

A Positive Dispersion Optimum for the Rounded Sample Mean: Sufficiency and Necessity of Finite Variance

Mark R. Diamond 

Angela O'Brien-Malone 

October 14, 2026

Abstract

A continuous quantity, observable only after nearest-lattice rounding, is estimated by the mean of n rounded readings of an imperfect instrument. Averaged over the position of the true value within a rounding cell, the mean-squared error of this estimator is a function of one dimensionless parameter: the pre-rounding standard deviation s , measured in cell widths. A perfectly repeatable instrument ($s = 0$) returns the same rounded value on every reading, so averaging provides no additional information; its cell-averaged error is the rounding floor $1/12$. We show that this floor is never optimal once more than one reading is taken. For *any* centred pre-rounding law with finite, non-zero variance and any sample size $n > 1$, some strictly positive dispersion yields a smaller cell-averaged error than zero, so the global minimizer lies away from the no-scatter endpoint; and each of the three hypotheses— $n > 1$, finite variance, non-degeneracy—is individually necessary, so existence is *exactly* characterized. The mechanism is elementary: the autocorrelation of the rounding error has a *corner* at zero lag, so a little scatter relieves the rounding bias at first order in s while its variance contribution grows only at second order. We give the short corner-based proof and then derive the spectral form of the criterion, using it to compute the optimum for four laws. The Gaussian optimizer is unique and solves a Jacobi-theta equation. For the Laplace law the bias term has a hyperbolic closed form, and the optimizer is unique and characterized by the corresponding derivative equation. For the uniform and two-point laws the optimizer is explicit. The uniform law cancels every bias harmonic exactly at its first positive cancellation scale, which saturates the natural localization bound when $n = 2$; and the two-point law, whose criterion has several local wells once $n \geq 4$, shows that finite variance ensures existence but not a single-well objective.

Keywords: quantization error, rounded sample mean, rounding error, characteristic functions, mean-squared error, Fourier series.

2020 Mathematics Subject Classification: Primary 62F10; Secondary 94A12, 60E10, 42A16.

1 Introduction

From ancient astronomical practice onward, observers have sought both better instruments and better methods for identifying, reducing, bounding, and combining observational errors [8, 9]. The advantage of averaging was itself still contested in the eighteenth century: Simpson gave a probabilistic defence of the mean of repeated astronomical observations against the claim that a single observation taken with due care was equally reliable [10]. This history supports a natural limiting intuition: other things equal, greater precision and greater repeatability should improve a measurement. The question motivating this paper is whether

*Corresponding author: mark.diamond@yahoo.co.uk

that intuition remains valid when repeatability is considered together with finite numerical resolution. If all random unrepeatability could be removed, would a perfectly repeatable instrument necessarily be best?

Finite numerical reporting introduces a distinct limitation. An observation may concern a continuous quantity while being recorded only at fixed increments. Before this limitation was described in the language of quantization, closely related problems were treated through the graduation and interpolation of instrument scales, and through the rounding or grouping of recorded data [1, 2]. Sheppard’s analysis of data arranged according to equidistant divisions of a scale is an early statistical example [7]. In the present paper, finite resolution is modelled by nearest-lattice quantization: after rescaling, a continuous input is reported as the nearest point of a fixed lattice. This is a deliberately restricted model of the more general connection between finite-precision measurement and quantization [12, p. 1].

A perfectly repeatable instrument returns the same rounded value on every reading, so repeated observations cannot recover the unknown position of the true value within a rounding cell. An instrument with random unrepeatability may instead return neighbouring lattice values. We ask whether the average of n such rounded readings can have smaller cell-averaged squared error than the repeated output of the perfectly repeatable instrument.

Section 2 formulates the model and the cell-averaged criterion, and Section 4 reduces the existence question to the local behaviour of the rounding-error autocorrelation. The resulting theorem characterizes when a positive minimizing dispersion exists. The spectral representation in Section 6 is then used to study four representative scatter laws in Section 7. Section 8 explains the relation to dithered quantization, and Section 9 records the scope of the result and the remaining open questions.

The parameter s describes residual pre-rounding unrepeatability. It is not a deliberately injected dither or an independently controllable design variable. Accordingly, the optimum studied here is a property of the model and criterion, not a recommendation to add noise to an instrument.

2 Model and the cell-averaged criterion

All lengths are measured in cell widths, so the cell width is 1. We use the nearest-lattice rounding map

$$Q(x) = \lfloor x + \tfrac{1}{2} \rfloor, \quad (2.1)$$

which rounds half-integers upward. Any other fixed convention at half-integers would give the same cell-averaged quantities considered below. Let Z be a real random variable with

$$\mathbb{E}[Z] = 0, \quad \mathbb{E}[Z^2] = 1, \quad (2.2)$$

and let $Z_1, \dots, Z_n$ be independent copies of Z . For an unknown sub-cell position $u \in [-\frac{1}{2}, \frac{1}{2}]$ and a dispersion $s \geq 0$, put

$$Y_j(u, s) = Q(u + sZ_j), \quad \bar{Y}_n(u, s) = \frac{1}{n} \sum_{j=1}^n Y_j(u, s). \quad (2.3)$$

The criterion is the squared error of the rounded sample mean about the underlying position u , averaged over one representative cell:

$$\text{MSE}_n(s) = \int_{-1/2}^{1/2} \mathbb{E}[(\bar{Y}_n(u, s) - u)^2] \, du. \quad (2.4)$$

For a physical cell width h the model corresponds to observing $h Q(\eta/h + sZ)$ with $u = \eta/h$ taken modulo one; the dimensional cell-averaged error is then $h^2 \text{MSE}_n(s)$, and $s = \sigma/h$ is

the pre-rounding standard deviation expressed in cell widths. The magnitudes of σ and h thus never enter separately: the cell width contributes only the overall factor h^2 , so the shape of the cell-averaged error, and with it the location of its minimum, is governed by their ratio s alone.

It is convenient to separate quantization (*rounding*) from the identity. The periodic quantization error

$$q_\varepsilon(x) = x - Q(x) \quad (2.5)$$

equals x on $(-\frac{1}{2}, \frac{1}{2})$ and is extended with period one. Values at half-integers depend on the chosen tie convention, but all identities below are cell-integral identities and are therefore unaffected by changes on this null set. Thus, up to null sets, q_ε is an odd sawtooth with

$$\int_{-1/2}^{1/2} q_\varepsilon(u) du = 0, \quad \int_{-1/2}^{1/2} q_\varepsilon(u)^2 du = \frac{1}{12}, \quad (2.6)$$

and a single reading decomposes as

$$Y_j(u, s) - u = sZ_j - q_\varepsilon(u + sZ_j). \quad (2.7)$$

Throughout, $\varphi(t) = \mathbb{E}[e^{itZ}]$ denotes the characteristic function of Z , Z' an independent copy of Z , and $W = Z - Z'$ the symmetrized noise.

3 Variance cost *versus* bias reduction: the decomposition

The variance–bias split of the sample mean turns (2.4) into a contest between two cell averages.

Proposition 3.1 (Variance cost / bias-reduction decomposition). *Under (2.2), for every integer $n \geq 1$ and every $s \geq 0$,*

$$\text{MSE}_n(s) = \frac{1}{n} \left(s^2 + \frac{1}{12} \right) + \frac{n-1}{n} A(s), \quad A(s) := \int_{-1/2}^{1/2} b_s(u)^2 du, \quad b_s(u) := \mathbb{E}[q_\varepsilon(u + sZ)]. \quad (3.1)$$

Proof. Write $\Delta(u, s) = Y_1(u, s) - u$. For independent identically distributed readings the mean-squared error of the average splits as

$$\mathbb{E}[(\bar{Y}_n(u, s) - u)^2] = \frac{1}{n} \mathbb{E}[\Delta(u, s)^2] + \frac{n-1}{n} \mathbb{E}[\Delta(u, s)]^2. \quad (3.2)$$

We average the two terms over the cell separately. For the first, fix a realization z . By (2.7), the period-one shift-invariance of $\int_{-1/2}^{1/2} q_\varepsilon(\cdot + c)$, and (2.6),

$$\int_{-1/2}^{1/2} \{sz - q_\varepsilon(u + sz)\}^2 du = s^2 z^2 - 2sz \int_{-1/2}^{1/2} q_\varepsilon(u + sz) du + \int_{-1/2}^{1/2} q_\varepsilon(u + sz)^2 du = s^2 z^2 + \frac{1}{12}.$$

Taking the expectation over Z and using $\mathbb{E}[Z^2] = 1$ gives $\int_{-1/2}^{1/2} \mathbb{E}[\Delta(u, s)^2] du = s^2 + \frac{1}{12}$. For the second, since $\mathbb{E}[Z] = 0$ the linear part of (2.7) averages to zero, so $\mathbb{E}[\Delta(u, s)] = -\mathbb{E}[q_\varepsilon(u + sZ)] = -b_s(u)$, and the cell average of its square is $A(s)$ by definition. Substituting both into the cell average of (3.2) gives (3.1). $\square$

In (3.1), the s^2/n part is the *variance cost* of scatter; at $s = 0$ the first summand is only $1/(12n)$, and the full rounding floor $1/12$ is recovered only after adding the $A(0)$ contribution. The function A is the cell-averaged squared rounding bias; it carries all the noise-dependence of the bias-reduction. The two facts we need about it are immediate. First, q_ε has mean zero on each cell, so $b_0(u) = q_\varepsilon(u)$ and, by (2.6),

$$A(0) = \int_{-1/2}^{1/2} q_\varepsilon(u)^2 du = \frac{1}{12}, \quad \text{whence} \quad \text{MSE}_n(0) = \frac{1}{12}. \quad (3.3)$$

The value $\frac{1}{12}$ is the classical second moment of the rounding error over a cell [7]. Second, $|b_s(u)| \leq \frac{1}{2}$ always, so $A(s) \leq \frac{1}{4}$ and, as we show below, $A(s) \leq A(0)$: zero scatter maximizes the cell-averaged squared bias, and dispersion cannot increase it. Subtracting (3.3) from (3.1) isolates the comparison,

$$\text{MSE}_n(s) - \frac{1}{12} = \frac{s^2}{n} - \frac{n-1}{n} (A(0) - A(s)), \quad (3.4)$$

in which a positive dispersion improves on the no-scatter endpoint exactly when the bias reduction $A(0) - A(s)$, weighted by $(n-1)/n$, exceeds the variance cost s^2/n . The remaining question is the rate at which $A(0) - A(s)$ grows as s leaves zero. The next section answers this without any distributional assumption.

4 Existence via the corner

The reduction in the bias is governed by a single elementary function, the autocorrelation of the rounding error, and by one feature of it: a corner.

4.1 The squared bias as an autocorrelation averaged over a random lag

Lemma 4.1 (Autocorrelation form of the bias). *Let*

$$g(\delta) := \int_{-1/2}^{1/2} q_\varepsilon(v) q_\varepsilon(v + \delta) dv. \quad (4.1)$$

Then for every $s \geq 0$,

$$A(s) = \mathbb{E}[g(sW)], \quad W = Z - Z'. \quad (4.2)$$

Proof. Using $b_s(u) = \mathbb{E}[q_\varepsilon(u + sZ)]$ and an independent copy Z' to write the square as a product of two expectations,

$$b_s(u)^2 = \mathbb{E}[q_\varepsilon(u + sZ)] \mathbb{E}[q_\varepsilon(u + sZ')] = \mathbb{E}[q_\varepsilon(u + sZ) q_\varepsilon(u + sZ')],$$

the last step by independence. The integrand is bounded by $\frac{1}{4}$ on a bounded interval, so Fubini's theorem moves the cell integral inside the expectation:

$$A(s) = \mathbb{E} \left[\int_{-1/2}^{1/2} q_\varepsilon(u + sZ) q_\varepsilon(u + sZ') du \right].$$

For fixed realizations z, z' , substitute $v = u + sz$ and use the period-one shift-invariance of the integral of a period-one integrand; the inner integral becomes $\int_{-1/2}^{1/2} q_\varepsilon(v) q_\varepsilon(v + s(z' - z)) dv = g(s(z' - z))$, which depends on z, z' only through the lag. Since g is even (substitute $v \mapsto -v$ and use $q_\varepsilon(-x) = -q_\varepsilon(x)$) and $Z - Z'$ is symmetric, replacing $Z' - Z$ by $W = Z - Z'$ leaves the expectation unchanged, giving (4.2). $\square$

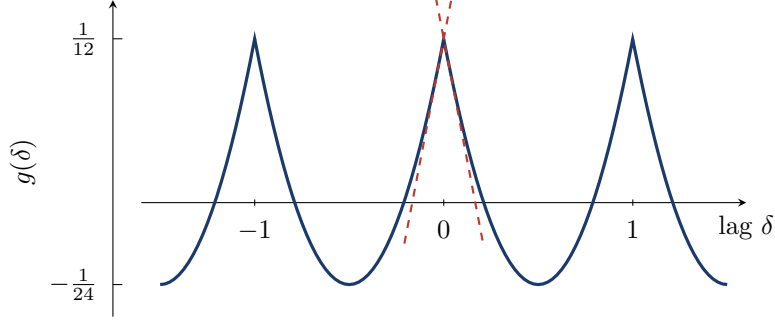


Figure 1: The rounding-error autocorrelation $g(\delta)$ of (4.3), period one with maximum $\frac{1}{12}$ and minimum $-\frac{1}{24}$. At each integer lag g has a downward *corner*; the red dashed lines are the one-sided tangents of slopes $+\frac{1}{2}$ (left) and $-\frac{1}{2}$ (right) at the peak $(0, \frac{1}{12})$. The vertical scale is greatly exaggerated.

4.2 The autocorrelation and its corner

Lemma 4.2 (Closed form and corner). *With $\langle \delta \rangle := \text{dist}(\delta, \mathbb{Z}) \in [0, \frac{1}{2}]$ the distance from δ to the nearest integer,*

$$g(\delta) = \frac{1}{12} - \frac{1}{2}\langle \delta \rangle + \frac{1}{2}\langle \delta \rangle^2, \quad \text{equivalently} \quad \frac{1}{12} - g(\delta) = \frac{1}{2}\langle \delta \rangle(1 - \langle \delta \rangle). \quad (4.3)$$

In particular g is continuous, has period one, attains its maximum $g(0) = \frac{1}{12}$ at the integers, and has the one-sided slopes

$$g'(0^+) = -\frac{1}{2}, \quad g'(0^-) = +\frac{1}{2}. \quad (4.4)$$

Proof. By evenness and periodicity it suffices to compute g on $[0, \frac{1}{2}]$. Fix $\delta \in [0, \frac{1}{2}]$. On the window $v \in [-\frac{1}{2}, \frac{1}{2})$ the first factor is $q_\varepsilon(v) = v$. The shifted argument $v + \delta$ crosses the half-integer discontinuity at $v = \frac{1}{2} - \delta \in [0, \frac{1}{2}]$, where $q_\varepsilon(v + \delta)$ drops by one: it equals $v + \delta$ for $v < \frac{1}{2} - \delta$ and $v + \delta - 1$ for $v \geq \frac{1}{2} - \delta$. Splitting (4.1) at this point,

$$g(\delta) = \int_{-1/2}^{1/2-\delta} v(v + \delta) dv + \int_{1/2-\delta}^{1/2} v(v + \delta - 1) dv = \frac{1}{2}\left(\frac{1}{2} - \delta\right)^2 - \frac{1}{24} = \frac{1}{12} - \frac{\delta}{2} + \frac{\delta^2}{2},$$

the cubic endpoint terms cancelling at the split point. On $[0, \frac{1}{2}]$ the nearest integer is 0, so $\langle \delta \rangle = \delta$ and this is (4.3); evenness and periodicity carry it to all δ through $\langle \delta \rangle$. On $(0, \frac{1}{2})$, $g(\delta) = \frac{1}{12} - \frac{\delta}{2} + \frac{\delta^2}{2}$ has $g'(\delta) = -\frac{1}{2} + \delta$; letting $\delta \downarrow 0$ gives $g'(0^+) = -\frac{1}{2}$, and evenness gives $g'(0^-) = +\frac{1}{2}$. $\square$

The parabolic term $\frac{1}{2}\langle \delta \rangle^2$ is flat at the origin, so all of the behaviour of g near a peak comes from the term $-\frac{1}{2}\langle \delta \rangle$: an inverted wedge with slopes $\pm\frac{1}{2}$ (Figure 1). A smooth peak could not do this. By Taylor's theorem a differentiable function with a vanishing derivative at its maximum falls only quadratically; the corner falls linearly. That gap—linear bias reduction against quadratic variance cost—is the entire content of the existence theorem.

4.3 Linear relief and the existence theorem

Lemma 4.3 (First-order reduction in the bias). *If $\mathbb{E}[Z^2] < \infty$ then $\mathbb{E}|W| < \infty$ and*

$$\frac{A(0) - A(s)}{s} \longrightarrow \frac{1}{2} \mathbb{E}|W| = \frac{1}{2} \mathbb{E}|Z - Z'| \quad (s \downarrow 0). \quad (4.5)$$

Proof. By (4.2) and $A(0) = \mathbb{E}[g(0)] = \frac{1}{12}$,

$$\frac{A(0) - A(s)}{s} = \mathbb{E} \left[\frac{\frac{1}{12} - g(sW)}{s} \right] = \mathbb{E} \left[\frac{\frac{1}{2} \langle sW \rangle (1 - \langle sW \rangle)}{s} \right],$$

using (4.3). For fixed w , once $s|w| < \frac{1}{2}$ we have $\langle sw \rangle = s|w|$, so the bracket equals $\frac{1}{2}|w|(1 - s|w|) \rightarrow \frac{1}{2}|w|$ as $s \downarrow 0$. Moreover $\langle x \rangle(1 - \langle x \rangle) \leq \langle x \rangle \leq |x|$ gives the uniform domination $0 \leq \frac{1}{2}\langle sw \rangle(1 - \langle sw \rangle)/s \leq \frac{1}{2}|w|$, an integrable bound since $\mathbb{E}|W| \leq 2\mathbb{E}|Z| < \infty$. Dominated convergence yields (4.5). $\square$

Theorem 4.4 (A positive optimum under finite variance). *Let Z satisfy (2.2) and let $n > 1$. Then MSE_n attains its global minimum on $[0, \infty)$. Let s_n^* denote any global minimizer. Then $s_n^* > 0$ and*

$$\boxed{\text{MSE}_n(s_n^*) < \text{MSE}_n(0) = \frac{1}{12}.} \quad (4.6)$$

Proof. The autocorrelation g is continuous and bounded, so by bounded convergence $A(s) = \mathbb{E}[g(sW)]$ is continuous, and hence so is MSE_n . From (3.1) and $A \geq 0$, $\text{MSE}_n(s) \geq s^2/n \rightarrow \infty$, so the continuous MSE_n attains a global minimum on $[0, \infty)$.

Combining the contest (3.4) with the rate at which the bias is reduced, (4.5),

$$\text{MSE}_n(s) - \frac{1}{12} = \frac{s^2}{n} - \frac{n-1}{2n} \mathbb{E}|W| s + o(s) \quad (s \downarrow 0). \quad (4.7)$$

Because $\mathbb{E}[Z^2] = 1$ the law is non-degenerate, so $\mathbb{E}|W| = \mathbb{E}|Z - Z'| > 0$; the negative linear term dominates the positive quadratic one, and the right-hand side is negative for all sufficiently small $s > 0$. By (3.4) every such s satisfies $\text{MSE}_n(s) < \text{MSE}_n(0)$. The endpoint $s = 0$ is therefore not a minimizer, so every global minimizer lies in $(0, \infty)$ and takes a value below $1/12$. $\square$

Remark 4.5 (Uniqueness is separate from existence). Uniqueness is not asserted in Theorem 4.4. It is established for the two-point and uniform laws in §7.1 and §7.4, and for Gaussian scale mixtures—including the Gaussian and Laplace laws—in Proposition 7.2.

The proof drew on finite variance twice—for coercivity and, through $\mathbb{E}|W| < \infty$, for the bias-reduction rate—and on non-degeneracy once. With $n > 1$ these three are also necessary, as we now show; existence is then characterized exactly.

5 The characterization

Proposition 5.1 (The hypotheses are necessary). *Let Z be centred and $n \geq 1$. A strictly improving positive dispersion—a value $s > 0$ with $\text{MSE}_n(s) < \text{MSE}_n(0)$ —exists if and only if*

$$n > 1 \quad \text{and} \quad 0 < \text{Var } Z < \infty. \quad (5.1)$$

Proof. Sufficiency follows from Theorem 4.4 by scaling. If $0 < \text{Var } Z = \tau^2 < \infty$, put $\tilde{Z} = Z/\tau$. Theorem 4.4 applied to $\tilde{Z}$ gives a $t > 0$ such that $\text{MSE}_n^{\tilde{Z}}(t) < \text{MSE}_n^{\tilde{Z}}(0)$; since the original law at dispersion $s = t/\tau$ has the same rounded readings as the normalized law at dispersion t , this gives $\text{MSE}_n^Z(t/\tau) < \text{MSE}_n^Z(0)$. For necessity, suppose one condition fails.

If $n = 1$, there is no second reading to average: the bias-reduction term drops out and $\text{MSE}_1(s) = \int_{-1/2}^{1/2} \mathbb{E}[\Delta(u, s)^2] du$. The first-term calculation in the proof of Proposition 3.1

invokes the normalization (2.2) only at its closing step; carried out for a centred Z of arbitrary finite variance it gives $\int_{-1/2}^{1/2} \mathbb{E}[\Delta(u, s)^2] du = s^2 \mathbb{E}[Z^2] + \frac{1}{12}$, whence

$$\text{MSE}_1(s) = s^2 \text{Var } Z + \frac{1}{12} \quad (\text{Var } Z < \infty),$$

so no positive s strictly improves on $s = 0$; if $\text{Var } Z = \infty$, the infinite-variance case below shows that $\text{MSE}_1(s) = \infty$ for every $s > 0$.

If $\text{Var } Z = 0$, then, since Z is centred, $Z = 0$ almost surely. Hence $Y_j(u, s) = Q(u)$ for every j and every $s \geq 0$, so

$$\text{MSE}_n(s) = \int_{-1/2}^{1/2} (Q(u) - u)^2 du = \int_{-1/2}^{1/2} q_\varepsilon(u)^2 du = \frac{1}{12}$$

for all $s \geq 0$. Thus no positive dispersion strictly improves on the endpoint.

If $\text{Var } Z = \infty$ (equivalently $\mathbb{E}[Z^2] = \infty$), fix $s > 0$ and $u \in [-\frac{1}{2}, \frac{1}{2}]$, and write $\Delta_j(u, s) = Y_j(u, s) - u$. The pointwise inequality $(a + b)^2 \geq \frac{1}{2}a^2 - b^2$, applied with $a = sZ_j$, $b = -q_\varepsilon(u + sZ_j)$ and $|q_\varepsilon| \leq \frac{1}{2}$, gives

$$\Delta_j(u, s)^2 = (sZ_j - q_\varepsilon(u + sZ_j))^2 \geq \frac{1}{2}s^2 Z_j^2 - \frac{1}{4}, \quad \text{hence} \quad \mathbb{E}[\Delta_j(u, s)^2] \geq \frac{1}{2}s^2 \mathbb{E}[Z^2] - \frac{1}{4} = \infty.$$

Each $\Delta_j(u, s)$ is almost surely finite, since sZ_j is finite and $|q_\varepsilon| \leq \frac{1}{2}$; hence $C := \sum_{j=2}^n \Delta_j(u, s)$ is a.s. finite and independent of $\Delta_1(u, s)$. Applying the same inequality once more, now with $a = \Delta_1(u, s)$ and $b = C$, and conditioning on the event $C = c$, gives for every finite c

$$\mathbb{E}[(\Delta_1(u, s) + c)^2] \geq \frac{1}{2} \mathbb{E}[\Delta_1(u, s)^2] - c^2 = \infty.$$

Since C is a.s. finite and independent of $\Delta_1(u, s)$, it follows that

$$\mathbb{E}[(\Delta_1(u, s) + C)^2 \mid C] = \infty \quad \text{a.s.}$$

Since $\bar{Y}_n(u, s) - u = \frac{1}{n}(\Delta_1(u, s) + C)$, taking expectations gives

$$\mathbb{E}[(\bar{Y}_n(u, s) - u)^2] = \frac{1}{n^2} \mathbb{E}[(\Delta_1(u, s) + C)^2] = \infty$$

for every u and every $s > 0$. Therefore $\text{MSE}_n(s) = \infty$ for every $s > 0$, whereas $\text{MSE}_n(0) = \frac{1}{12}$; no positive value improves on the endpoint.

In each case no strictly improving positive dispersion exists. $\square$

Theorem 4.4 and Proposition 5.1 together close the existence question.

Theorem 5.2 (Existence, exactly characterized). *For a centred pre-rounding law and $n \geq 1$, there exists a strictly improving positive dispersion,*

$$\exists s > 0 \quad \text{such that} \quad \text{MSE}_n(s) < \text{MSE}_n(0),$$

if and only if $n > 1$ and $0 < \text{Var } Z < \infty$.

Two remarks fix the logical shape of this result. First, the two directions have different characters. Sufficiency is *local*: it reads the sign of the one-sided derivative

$$\left. \frac{d}{ds} \right|_{s=0+} \text{MSE}_n(s) = -\frac{n-1}{n} \cdot \frac{1}{2} \mathbb{E}|W| < 0 \quad (5.2)$$

straight off the corner, with no need to know where the minimum is. Necessity is *global*: each failure mode is excluded by a statement about the whole function MSE_n , not its germ at the origin. The asymmetry is worth stating plainly, because it is what makes the characterization usable: to certify that scatter helps one differentiates once at the endpoint, whereas to certify that it cannot one must control the entire curve.

Remark 5.3 (Centredness is a normalization, not a hypothesis). The assumption $\mathbb{E}[Z] = 0$ is a normalization. Write $Z = m + Z^\circ$, where Z° is centred. The mean error of one rounded reading is

$$B_s(u) := \mathbb{E}[Y_1(u, s) - u] = sm - \mathbb{E}[q_\varepsilon(u + sm + sZ^\circ)],$$

so, by the period-one shift-invariance of the cell integral and $\int_{-1/2}^{1/2} \mathbb{E}[q_\varepsilon(v + sZ^\circ)] dv = 0$,

$$\int_{-1/2}^{1/2} B_s(u)^2 du = \int_{-1/2}^{1/2} (sm - \mathbb{E}[q_\varepsilon(v + sZ^\circ)])^2 dv = s^2 m^2 + A_{Z^\circ}(s).$$

The cost term carries the same factor: by the cell-integral calculation in the proof of Proposition 3.1, without normalising, $\int_{-1/2}^{1/2} \mathbb{E}[\Delta(u, s)^2] du = s^2 \mathbb{E}[Z^2] + \frac{1}{12} = s^2(m^2 + \text{Var } Z^\circ) + \frac{1}{12}$. Assembling the cell average of the variance-bias split (3.2), the mean enters both terms and collects as $\frac{1}{n}s^2 m^2 + \frac{n-1}{n}s^2 m^2 = s^2 m^2$, leaving

$$\text{MSE}_n(s) = s^2 m^2 + \text{MSE}_n^{Z^\circ}(s),$$

where $\text{MSE}_n^{Z^\circ}$ is the cell-averaged error (2.4) of the centred law Z° . Thus a non-zero mean adds the pure quadratic cost $s^2 m^2$ to the centred problem; since $s^2 m^2$ vanishes to second order at $s = 0$, it can shift the optimizer toward zero but cannot create or destroy the first-order reduction in the bias. The substantive condition is $0 < \text{Var } Z < \infty$.

With existence settled, the rest of the note computes the optimum. For that we need the spectral form of A .

6 The spectral form and a localization bound

The bridge from the corner to closed-form optima is the Fourier series of g .

Lemma 6.1 (Spectral form of the bias functional). *For every $s \geq 0$,*

$$A(s) = \frac{1}{2\pi^2} \sum_{k=1}^{\infty} \frac{|\varphi(2\pi ks)|^2}{k^2}, \quad A(0) = \frac{1}{2\pi^2} \sum_{k \geq 1} \frac{1}{k^2} = \frac{1}{12}. \quad (6.1)$$

Moreover, writing $|\varphi(t)|^2 = \mathbb{E}[\cos(tW)]$,

$$A(0) - A(s) = \frac{1}{2\pi^2} \sum_{k \geq 1} \frac{1 - \mathbb{E}[\cos(2\pi ksW)]}{k^2} = \mathbb{E}\left[\frac{1}{2} \langle sW \rangle (1 - \langle sW \rangle)\right]. \quad (6.2)$$

Proof. In $L^2[-\frac{1}{2}, \frac{1}{2}]$ the sawtooth has the sine series $q_\varepsilon(x) = \sum_{k \geq 1} \beta_k \sin(2\pi kx)$ with $\beta_k = (-1)^{k+1}/(\pi k)$, so $\beta_k^2 = 1/(\pi^2 k^2)$. Sine orthogonality gives $\int_{-1/2}^{1/2} \sin(2\pi kv) \sin(2\pi j(v+\delta)) dv = \frac{1}{2} \delta_{jk} \cos(2\pi k\delta)$, hence

$$g(\delta) = \sum_{k \geq 1} \frac{\beta_k^2}{2} \cos(2\pi k\delta) = \frac{1}{2\pi^2} \sum_{k \geq 1} \frac{\cos(2\pi k\delta)}{k^2}. \quad (6.3)$$

This series converges absolutely and uniformly (dominated by $\sum k^{-2}$), so expectation passes through the sum in $A(s) = \mathbb{E}[g(sW)]$; with $\mathbb{E}[\cos(2\pi ksW)] = |\varphi(2\pi ks)|^2$ from $|\varphi(t)|^2 = \mathbb{E}[e^{itW}]$ this gives (6.1). Subtracting from $A(0) = g(0) = \frac{1}{12}$ gives the first equality of (6.2). Applying the classical evaluation

$$\sum_{k \geq 1} \frac{1 - \cos(2\pi kx)}{k^2} = \pi^2 \langle x \rangle (1 - \langle x \rangle) \quad (6.4)$$

inside the expectation (legitimate by Tonelli, the summand being non-negative) gives the second equality, which is the deficit form of (4.3) and recovers Lemma 4.3 on dividing by s . $\square$

Identity (6.4) is the single welded joint between the two routes: read left to right it is the spectral series; read right to left it is the corner term $\frac{1}{2}\langle x \rangle(1 - \langle x \rangle)$ whose linear cusp drove the existence proof. The same fact does both jobs.

The spectral form makes one universal bound on the optimizer immediate.

Corollary 6.2 (Localization). *Under the hypotheses of Theorem 4.4, every strictly improving positive dispersion—in particular every global minimizer s_n^* —satisfies*

$$(s_n^*)^2 < \frac{n-1}{12}. \quad (6.5)$$

Proof. Improvement requires that the reduction in the bias outweighs the variance cost in (3.4): $s^2/n < \frac{n-1}{n}(A(0) - A(s))$. But $A(0) - A(s) \leq A(0) = \frac{1}{12}$, so $s^2 < (n-1)/12$. $\square$

The bound (6.5) is honest but loose: its scale $\sqrt{(n-1)/12}$ grows like $\sqrt{n}$, while the optimizers we compute below grow far more slowly. For the Gaussian law the optimizer grows like $\sqrt{\log n}$ (§7), whereas the uniform and two-point optimizers converge to finite limits. Rapid decay of φ alone does not force the Gaussian rate. If φ vanishes outside a bounded interval $[-T, T]$ —as it does, after rescaling to unit variance, for the centred finite-variance law with density proportional to $(\sin(x/2)/(x/2))^4$, whose characteristic function is the self-convolution of a triangle—then every term of (6.1) vanishes for $2\pi s \geq T$, so $A(s) = 0$ and MSE_n is strictly increasing there; hence $s_n^* \leq T/(2\pi)$ for every n . The localization scale is nonetheless the correct endpoint scale: for the uniform law, the exact-debiasing point reaches the boundary when $n = 2$, although there it ties the floor rather than strictly improving.

7 A gallery of laws

We compute the optimum for four centred, unit-variance laws. They are arranged to do four different jobs: the two-point law furnishes the cheapest witness, an exact rational optimizer, and a counterexample to single-well behaviour all at once; the Gaussian is the canonical smooth, uniquely solvable case; the Laplace shows that a heavier tail keeps uniqueness while changing the special function; and the uniform cancels the bias exactly and marks the boundary by saturating the localization bound. Throughout, the *bias-reduction rate* is $\frac{1}{2}\mathbb{E}|W|$, the slope from Lemma 4.3. Optimizers in Table 2 were obtained by minimizing (3.1) numerically and agree with the closed forms below to the digits shown.

When the symmetrized noise is bounded, the criterion is exactly quadratic on an initial interval, and the two bounded laws of the gallery share a single closed-form optimizer.

Lemma 7.1 (Bounded symmetrized noise). *Suppose $|W| \leq w$ almost surely for some $w > 0$, and put*

$$c_n := \frac{(n-1)\mathbb{E}|W|}{4n}.$$

Then for $0 \leq s \leq 1/w$,

$$\text{MSE}_n(s) - \frac{1}{12} = s^2 - 2c_n s. \quad (7.1)$$

Consequently, if $c_n \leq 1/w$, the unique minimizer of MSE_n over $[0, 1/w]$ is c_n , with $\text{MSE}_n(c_n) = \frac{1}{12} - c_n^2$.

Proof. The function $x \mapsto x(1-x)$ is symmetric about $x = \frac{1}{2}$, so for $|\delta| \leq 1$ one has $\langle \delta \rangle(1 - \langle \delta \rangle) = |\delta|(1 - |\delta|)$ whether $\langle \delta \rangle = |\delta|$ or $\langle \delta \rangle = 1 - |\delta|$. For $0 \leq s \leq 1/w$ we have $|sW| \leq 1$, so by (6.2) and $\mathbb{E}[W^2] = 2$,

$$A(0) - A(s) = \mathbb{E}\left[\frac{1}{2}s|W|(1 - s|W|)\right] = \frac{s}{2}\mathbb{E}|W| - s^2.$$

Substitution into (3.4) gives $\text{MSE}_n(s) - \frac{1}{12} = \frac{s^2}{n} - \frac{n-1}{n}\left(\frac{s}{2}\mathbb{E}|W| - s^2\right) = s^2 - 2c_n s$, which is (7.1). This parabola has its vertex at $s = c_n$ and value $-c_n^2$ there. $\square$

7.1 Two points: the minimal witness, exactly

Let $Z = \pm 1$ with probability $\frac{1}{2}$ each—the cheapest non-degenerate centred unit-variance law. Then $\varphi(t) = \cos t$, $W = Z - Z' \in \{-2, 0, 2\}$ with $\Pr(W = 0) = \frac{1}{2}$ and $\Pr(W = \pm 2) = \frac{1}{4}$, and $\mathbb{E}|W| = 1$, so the relief rate is exactly $\frac{1}{2}$. Because W is supported on $\{0, \pm 2\}$, (4.2) gives A in closed form with no series at all: $A(s) = \frac{1}{2}g(0) + \frac{1}{2}g(2s)$. For $s \in [0, \frac{1}{4}]$ one has $2s \in [0, \frac{1}{2}]$, so $\langle 2s \rangle = 2s$ and

$$A(s) = \frac{1}{2} \cdot \frac{1}{12} + \frac{1}{2}\left(\frac{1}{12} - s + 2s^2\right) = s^2 - \frac{s}{2} + \frac{1}{12}, \quad s \in [0, \frac{1}{4}]. \quad (7.2)$$

Substituting into (3.4), the contest collapses to a single parabola,

$$\text{MSE}_n(s) - \frac{1}{12} = s^2 - \frac{n-1}{2n}s,$$

which is (7.1) of Lemma 7.1 with $w = 2$ and $c_n = (n-1)/(4n)$. Its minimum over $[0, \frac{1}{4}]$ is attained at

$$s_n^* = \frac{n-1}{4n} \in (0, \frac{1}{4}), \quad \text{MSE}_n(s_n^*) = \frac{1}{12} - \frac{(n-1)^2}{16n^2}. \quad (7.3)$$

This is a fully elementary, rational optimum: no special function is involved, it is valid for every $n > 1$, and, since $s_n^* < \frac{1}{4}$, it always lies inside $[0, \frac{1}{4}]$.

It is also the global minimizer. With

$$r(s) = \text{dist}(2s, \mathbb{Z}),$$

one has $A(0) - A(s) = \frac{1}{4}r(s)(1 - r(s))$, so by (3.4)

$$n\left(\text{MSE}_n(s) - \frac{1}{12}\right) = s^2 - \frac{n-1}{4}r(s)(1 - r(s)) \geq s^2 - \frac{n-1}{16}, \quad (7.4)$$

because $r(1-r) \leq \frac{1}{4}$. By (7.3) the minimum over $[0, \frac{1}{4}]$ has $n(\text{MSE}_n(s_n^*) - \frac{1}{12}) = -(n-1)^2/(16n)$, and the lower bound in (7.4) exceeds this whenever $s^2 > (n-1)/(16n)$. Since $(n-1)/(16n) < \frac{1}{16}$, every $s \geq \frac{1}{4}$ gives a strictly larger value than s_n^* . On the remaining interval $[0, \frac{1}{4}]$ the criterion is the parabola above, so s_n^* is the unique global minimizer.

Yet

$$|\varphi(2\pi ks)|^2 = \cos^2(2\pi ks)$$

does not decay. Each harmonic revives with period $\frac{1}{2}$ in s , and for larger n the criterion develops later local wells. For each $\ell \in \mathbb{Z}_{\geq 0}$ parameterize the ℓ -th arch by

$$s = \frac{\ell}{2} + x, \quad 0 \leq x \leq \frac{1}{2}.$$

On the whole arch $r(s)(1 - r(s)) = 2x(1 - 2x)$, whether $r = 2x$ or $r = 1 - 2x$, so

$$n\left(\text{MSE}_n(s) - \frac{1}{12}\right) = nx^2 + \left(\ell - \frac{n-1}{2}\right)x + \frac{\ell^2}{4}, \quad 0 \leq x \leq \frac{1}{2}.$$

This parabola has an interior minimum, at $x = (\frac{n-1}{2} - \ell)/(2n) \in (0, \frac{1}{4})$, exactly when $\ell < (n-1)/2$; otherwise it is increasing on the arch. At a junction $s = \ell/2$ with $\ell \geq 1$ the left derivative, taken from arch $\ell-1$ at $x = \frac{1}{2}$, is $\frac{n-1}{2} + \ell > 0$, so a junction is never a local minimum. Hence MSE_n has exactly $\lfloor n/2 \rfloor$ local minima on $(0, \infty)$, one on each arch $\ell = 0, 1, \dots, \lfloor n/2 \rfloor - 1$. For $n = 2, 3$ the criterion is single-welled; for $n \geq 4$ it has several wells, of which the first, (7.3), contains the unique global minimizer. The two-point law therefore shows that finite variance ensures existence but does not imply a globally simple objective. For plots of MSE_n against s for two-point scatter, see figure 2.

7.2 Gaussian: the theta optimizer

For $Z \sim \mathcal{N}(0, 1)$, $\varphi(t) = e^{-t^2/2}$ and $|\varphi(2\pi ks)|^2 = e^{-4\pi^2 k^2 s^2}$, so (6.1) reads

$$A(s) = \frac{1}{2\pi^2} \sum_{k \geq 1} \frac{e^{-4\pi^2 k^2 s^2}}{k^2}. \quad (7.5)$$

The damping is faster than any power of k , so A is smooth on $(0, \infty)$ and may be differentiated term by term. Writing $H(s) := \sum_{k \geq 1} e^{-4\pi^2 k^2 s^2}$, one finds $A'(s) = -4s H(s)$ and hence, from (3.1), $\text{MSE}'_n(s) = \frac{2s}{n} (1 - 2(n-1)H(s))$. Since H is strictly decreasing from $+\infty$ to 0, the bracket increases strictly from $-\infty$ to 1, vanishing once: the optimizer is *unique* and solves

$$H(s_n^*) = \frac{1}{2(n-1)}, \quad \text{equivalently} \quad \vartheta_3(0, e^{-4\pi^2 (s_n^*)^2}) = \frac{n}{n-1}, \quad (7.6)$$

where $\vartheta_3(0, r) = 1 + 2 \sum_{j \geq 1} r^{j^2}$ and we used $\vartheta_3(0, r) = 1 + 2H$.

Keeping only the dominant $k = 1$ term gives the large- n asymptotic

$$s_n^* \sim \frac{\sqrt{\log(2(n-1))}}{2\pi} \quad (n \rightarrow \infty), \quad (7.7)$$

the slow $\sqrt{\log n}$ growth promised in §6; it already agrees with the exact root to three decimals by $n = 10$ (Table 1). The bias-reduction rate is $\frac{1}{2} \mathbb{E}|W| = \frac{1}{2} \cdot 2/\sqrt{\pi} = 1/\sqrt{\pi} \approx 0.5642$, since $W \sim \mathcal{N}(0, 2)$.

n	s_n^* (exact root)	$\text{MSE}_n(s_n^*)$	asymptotic (7.7)
2	0.14105	0.06344	0.13251
3	0.18839	0.04796	0.18739
5	0.22961	0.03227	0.22951
10	0.27059	0.01819	0.27058
100	0.36600	0.00243	0.36600
1000	0.43876	0.00030	0.43876

Table 1: Gaussian noise: the unique optimizer s_n^* solving the theta equation (7.6), the cell-averaged error there, and the asymptotic (7.7). The optimizer grows like $\sqrt{\log n}$; the asymptotic is accurate to three decimals from $n = 10$.

For plots of MSE_n against s for Gaussian scatter, see figure 3.

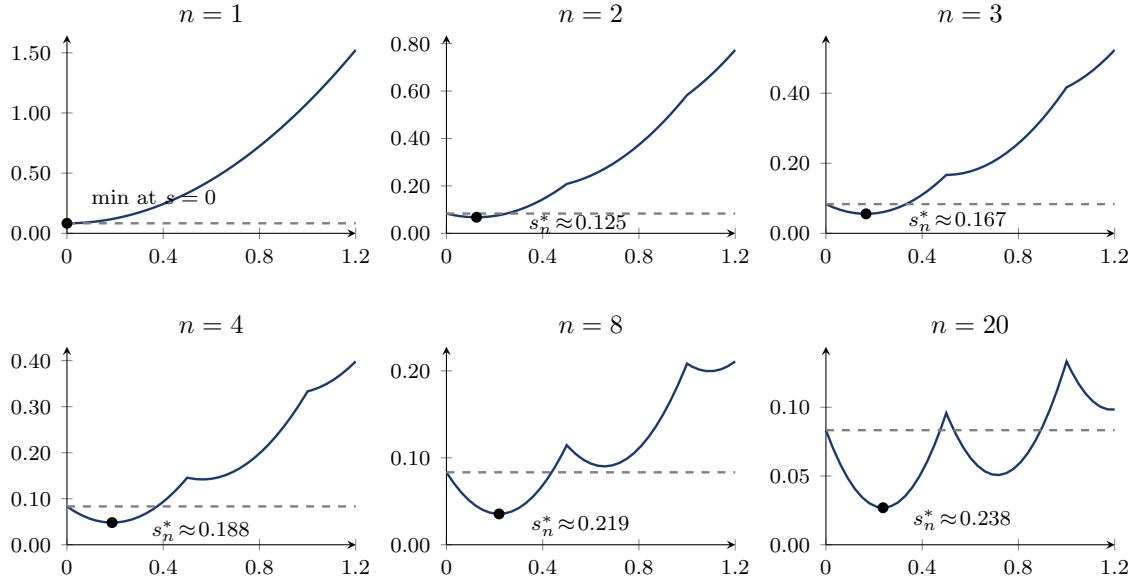


Figure 2: Two-point noise ($Z = \pm 1$ with equal probability). Cell-averaged $\text{MSE}_n(s)$ as a function of the scatter level s measured in cell widths, for $n \in \{1, 2, 3, 4, 8, 20\}$. The dashed line marks the zero-scatter value $\text{MSE}_n(0) = 1/12$; the dot marks the global minimizer. The minimizer follows the exact rational form $s_n^* = (n-1)/(4n)$ with value $1/12 - (n-1)^2/(16n^2)$. Because $|\varphi(2\pi ks)|^2 = \cos^2(2\pi ks)$ does not decay, the harmonics revive with period $\frac{1}{2}$ in s ; the curve has exactly $\lfloor n/2 \rfloor$ local wells, so several wells appear from $n = 4$ onward (here $n = 4, 8, 20$), and the first well always contains the global minimizer. Vertical scales differ between panels.

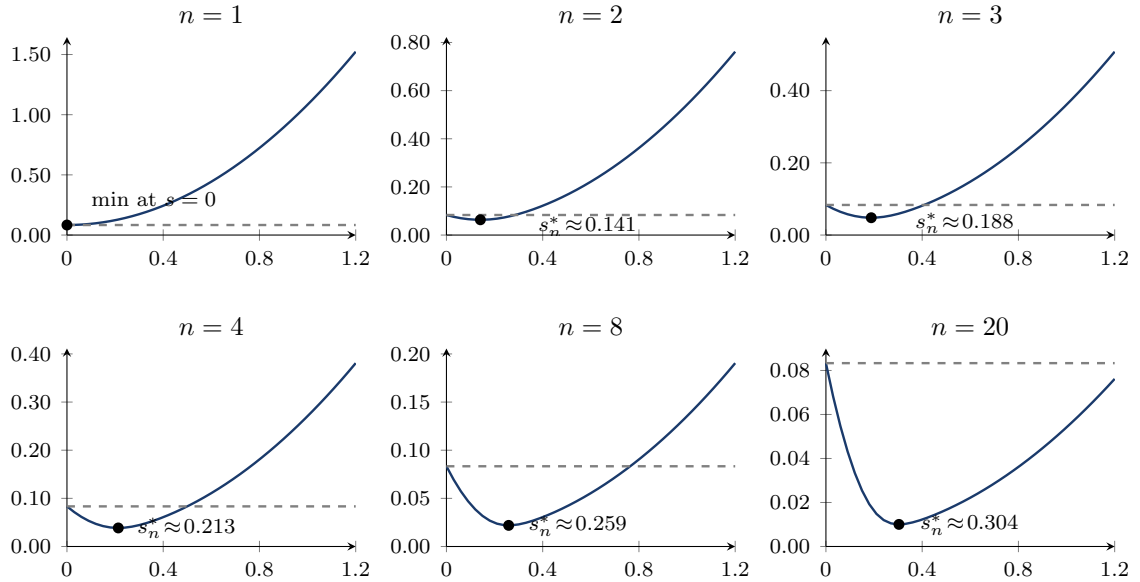


Figure 3: Gaussian noise. Cell-averaged $\text{MSE}_n(s)$ as a function of the scatter level s measured in cell widths, for $n \in \{1, 2, 3, 4, 8, 20\}$. The dashed line marks the zero-scatter value $\text{MSE}_n(0) = 1/12$; the dot marks the global minimizer. For $n = 1$ the minimum is at $s = 0$; for every $n > 1$ it is strictly interior, and the minimum value lies strictly below $1/12$. Vertical scales differ between panels.

7.3 Laplace: a hyperbolic bias term, still a unique optimizer

The unit-variance Laplace law has $\varphi(t) = (1 + t^2/2)^{-1}$, so $|\varphi(2\pi ks)|^2 = (1 + 2\pi^2 k^2 s^2)^{-2}$ and, for $s > 0$, the series (6.1) sums to a hyperbolic closed form: with $a = 1/(s\sqrt{2})$,

$$A(s) = s^2 + \frac{1}{12} - \frac{3}{8a} \coth a - \frac{1}{8} \operatorname{csch}^2 a. \quad (7.8)$$

(The right-hand side is defined for $s > 0$, where $a = 1/(s\sqrt{2})$ is finite; as $s \downarrow 0$ one has $a \rightarrow \infty$, $\frac{3}{8a} \coth a \rightarrow 0$ and $\frac{1}{8} \operatorname{csch}^2 a \rightarrow 0$, so it extends continuously to $A(0) = \frac{1}{12}$, agreeing with the series. We verified (7.8) against the defining series to better than 10^{-25} .) Where the Gaussian produces a theta function, the heavier-tailed Laplace produces $\coth$; the contrast is a feature, not an accident, and traces to the analytic class of $|\varphi|^2$. For Laplace noise $|\varphi(t)|^2$ is a completely monotone function of t^2 —equivalently $W = Z - Z'$ is a scale mixture of Gaussians—and that single structural property forces uniqueness, by the following general mechanism.

Proposition 7.2 (Uniqueness for Gaussian scale mixtures). *Suppose $W = Z - Z'$ is a Gaussian scale mixture; that is, $|\varphi(t)|^2 = \int_0^\infty e^{-\lambda t^2} d\mu(\lambda)$ for a finite measure μ with $\int_0^\infty \lambda d\mu(\lambda) < \infty$. Then for every $n > 1$ the optimizer of MSE_n is unique.*

Proof. By (6.1) and the representation of $|\varphi|^2$,

$$A(s) = \int_0^\infty A_G(s\sqrt{\lambda}) d\mu(\lambda),$$

where A_G is the Gaussian functional (7.5) (legitimate by Tonelli, the summand being non-negative). Fix a compact interval $[s_1, s_2] \subset (0, \infty)$. For $s \in [s_1, s_2]$ the s -derivative of the integrand is $-4s \lambda H(s\sqrt{\lambda})$, and since H is decreasing, $0 \leq \lambda H(s\sqrt{\lambda}) \leq \lambda H(s_1\sqrt{\lambda}) =: G(\lambda)$. The bound G is bounded on $(0, \infty)$ —it tends to 0 as $\lambda \downarrow 0$ (since $H(r) \sim 1/(4\sqrt{\pi}r)$, so $\lambda H(s_1\sqrt{\lambda}) \sim \sqrt{\lambda}/(4\sqrt{\pi}s_1)$) and as $\lambda \rightarrow \infty$ (where H decays super-polynomially)—hence μ -integrable, μ being finite. One may therefore differentiate under the integral; with $A'_G(r) = -4r H(r)$ this gives, for $s > 0$,

$$\operatorname{MSE}'_n(s) = \frac{2s}{n} \left(1 - 2(n-1) H_\mu(s) \right), \quad H_\mu(s) := \int_0^\infty \lambda H(s\sqrt{\lambda}) d\mu(\lambda).$$

For each $\lambda > 0$ the map $s \mapsto H(s\sqrt{\lambda})$ is strictly decreasing from $+\infty$ to 0; the fibre $\lambda = 0$ contributes $\lambda H(s\sqrt{\lambda}) \equiv 0$ and is inert. Because Z is non-degenerate, $\mu((0, \infty)) > 0$, so H_μ is strictly decreasing on $(0, \infty)$, with $H_\mu(s) \rightarrow +\infty$ as $s \downarrow 0$ (monotone convergence) and $H_\mu(s) \rightarrow 0$ as $s \rightarrow \infty$ (dominated convergence, via G). Hence the bracket $1 - 2(n-1)H_\mu(s)$ is strictly increasing from $-\infty$ to 1 and vanishes exactly once; MSE'_n changes sign once on $(0, \infty)$, from $-$ to $+$, so the minimizer is unique. $\square$

Proposition 7.2 covers both the Gaussian ($\mu = \delta_1$) and the Laplace; for the latter the optimizer is the unique root of $\operatorname{MSE}'_n = 0$ with A given by (7.8). The bias-reduction rate is $\frac{1}{2}\mathbb{E}|W| = \frac{3}{4\sqrt{2}} \approx 0.530$, slightly below the Gaussian's $1/\sqrt{\pi} \approx 0.564$. The two optimizers track closely but do not nest: the smaller bias-reduction rate places the Laplace optimizer marginally *below* the Gaussian one for small n (at $n = 2, 3, 4$; Table 2), while the polynomial—rather than Gaussian—decay of $|\varphi(2\pi ks)|^2$ keeps relief alive at larger scales, so the Laplace optimizer overtakes the Gaussian near $n = 5$ and thereafter sits higher and grows faster. The crossover is a spectral-tail effect, not a relief-rate effect. A clean necessary-and-sufficient condition for uniqueness—a statement about the shape of A alone, neither requiring complete monotonicity nor excluded by the two-point oscillation—remains open. For plots of MSE_n against s for Laplacian scatter, see figure 4.

7.4 Uniform: exact debiasing and the localization scale

For $Z \sim \text{Unif}[-\sqrt{3}, \sqrt{3}]$ (variance one), $\varphi(t) = \sin(\sqrt{3}t)/(\sqrt{3}t)$, so

$$|\varphi(2\pi ks)|^2 = \left(\frac{\sin(2\sqrt{3}\pi ks)}{2\sqrt{3}\pi ks} \right)^2,$$

which does not decay monotonically but *vanishes* whenever $2\sqrt{3}s \in \mathbb{Z}$. Consequently every harmonic is annihilated at each positive integer multiple $s = ms_0$ ($m = 1, 2, \dots$) of the first positive cancellation scale

$$s_0 = \frac{1}{2\sqrt{3}}, \quad s_0^2 = \frac{1}{12}. \quad (7.9)$$

At s_0 , $|\varphi(2\pi ks_0)|^2 = \sin^2(\pi k)/(\pi k)^2 = 0$ for every $k \geq 1$, so $A(s_0) = 0$: the cell-averaged rounding bias is cancelled exactly. There

$$\text{MSE}_n(s_0) = \frac{1}{n} \left(\frac{1}{12} + \frac{1}{12} \right) = \frac{1}{6n}, \quad (7.10)$$

equal to the floor $\frac{1}{12}$ at $n = 2$ and strictly below it for every $n > 2$. Two things are worth noting. First, the tie at $n = 2$ occurs at $s_0^2 = \frac{1}{12} = (n-1)/12$: the uniform law saturates the localization bound (6.5), so its scale $\sqrt{(n-1)/12}$ is realized, not merely an artefact of the proof. Second, s_0 is a *debiasing* point, not the global optimizer, which is explicit. Here $|W| \leq 2\sqrt{3}$ and $\mathbb{E}|W| = 2/\sqrt{3}$, so Lemma 7.1 applies with $w = 2\sqrt{3}$, $1/w = s_0$ and $c_n = (n-1)/(2\sqrt{3}n) < s_0$: on $[0, s_0]$,

$$\text{MSE}_n(s) - \frac{1}{12} = s^2 - \frac{n-1}{\sqrt{3}n} s,$$

uniquely minimized at

$$s_n^* = \frac{n-1}{2\sqrt{3}n} = \frac{n-1}{n} s_0, \quad \text{MSE}_n(s_n^*) = \frac{1}{12} - \frac{(n-1)^2}{12n^2} = \frac{2n-1}{12n^2}. \quad (7.11)$$

On the other side, for every $s \geq s_0$,

$$\text{MSE}_n(s) = \frac{s^2}{n} + \frac{1}{12n} + \frac{n-1}{n} A(s) \geq \frac{s^2}{n} + \frac{1}{12n} \geq \frac{s_0^2}{n} + \frac{1}{12n} = \frac{1}{6n} > \frac{2n-1}{12n^2},$$

so no $s \geq s_0$ competes, and (7.11) is the unique global minimizer. It lies strictly below the debiasing scale for every $n > 1$, rising to s_0 as $n \rightarrow \infty$; in particular the localization bound (6.5) is strict for the optimizer even when $n = 2$, where only the debiasing point attains it (Table 2). The bias-reduction rate is $\frac{1}{2}\mathbb{E}|W| = 1/\sqrt{3} \approx 0.5774$.

The simultaneous cancellation of all bias harmonics is the familiar uniform- dither condition of Schuchman [5]; a modern survey appears in [3]. In the present calculation it follows directly from the uniform characteristic function and is the saturation endpoint of Corollary 6.2. Section 8 distinguishes this mathematical overlap from the estimation question studied here. For plots of MSE_n against s with uniform scatter, see figure 5.

8 Relation to dithered quantization

The same integrals and lattice-frequency sums occur in dither theory, but the role of the randomness and the inferential objective are different. In the present model, sZ represents pre-existing measurement unrepeatability. The criterion is the cell-averaged mean-squared

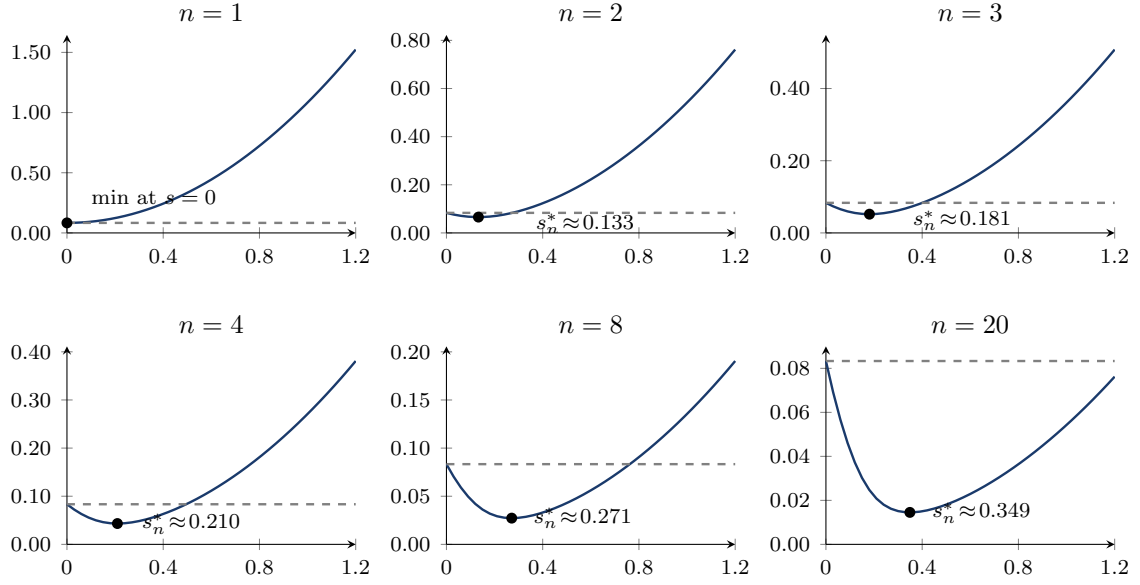


Figure 4: Laplace noise (unit variance, $\varphi(t) = (1 + t^2/2)^{-1}$). Cell-averaged $\text{MSE}_n(s)$ as a function of the scatter level s measured in cell widths, for $n \in \{1, 2, 3, 4, 8, 20\}$. The dashed line marks the zero-scatter value $\text{MSE}_n(0) = 1/12$; the dot marks the global minimizer. The slightly smaller bias-reduction rate places the optimizer marginally below the Gaussian for small n (visible for $n = 2, 3, 4$), while the heavier, polynomially decaying spectral tail makes it overtake the Gaussian near $n = 5$ and grow somewhat faster thereafter. Vertical scales differ between panels.

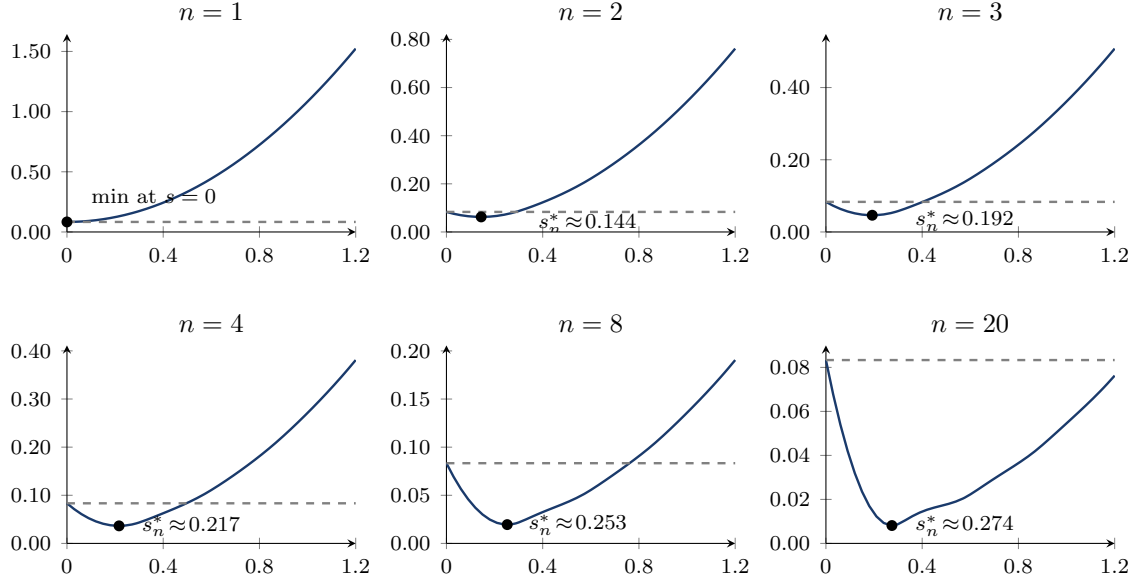


Figure 5: Uniform noise ($Z \sim \text{Unif}[-\sqrt{3}, \sqrt{3}]$, variance one). Cell-averaged $\text{MSE}_n(s)$ as a function of the scatter level s measured in cell widths, for $n \in \{1, 2, 3, 4, 8, 20\}$. The dashed line marks the zero-scatter value $\text{MSE}_n(0) = 1/12$; the dot marks the global minimizer. The spectrum $|\varphi(2\pi ks)|^2$ has nulls, so every bias harmonic is annihilated at $s_0 = 1/(2\sqrt{3}) \approx 0.289$ (where the bias vanishes and $\text{MSE}_n = 1/(6n)$), and again at every integer multiple of s_0 . The global optimizer is $s_n^* = (n-1)s_0/n$, strictly below s_0 , with value $(2n-1)/(12n^2)$. The optimizer barely grows past $n = 8$, reflecting that the bias is already nearly cancelled. Vertical scales differ between panels.

		$n = 2$	$n = 5$	$n = 10$
Two-point	s_n^*	0.1250	0.2000	0.2250
	$\text{MSE}_n(s_n^*)$	0.0677	0.0433	0.0327
Gaussian	s_n^*	0.1410	0.2296	0.2706
	$\text{MSE}_n(s_n^*)$	0.0634	0.0323	0.0182
Laplace	s_n^*	0.1329	0.2302	0.2902
	$\text{MSE}_n(s_n^*)$	0.0657	0.0374	0.0234
Uniform	s_n^*	0.1443	0.2309	0.2598
	$\text{MSE}_n(s_n^*)$	0.0625	0.0300	0.0158

Table 2: The global minimizer s_n^* (in cell widths) and the cell-averaged error there, for four unit-variance laws and three sample sizes. Every value lies strictly below the zero-dispersion floor $\text{MSE}_n(0) = 1/12 \approx 0.0833$, as Theorem 4.4 requires. The two-point and uniform entries match the exact forms (7.3) and (7.11); the optimizers are of comparable size across laws and grow only slowly with n .

error of the average of n rounded readings about the unknown sub-cell value. No random signal is introduced by design, and no realization of the scatter is available for subtraction.

In subtractive dithering, a known random signal is added before quantization and subtracted after quantization in order to control the statistics of the quantization error. Schuchman’s condition states that this error is independent of the signal precisely when the dither characteristic function vanishes at every nonzero lattice frequency; the uniform law is the canonical example [5]. See [3] for a survey and [12] for a broader treatment. The exact-cancellation scale in Section 7.4 satisfies the same harmonic condition, which explains the mathematical overlap.

The questions are nevertheless distinct. Dither theory concerns a designed randomization and properties of the resulting quantization error; the present paper concerns estimation by repeated rounded measurements subject to uncontrolled scatter. Schuchman also observed that saturation prevents full independence for a finite-level quantizer. We assume an infinite lattice, so that boundary does not arise. To our knowledge, the specific existence characterization in Theorem 5.2 has not previously been given.

9 Scope and open questions

For the model of Section 2, Theorem 5.2 settles the existence question. The cell-averaged error of the rounded sample mean is strictly below its zero-dispersion value $\frac{1}{12}$ for some positive dispersion—equivalently, every global minimizer lies away from $s = 0$ —if and only if $n > 1$ and the pre-rounding law is centred with finite, non-zero variance. No density, smoothness, or additional tail condition is required. The strict local improvement follows from the corner of the rounding-error autocorrelation: the bias is reduced at first order in s , whereas the variance contribution is second order. Continuity and the quadratic variance term then give attainment of the global minimum.

Finite variance alone does not imply uniqueness. The local bias-reduction rate $\frac{1}{2}\mathbb{E}|W|$ determines the behaviour near zero, whereas uniqueness depends on the global shape of A . When $|\varphi|^2$ is completely monotone in t^2 —equivalently, when W is a Gaussian scale mixture, as in the Gaussian and Laplace examples—the optimizer is unique (Proposition 7.2). When $|\varphi|^2$ oscillates, as for the two-point law, the criterion can have several local wells (for that law, once $n \geq 4$), although its first well contains the unique global minimizer. A useful open

problem is to identify verifiable structural conditions on A , or equivalently on $|\varphi|^2$, that are necessary and sufficient for uniqueness. The growth of the optimizer with n is governed by the decay of $|\varphi|^2$ rather than by the tails of the law itself: the bounded uniform and two-point laws have optimizers converging to finite limits, the Gaussian optimizer grows like $\sqrt{\log n}$, and a characteristic function of bounded support keeps the optimizer bounded (§6). It remains to identify spectral-decay conditions under which the Gaussian $\sqrt{\log n}$ rate holds; decay of φ faster than every inverse power is not sufficient.

Several assumptions delimit the result. The comparison is based on the cell *average*; it does not assert improvement at every fixed sub-cell position. The quantizer is an infinite nearest-integer lattice, so saturation and finite-range effects are excluded; truncation can be reduced to a shifted version of the same calculation. The readings are independent and identically distributed within a centred scale family, and the estimator is their sample mean. Other dependence structures, non-scale families, and estimators tailored to the rounded data may alter or remove the optimum. Finally, s represents residual unrepeatability rather than a controllable intervention. The result is therefore not a prescription to add noise to a measuring instrument. Within this model, it establishes a non-monotone relation between repeatability and cell-averaged estimation error.

Historical note and author contributions

The problem considered in this paper has its origin in Lumsden’s [4] “flogging wall” thought experiment, first encountered by MRD as an undergraduate student in 1979. Lumsden used the flogging wall to test the counterintuitive possibility that, in some circumstances, an unreliable psychometric instrument might outperform a reliable one. Between 1989 and 2014, Daniel D. Reidpath (ORCID: 0000-0002-8796-0420) and MRD intermittently explored a variety of formalized simulation models of the flogging wall which strongly suggested that instrument unreliability could be advantageous. Two unpublished papers were developed on the basis of the simulation results.

In 2023 MRD reformulated the problem in terms of a weighing scale that *truncates* to finite precision, as opposed to the simpler rounding formulation used here. AOBM observed that this finite precision formulation could be expressed naturally as a question about quantization; see for example [12, p. 1]. Our first formalization was restricted to Gaussian scatter using, as a starting point, the treatment of quantized distributions given by Schwartz [6], and drawing also on the quantization framework of Widrow [11]. This led to a moment-based derivation of a formula for the mean-squared error $\text{MSE}_n(u, s)$ around a specific u , followed by integration over $u \in [-\frac{1}{2}, \frac{1}{2})$, differentiation with respect to s , and a Jacobi theta-function expression for the minimizing scale. That work was developed into a short unpublished manuscript.

The present paper arose from the attempt to identify which features of the Gaussian calculation were essential. In particular, MRD conjectured from the moment-based model that positive finite variance might be the necessary and sufficient condition for the existence of a strictly positive minimizer of the cell-averaged mean-squared error, for $n > 1$. The central simplification in the present paper is the replacement of the Gaussian moment calculation by the autocorrelation formulation used above. This change separates the existence argument from the special analytic form of the Gaussian density, and is what allows the result to be stated for the more general class of distributions considered here.

Figures were initially developed in Wolfram Mathematica and then translated into TikZ for inclusion in the manuscript. The mathematical arguments, exposition, and final form of the paper are the responsibility of the authors.

Acknowledgements and AI tool disclosure

The authors used several generative AI tools during the development of this paper, including OpenAI ChatGPT, Anthropic Claude, Google Gemini, and Moonshot AI Kimi. The tools were used to check algebraic derivations from the quantization literature, to test and expand justifications involving Fourier series, convergence, and exchanges of limits, sums, and integrals, and to help identify simpler proof strategies from earlier moment-based formulations. In particular, ChatGPT and Claude independently suggested that it might be possible to simplify the existence argument by working with an autocorrelation function rather than with the moment formula derived from the Gaussian model.

AI tools were also used for proofreading and for checks of notational consistency. Google Gemini and Moonshot Kimi were particularly helpful with LaTeX assistance, and the translation of Wolfram Mathematica plots into TikZ. All AI-assisted mathematical suggestions were reviewed, revised, and checked by the authors. The text, mathematical claims, and final editorial decisions are those of the authors, who take responsibility for any remaining errors.

References

- [1] A. Chapman. *Dividing the Circle: The Development of Critical Angular Measurement in Astronomy, 1500–1850*. 2nd edition, John Wiley & Sons, Chichester, 1995.
- [2] A. Hald. On the history of the correction for grouping, 1873–1922. *Scandinavian Journal of Statistics*, 28(3):417–428, 2001. doi:10.1111/1467-9469.00245.
- [3] S. P. Lipshitz, R. A. Wannamaker, and J. Vanderkooy. Quantization and dither: a theoretical survey. *Journal of the Audio Engineering Society*, 40(5):355–375, 1992.
- [4] J. Lumsden. Test theory. *Annual Review of Psychology*, 27(1):251–280, 1976.
- [5] L. Schuchman. Dither signals and their effect on quantization noise. *IEEE Transactions on Communication Technology*, COM-12(4):162–165, December 1964. doi:10.1109/TCOM.1964.1088973.
- [6] L. M. Schwartz. Round-off error. *Analytical Chemistry*, 52(7):1141–47. doi:10.1021/ac50057a033.
- [7] W. F. Sheppard. On the calculation of the most probable values of frequency-constants, for data arranged according to equidistant division of a scale. *Proceedings of the London Mathematical Society*, s1-29(1):353–380, 1898. doi:10.1112/plms/s1-29.1.353.
- [8] O. B. Sheynin. Mathematical treatment of astronomical observations: A historical essay. *Archive for History of Exact Sciences*, 11(2–3):97–126, 1973.
- [9] O. B. Sheynin. The true value of a measured constant and the theory of errors. *Historia Scientiarum*, 17(1):38–48, 2007.
- [10] T. Simpson. A letter to the Right Honourable George Earl of Macclesfield, President of the Royal Society, on the advantage of taking the mean of a number of observations, in practical astronomy. *Philosophical Transactions of the Royal Society of London*, 49:82–93, 1755. doi:10.1098/rstl.1755.0020.
- [11] B. Widrow. Statistical analysis of amplitude-quantized sampled-data systems. *Transactions of the American Institute of Electrical Engineers, Part II: Applications and Industry*, 79(6):555–568, 1961. doi:10.1109/TAI.1961.6371702.

- [12] B. Widrow and I. Kollár. *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*. Cambridge University Press, Cambridge and New York, 2008.