not provide a one-shot scalar refinement at the rates in (1).

Parametric models. Kipnis and Duchi (2022) characterize exact asymptotic efficiency for symmetric log-concave location families. Their concrete nonadaptive lower bound assumes bounded complexity of quantizer level sets. Kumar and Vatedka (2026) prove an exact nonadaptive versus adaptive MSE gap for several unknown-variance location-scale families. These are constant-efficiency statements for known families. They do not contradict equality of minimax orders over $\mathcal{D}(k,\lambda,\sigma)$.

Why periodic reuse was incomplete. The open-problem paper already identifies periodic queries as a

Table 1: One-dimensional finite-moment mean estimation. The refinement term $\mathfrak{r}_k$ is defined in (1). The interval-query entry is a lower bound that already holds for distributions supported on an interval of length σ. One transition means that refinement queries are chosen after localization.

Protocol	Query family	Adaptive transitions	Sample complexity or lower bound
Lau and Scarlett (2026b)	thresholds	sequential	$\Theta_k(\log(\lambda/\sigma) + \mathfrak{r}_k)$
Lau and Scarlett (2026b)	general	one transition	$\Theta_k(\log(\lambda/\sigma) + \mathfrak{r}_k)$
Lau and Scarlett (2026b) lower bound	intervals	zero	$\Omega((\lambda\sigma/\epsilon^2)\log(1/\delta))$
Theorem 1	general	zero	$\Theta_k(\log(\lambda/\sigma) + \mathfrak{r}_k)$

plausible route for light tails (Lau and Scarlett, 2026a). A fixed small period gives local resolution everywhere, but an outlier can wrap into the same residue as a central sample. A fixed large period controls wrapping, but stochastic quantization pays variance proportional to the square of that period. Our telescoping construction separates these roles. The base period is fixed at $\Theta(\sigma)$. Larger periods enter only through locally constant differences whose variance is multiplied by a tail probability.

4 A Prequeried Modulo Dictionary

Nonadaptive localization. Use an independent sample block for the deterministic codebook protocol of Lau and Scarlett (2026b, Theorem 16). With probability at least $1-\delta/2$, it returns an interval I containing μ with $|I| \le 100\sigma$. Its query cost is

$$O\left(\log\frac{\lambda}{\sigma} + \log\frac{1}{\delta}\right).$$

Let c be the midpoint. On the localization event

$$|c-\mu| \le 50\sigma. \tag{2}$$

All refinement samples below are independent of this block, and every candidate refinement query is issued concurrently with localization.

Padded shifts. Set $P_0 = 600\sigma$ and $P_i = 2^i P_0$. At period P, prequery the three candidate shifts

$$\mathcal{S}_P = \{0, P/3, 2P/3\}.$$

After computing c, the decoder chooses $s_i \in \mathcal{S}_{P_i}$ maximizing $\mathrm{dist}(c, s_i + P_i\mathbb{Z})$. Three equally spaced boundary lattices guarantee

$$\mathrm{dist}(c, s_i + P_i\mathbb{Z}) \ge P_i/3.$$

Combining this with (2) gives

$$\mathrm{dist}(\mu, s_i + P_i\mathbb{Z}) \ge P_i/4. \tag{3}$$

The choice is data dependent only at the decoder. Queries for all candidates were fixed in advance.

Protocol at a glance. For each independent refinement repetition, the fixed query list contains:

1. dithered remainder queries for all three base shifts,
2. level-set queries for all nine adjacent-shift pairs at every scale,
3. no query that depends on the localization bits.

The decoder localizes once, selects one padded shift per scale, telescopes the corresponding empirical increments, unwraps at the top scale, and takes a median across repetitions.

Base remainder. For every $s \in \mathcal{S}_{P_0}$, draw independent public $U_j \sim \mathrm{Unif}[0, P_0]$ before querying and record

$$Y_j = \mathbf{1}\{R_{P_0,s}(X_j) \ge U_j\}.$$

Then $\mathbb{E}[P_0 Y_j \mid X_j] = R_{P_0,s}(X_j)$. Thus

$$\widehat{B}_s = \frac{P_0}{n_0}\sum_{j=1}^{n_0} Y_j$$

is unbiased for $\mathbb{E}R_{P_0,s}(X)$ and

$$\mathrm{Var}(\widehat{B}_s) \le \frac{P_0^2}{4n_0}. \tag{4}$$

Adjacent increments. For $s \in \mathcal{S}_P$ and $t \in \mathcal{S}_{2P}$, direct expansion gives

$$R_{2P,t}(x) - R_{P,s}(x) = s - t + PJ_{P,s,t}(x), \tag{5}$$

$$J_{P,s,t}(x) = \left\lfloor\frac{x-s}{P}\right\rfloor - 2\left\lfloor\frac{x-t}{2P}\right\rfloor.$$

The function $J_{P,s,t}$ is $2P$-periodic and takes at most three values, all in $\{-1,0,1,2,3\}$. For every one of the nine candidate pairs, we preallocate fresh samples to the five level-set queries

$$\mathbf{1}\{J_{P,s,t}(X) = \ell\}, \qquad \ell = -1,\dots,3.$$

The exact alphabets are

	$t/P=0$	$t/P=2/3$	$t/P=4/3$
$s/P=0$	$\{0,1\}$	$\{0,1,2\}$	$\{1,2,3\}$
$s/P=1/3$	$\{-1,0,1\}$	$\{0,1,2\}$	$\{1,2\}$
$s/P=2/3$	$\{-1,0,1\}$	$\{0,1\}$	$\{0,1,2\}$.

The short alphabet makes each increment estimable by a constant-size one-bit dictionary.

For the selected pair, let $\ell_0 = J_{P,s,t}(c)$ and $p_\ell = \mathbb{P}\{J_{P,s,t}(X) = \ell\}$. Use

$$\mathbb{E}J_{P,s,t}(X) = \ell_0 + \sum_{\ell \neq \ell_0} (\ell - \ell_0) p_\ell \tag{6}$$

to estimate (5). This is conditionally unbiased given the localization transcript. More importantly, the discontinuities of J lie on the two selected boundary lattices. Equation (3) implies

$$J_{P,s,t}(X) \neq \ell_0 \implies |X - \mu| \geq P/4.$$

Independent empirical level probabilities based on n_i samples each therefore give

$$\mathrm{Var}(\widehat{D}_i \mid c) \leq \frac{CP_i^2}{n_i} \mathbb{P}\{|X - \mu| \geq P_i/4\}. \tag{7}$$

This local variance bound is the core of the protocol.

Telescoping decoder. For any shifts, pointwise in x,

$$R_{P_L,s_L}(x) = R_{P_0,s_0}(x) + \sum_{i=0}^{L-1} \{R_{P_{i+1},s_{i+1}}(x) - R_{P_i,s_i}(x)\}. \tag{8}$$

Let

$$A = s_L + P_L \left\lfloor \frac{c - s_L}{P_L} \right\rfloor.$$

One refinement estimate is

$$\widetilde{\mu} = A + \widehat{B}_{s_0} + \sum_{i=0}^{L-1} \widehat{D}_i.$$

We repeat the full prequeried refinement dictionary independently and return the median of the resulting $\widetilde{\mu}$ values.

5 Why One Moment Pays for Every Scale

Set $\kappa = \min\{k, 3\}$, so $\mathbb{E}|X - \mu|^\kappa \leq \sigma^\kappa$. Choose the smallest L satisfying

$$P_L \geq C_\kappa \sigma \left(\frac{\sigma}{\epsilon}\right)^{1/(\kappa-1)}. \tag{9}$$

Lemma 2 (Top-scale alias). *On the localization event,*

$$|\mu - \{A + \mathbb{E}R_{P_L,s_L}(X)\}| \leq \frac{4^\kappa \sigma^\kappa}{P_L^{\kappa-1}}.$$

The selected cell is $[A, A + P_L)$, and (3) places μ in its middle half. Unwrapping is exact for X in this cell. Outside it,

$$|X - \{A + R_{P_L,s_L}(X)\}| \leq 4|X - \mu| \mathbf{1}\{|X - \mu| \geq P_L/4\}.$$

Taking expectations and applying the κ-moment bound proves the lemma. The choice (9) makes this bias at most $\epsilon/4$.

For $r = \sigma/\epsilon$, allocate

$$n_0 = \lceil Cr^2 \rceil, \qquad n_i = \left\lceil Cr^2 2^{i(2-\kappa)} \right\rceil. \tag{10}$$

Equations (4) and (7) give

$$\mathrm{Var}(\widetilde{\mu} \mid c) \leq C' \epsilon^2 \left[1 + \sum_{i=0}^{L-1} 2^{i\kappa} \mathbb{P}\{|X - \mu| \geq 150\sigma 2^i\}\right]. \tag{11}$$

The apparent sum over scales costs only one moment. For every $z \in \mathbb{R}$,

$$\sum_{i:\, 150\sigma 2^i \leq |z|} 2^{i\kappa} \leq C_\kappa \left(\frac{|z|}{\sigma}\right)^\kappa. \tag{12}$$

Taking expectation with $z = X - \mu$ bounds the bracket in (11) by a κ-dependent constant. Increasing C in (10) makes one repetition accurate to $\epsilon/4$ around its expectation with probability at least $3/4$.

Condition on the complete localization transcript and its success. The selected dictionary path is now fixed, while refinement repetitions remain independent. A median of $O(\log(1/\delta))$ repetitions has sampling error at most $\epsilon/4$ with conditional probability at least $1-\delta/2$. Lemma 2 and the localization failure probability finish the accuracy proof.

It remains to count queries. Candidate shifts, pairs, and levels contribute only absolute constants. One repetition uses

$$O\left(r^2 + \sum_{i=0}^{L-1} r^2 2^{i(2-\kappa)} + L\right) \tag{13}$$

queries. Since $2^L = \Theta_\kappa(r^{1/(\kappa-1)})$, expression (13) is $O_k(r^{k/(k-1)})$ for $1 < k < 2$, $O(r^2 \log r)$ for $k = 2$, and $O_k(r^2)$ for $k > 2$. Localization's $O(\log(1/\delta))$ term is absorbed by refinement. This proves Theorem 1.

6 Discussion

The result removes interaction only for arbitrary measurable quantizers. Threshold and interval lower bounds remain intact (Lau and Scarlett, 2026b). Our level sets are periodic unions of intervals and deliberately violate spatial locality.

The theorem is also an order statement. Parametric work proves that nonadaptive one-bit protocols can lose exact asymptotic efficiency for symmetric log-concave and location-scale families (Kipnis and Duchi, 2022; Kumar and Vatedka, 2026). Such constant gaps are compatible with the rate equality proved here.

The construction answers the scalar finite-moment question, but leaves several directions open. The localization codebook is sample optimal but not designed for fastest decoding over a very large search range. The constant factor from prequerying every shift pair is also conservative. More broadly, it remains unknown when decoder-side selection from a fixed query dictionary can replace interaction in multivariate and structured estimation.

References

Abdalla, P. and Chen, J. (2026). Robust mean estimation under quantization. *arXiv preprint arXiv:2601.07074*.

Boufounos, P. T. (2012). Universal rate-efficient scalar quantization. *IEEE Transactions on Information Theory*, 58(3):1861–1872.

Cai, T. T. and Wei, H. (2024). Distributed gaussian mean estimation under communication constraints: Optimal rates and communication-efficient algorithms. *Journal of Machine Learning Research*, 25(37):1–63.

Chen, J., Wang, C.-L., Ng, M. K., and Wang, D. (2023). High dimensional statistical estimation under uniformly dithered one-bit quantization. *IEEE Transactions on Information Theory*, 69(8):5151–5187.

Kipnis, A. and Duchi, J. C. (2022). Mean estimation from one-bit measurements. *IEEE Transactions on Information Theory*, 68(9):6276–6296.

Kumar, R. and Vatedka, S. (2026). One-bit distributed mean estimation with unknown variance. *Transactions on Machine Learning Research*.

Lau, I. and Scarlett, J. (2026a). Open problem: Is interaction necessary for order-optimal 1-bit mean estimation? *arXiv preprint arXiv:2607.02896*.

Lau, I. and Scarlett, J. (2026b). Order-optimal sequential 1-bit mean estimation in general tail regimes. *arXiv preprint arXiv:2604.07796*.

Lau, I. and Scarlett, J. (2026c). Sequential 1-bit mean estimation with near-optimal sample complexity. In *International Conference on Artificial Intelligence and Statistics*.

Mayekar, P., Suresh, A. T., and Tyagi, H. (2021). Wyner–ziv estimators: Efficient distributed mean estimation with side-information. In *International Conference on Artificial Intelligence and Statistics*, volume 130, pages 3502–3510. PMLR.

This supplement gives a complete specification and proof. All logarithms are natural. Constants may depend on the fixed moment exponent k, but never on $\lambda, \sigma, \epsilon$, or δ.

A Protocol

Let $D \in \mathcal{D}(k, \lambda, \sigma)$ have mean μ. Put

$$\kappa = \min\{k, 3\}, \qquad r = \frac{\sigma}{\epsilon}.$$

Lyapunov's inequality gives

$$\mathbb{E}|X - \mu|^\kappa \le \sigma^\kappa. \tag{14}$$

Localization. Apply the deterministic nonadaptive localization protocol in Theorem 16 of Lau and Scarlett (2026) with failure probability $\delta/2$. It uses

$$n_{\text{loc}} \le C \left(\log \frac{\lambda}{\sigma} + \log \frac{2}{\delta} + 1 \right)$$

fresh samples. Except on an event of probability at most $\delta/2$, it returns an interval I containing μ with length at most 100σ. The value 100 follows directly from their proof: the returned interval is the union of at most five bins, each of width at most 20σ.

Let $\mathcal{T}_{\text{loc}}$ be the complete localization transcript and let c be the midpoint of I. Define

$$\mathcal{E}_{\text{loc}} = \{\mu \in I, \ |I| \le 100\sigma\}.$$

On this event,

$$|c - \mu| \le 50\sigma. \tag{15}$$

Scales. Set

$$P_0 = 600\sigma, \quad P_i = 2^i P_0, \quad \mathcal{S}_P = \{0, P/3, 2P/3\}.$$

Let L be the smallest nonnegative integer such that

$$P_L \ge C_{\text{top},\kappa} \sigma r^{1/(\kappa-1)}. \tag{16}$$

The decoder selects

$$s_i \in \arg\max_{s \in \mathcal{S}_{P_i}} \operatorname{dist}(c, s + P_i\mathbb{Z}), \qquad 0 \le i \le L, \tag{17}$$

with a deterministic tie rule.

Base blocks. Define

$$R_{P,s}(x) = x - s - P \left\lfloor \frac{x - s}{P} \right\rfloor.$$

For every $s \in \mathcal{S}_{P_0}$, allocate n_0 fresh samples. Before any data are observed, draw independent public $U_{s,j} \sim \text{Unif}[0, P_0]$ and query

$$Q_{s,j}(x) = \mathbf{1}\{R_{P_0,s}(x) \ge U_{s,j}\}.$$

For the selected shift, define

$$\widehat{B} = \frac{P_0}{n_0} \sum_{j=1}^{n_0} Q_{s_0,j}(X_{s_0,j}).$$

Increment blocks. For every $0 \le i < L$, every $(s,t) \in \mathcal{S}_{P_i} \times \mathcal{S}_{P_{i+1}}$, and every $\ell \in \{-1, 0, 1, 2, 3\}$, allocate n_i fresh samples and query

$$Q_{i,s,t,\ell}(x) = \mathbf{1}\{J_{P_i,s,t}(x) = \ell\},$$

where

$$J_{P,s,t}(x) = \left\lfloor \frac{x-s}{P} \right\rfloor - 2\left\lfloor \frac{x-t}{2P} \right\rfloor.$$

Let $\widehat{p}_{i,s,t,\ell}$ be the empirical average in this block. For the selected pair, set $\ell_i^0 = J_{P_i,s_i,s_{i+1}}(c)$ and

$$\begin{aligned} \widehat{D}_i &= s_i - s_{i+1} + P_i \widehat{J}_i, \\ \widehat{J}_i &= \ell_i^0 + \sum_{\ell \ne \ell_i^0} (\ell - \ell_i^0) \widehat{p}_{i,s_i,s_{i+1},\ell}. \end{aligned} \tag{18}$$

Unwrapping and amplification. Define

$$A_L = s_L + P_L \left\lfloor \frac{c - s_L}{P_L} \right\rfloor$$

and

$$\widetilde{\mu} = A_L + \widehat{B} + \sum_{i=0}^{L-1} \widehat{D}_i. \tag{19}$$

Use $K = \lceil C_{\mathrm{med}} \log(2/\delta) \rceil$ independent copies of every refinement block and return the median of their outputs. Every localization and refinement query, scale, candidate shift, level, sample assignment, and public random variable is fixed before any response is observed. The localization transcript affects only decoder arithmetic.

B Geometry and Algebra

Lemma 3 (Three-shift padding)**.** *For every $P > 0$ and $c \in \mathbb{R}$,*

$$\max_{s \in \mathcal{S}_P} \operatorname{dist}(c, s + P\mathbb{Z}) \ge P/3.$$

On $\mathcal{E}_{\mathrm{loc}}$, every selected shift satisfies

$$\operatorname{dist}(\mu, s_i + P_i\mathbb{Z}) \ge P_i/4.$$

Proof. Modulo P, the three boundary points are equally spaced, so the farthest is at circular distance at least $P/3$. Equation (15) gives $|c-\mu| \le 50\sigma = P_0/12 \le P_i/12$. The triangle inequality yields $P_i/3 - P_i/12 = P_i/4$. $\square$

Lemma 4 (Finite adjacent alphabet)**.** *For $s \in \mathcal{S}_P$ and $t \in \mathcal{S}_{2P}$,*

$$R_{2P,t}(x) - R_{P,s}(x) = s - t + P J_{P,s,t}(x). \tag{20}$$

The function J is $2P$-periodic and takes at most three values. Writing $a = s/P$ and $b = t/P$, its exact ranges are

	$b = 0$	$b = 2/3$	$b = 4/3$
$a = 0$	$\{0, 1\}$	$\{0, 1, 2\}$	$\{1, 2, 3\}$
$a = 1/3$	$\{-1, 0, 1\}$	$\{0, 1, 2\}$	$\{1, 2\}$
$a = 2/3$	$\{-1, 0, 1\}$	$\{0, 1\}$	$\{0, 1, 2\}$.

Proof. Expansion proves (20). Adding $2P$ to x increases the first floor by two and the second by one, proving periodicity. Divide by P, restrict to $[0, 2)$, and evaluate the floors on the breakpoints $\{a, a+1, b\}$ modulo two to obtain the table. $\square$

Lemma 5 (Local constancy)**.** *On $\mathcal{E}_{\mathrm{loc}}$,*

$$J_{P_i,s_i,s_{i+1}}(X) \ne J_{P_i,s_i,s_{i+1}}(c) \implies |X - \mu| \ge P_i/4.$$

Proof. The discontinuities of $J_{P,s,t}$ lie in $(s + P\mathbb{Z}) \cup (t + 2P\mathbb{Z})$. Lemma 3 places the first lattice at distance at least $P/4$ from μ and the second at distance at least $P/2$. The points c and μ lie in the same component of their complement because $|c-\mu| \le P/12$. If $J(X) \ne J(c)$, the segment from c to X crosses a discontinuity at distance at least $P/4$ from μ. The claim follows. $\square$

Lemma 6 (Telescoping and top cell)**.** *For arbitrary shifts,*

$$R_{P_L,s_L}(x) = R_{P_0,s_0}(x) + \sum_{i=0}^{L-1} \{R_{P_{i+1},s_{i+1}}(x) - R_{P_i,s_i}(x)\}.$$

On $\mathcal{E}_{\mathrm{loc}}$, $\mu \in [A_L + P_L/4, A_L + 3P_L/4]$.

Proof. The first claim telescopes. The definition of A_L and selected padding give $c \in [A_L + P_L/3, A_L + 2P_L/3]$. Now use $|c - \mu| \le P_L/12$. $\square$

C Bias and Variance

Lemma 7 (Conditional unbiasedness)**.** *Conditional on any localization transcript,*

$$\mathbb{E}[\widetilde{\mu} \mid \mathcal{T}_{\mathrm{loc}}] = A_L + \mathbb{E}R_{P_L,s_L}(X).$$

Proof. For fixed x,

$$\mathbb{P}_U\{U \le R_{P_0,s}(x)\} = R_{P_0,s}(x)/P_0,$$

so the selected base estimator is unbiased. For a selected increment, with $p_\ell = \mathbb{P}\{J(X) = \ell\}$ and $\ell_0 = J(c)$,

$$\mathbb{E}J(X) = \ell_0 + \sum_{\ell \ne \ell_0} (\ell - \ell_0)p_\ell.$$

Thus (18) and (20) give an unbiased increment estimator. Refinement blocks are independent of $\mathcal{T}_{\mathrm{loc}}$. Lemma 6 finishes the proof. $\square$

Lemma 8 (Top-scale alias)**.** *On $\mathcal{E}_{\mathrm{loc}}$,*

$$|\mu - \{A_L + \mathbb{E}R_{P_L,s_L}(X)\}| \le \frac{4^\kappa \sigma^\kappa}{P_L^{\kappa-1}}.$$

Proof. Write $P = P_L$, $s = s_L$, and $A = A_L$. Unwrapping is exact for $x \in [A, A + P)$. Since μ lies in the middle half of this cell, for all $x \in \mathbb{R}$ the following bound holds:

$$|x - \{A + R_{P,s}(x)\}| \le 4|x - \mu|\mathbf{1}\{|x - \mu| \ge P/4\}.$$

Indeed, write $x \in [A + mP, A + (m + 1)P)$ for $m \in \mathbb{Z}$. The left side is $|m|P$. It vanishes for $m = 0$. If $m \ge 1$, then $x - \mu \ge (m - 3/4)P \ge mP/4$, and if $m \le -1$, then $\mu - x \ge (|m| - 1/4)P \ge |m|P/4$. In either nonzero case, $|x - \mu| \ge P/4$ as well. Taking expectations and using (14) gives

$$\begin{aligned} |\mu - \{A + \mathbb{E}R_{P,s}(X)\}| &\le 4\mathbb{E}[|X - \mu|\mathbf{1}\{|X - \mu| \ge P/4\}] \\ &\le 4\sigma^\kappa (P/4)^{1-\kappa}, \end{aligned}$$

which is the claim. $\square$

Lemma 9 (Tail-sensitive variance)**.** *Conditional on a successful localization transcript,*

$$\mathrm{Var}(\widehat{D}_i \mid \mathcal{T}_{\mathrm{loc}}) \le \frac{64P_i^2}{n_i}\mathbb{P}\{|X - \mu| \ge P_i/4\}.$$

Proof. The nonbaseline probabilities in (18) use independent blocks. Lemma 4 gives $|\ell - \ell_i^0| \le 4$. Therefore

$$\begin{aligned}\operatorname{Var}(\widehat{D}_i \mid \mathcal{T}_{\text{loc}}) &\le \frac{16P_i^2}{n_i} \sum_{\ell \neq \ell_i^0} p_\ell(1-p_\ell) \\ &\le \frac{64P_i^2}{n_i} \mathbb{P}\{J(X) \neq \ell_i^0\}.\end{aligned}$$

Apply Lemma 5. □

Lemma 10 (One-moment tail sum)**.** *For every $\kappa > 0$ and $a > 0$,*

$$\sum_{i=0}^{\infty} 2^{i\kappa} \mathbb{P}\{|X-\mu| \ge a\sigma 2^i\} \le \frac{2^\kappa}{2^\kappa - 1} a^{-\kappa} \frac{\mathbb{E}|X-\mu|^\kappa}{\sigma^\kappa}.$$

Proof. By Tonelli, the left side equals

$$\mathbb{E}\left[\sum_{i:\, a\sigma 2^i \le |X-\mu|} 2^{i\kappa}\right].$$

The inner geometric sum is at most

$$\frac{2^\kappa}{2^\kappa - 1}\left(\frac{|X-\mu|}{a\sigma}\right)^\kappa.$$

Take expectations. □

D Proof of the Main Theorem

Proof. Choose $C_{\text{top},\kappa}$ in (16) so that Lemma 8 is at most $\epsilon/4$. Allocate

$$n_0 = \lceil C_{\text{var},\kappa} r^2 \rceil, \qquad n_i = \left\lceil C_{\text{var},\kappa} r^2 2^{i(2-\kappa)} \right\rceil.$$

The selected base estimate is a rescaled Bernoulli average, hence

$$\operatorname{Var}(\widehat{B} \mid \mathcal{T}_{\text{loc}}) \le \frac{P_0^2}{4n_0} \le \frac{600^2}{4C_{\text{var},\kappa}} \epsilon^2.$$

All selected component blocks remain independent after conditioning on $\mathcal{T}_{\text{loc}}$. Lemma 9 gives

$$\begin{aligned}\sum_{i=0}^{L-1} \operatorname{Var}(\widehat{D}_i \mid \mathcal{T}_{\text{loc}}) &\le \frac{C_\kappa \epsilon^2}{C_{\text{var},\kappa}} \sum_{i=0}^{L-1} 2^{i\kappa} \mathbb{P}\{|X-\mu| \ge 150\sigma 2^i\} \\ &\le \frac{C'_\kappa}{C_{\text{var},\kappa}} \epsilon^2,\end{aligned}$$

where the last line uses Lemma 10. Choose $C_{\text{var},\kappa}$ so that $\operatorname{Var}(\widetilde{\mu} \mid \mathcal{T}_{\text{loc}}) \le \epsilon^2/64$. Chebyshev's inequality makes one repetition accurate to $\epsilon/4$ around its conditional mean with probability at least $3/4$.

Conditional on the complete successful localization transcript, the dictionary path is fixed and the K repetitions remain independent. A binomial tail bound makes their median accurate to $\epsilon/4$ around its conditional mean with probability at least $1-\delta/2$. The alias bias is at most $\epsilon/4$. Adding the localization failure probability proves the accuracy statement with room in the constants.

There are three base candidates, nine shift pairs per scale, and five levels per pair. One repetition therefore costs

$$C_\kappa\left[r^2 + \sum_{i=0}^{L-1} r^2 2^{i(2-\kappa)} + L\right].$$

Minimality in (16) gives $2^L = \Theta_\kappa(r^{1/(\kappa-1)})$ after choosing the theorem's constant c_k sufficiently small. For $1 < k < 2$, the geometric sum is $\Theta_k(r^{k/(k-1)})$. For $k = 2$, it is $O(r^2 \log r)$. For $k > 2$, it is $O_k(r^2)$. The ceiling contribution L is absorbed.

Multiplication by $K = O(\log(1/\delta))$ gives the claimed refinement rates. The localization term $\log(1/\delta)$ is absorbed by refinement, while the remaining localization cost is $O(\log(\lambda/\sigma))$ up to an additive constant. This completes the upper bound. The matching lower bound is Theorem 9 of Lau and Scarlett (2026), which allows arbitrary adaptive quantizers and therefore applies to the nonadaptive subclass for the high-confidence range $\delta < \delta_0$ stated there. $\square$

E Scope of the Lower Bound

The protocol uses measurable periodic unions of intervals. It does not use threshold or single-interval queries. Existing linear-in-λ/σ lower bounds for nonadaptive threshold and interval queries therefore remain intact.