NONSMOOTH NONCONVEX OPTIMIZATION FROM NOISY COMPARISONS: UNBIASED GAP ESTIMATION AND THE PRICE OF PREFERENCE FEEDBACK

Anonymous authors
Paper under double-blind review

ABSTRACT

We study the minimization of L-Lipschitz, possibly nonsmooth and nonconvex $f : \mathbb{R}^d \to \mathbb{R}$ from noisy pairwise comparisons: comparing x and y returns a bit equal to 1 with probability $\sigma((f(y) - f(x))/\tau)$, for a known link σ and temperature τ, and the goal is a (δ, ε)-Goldstein stationary point. We give an exactly unbiased estimator of the gap $f(y) - f(x)$ whenever σ^{-1} is analytic on a lens in the complex plane, as for the logistic, probit and cauchit links: compare the pair until the outcome changes and return $\pm\tau$ times a partial sum of the expansion coefficients of the inverse link, the harmonic number of the run length for the logistic link. It uses at most $1+1/(1-\alpha)$ comparisons in expectation, α the largest queried success probability, and has second moment of order the gap squared plus τ^2. With randomized smoothing and the online-to-nonconvex conversion this gives $O\big(d\,(1 + d\,C_{\mathrm{link}}/\kappa^2)\,L^2\Delta_f/(\delta\varepsilon^3)\big)$ comparisons when $\Delta_f \geq L\delta$, where κ is the ratio of the largest queried gap to τ, which the smoothing radius caps at a constant; relative to the optimal rate $O(dL^2\Delta_f/(\delta\varepsilon^3))$ for noisy value oracles, comparisons cost a factor d/κ^2. We prove that any comparison-based method needs $\Omega(d^2\tau^2/(\varepsilon^2\delta^2))$ comparisons on a smooth convex subclass once $d \gtrsim (\varepsilon\delta/\tau)^2$: in the comparison-noise-dominated regime the bounds agree in d and τ, and a factor $\Delta_f/(\delta\varepsilon)$ remains open. Experiments illustrate the estimator's cost and variance and the d^2 growth of the gradient-estimator second moment.

1 INTRODUCTION

Comparison-based optimization models settings in which an algorithm cannot observe values of an objective f but can ask which of two candidates is better and receive a noisy answer, as in dueling bandits, human-in-the-loop tuning, and preference signals in RLHF-style pipelines (Saha et al., 2021; Tang et al., 2024; Chen et al., 2025). For smooth objectives, convergence guarantees under noisy comparisons are available in the convex (Saha et al., 2021; 2025) and the nonconvex (Bergou et al., 2020; El Bakkali & Saadi, 2025; El Bakkali et al., 2026) regimes. The nonsmooth nonconvex regime, where gradients need not exist and the standard solution concept is (δ, ε)-Goldstein stationarity (Goldstein, 1977; Zhang et al., 2020), has, to our knowledge, not been treated with comparison feedback prior to this work. With noisy *value* feedback it is well understood: randomized smoothing and two-point differences give $O(d^{3/2}\delta^{-1}\varepsilon^{-4})$ evaluations (Lin et al., 2022), improved to $O(d^{3/2}\delta^{-1}\varepsilon^{-3})$ by variance reduction (Chen et al., 2023) and to the dimension-optimal $O(d\,\delta^{-1}\varepsilon^{-3})$ by Kornowski & Shamir (2023) through the online-to-nonconvex conversion of Cutkosky et al. (2023). Those methods need the magnitude of value differences, which comparisons do not reveal.

We work with the standard oracle of preference-based optimization: for $x, y \in \mathbb{R}^d$ the oracle returns $Y \in \{0, 1\}$ with $\mathbb{P}(Y = 1) = \sigma\big((f(y) - f(x))/\tau\big)$, where the link $\sigma : \mathbb{R} \to (0, 1)$ is known, strictly increasing and symmetric, and $\tau > 0$ is a temperature. The central quantity throughout is

$$\kappa \;:=\; \frac{B}{\tau}, \qquad B := \sup \text{ of the queried gaps } |f(y) - f(x)|,$$

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.

Table 1: Oracle complexity for a (δ, ε)-Goldstein stationary point of an L-Lipschitz f with $f(x_0) - \inf f \le \Delta_f$, assuming $\Delta_f \ge L\delta$; constants omitted. For comparisons, $\kappa = \min\{\kappa^\star, L\delta/\tau\}$ (O2NC) or $2L\delta/\tau$ (SGD), $C_{\mathrm{link}} = C_{\mathrm{link}}(\sigma, \kappa)$, and $\mathbb{E}[N] \le 1 + 1/(1-\alpha)$ with $\alpha = \sigma(\kappa)$ is the expected number of comparisons per gap estimate.

Oracle	Method	Complexity	Reference
values, exact or Lipschitz per sample	smoothing + SGD	$d^{3/2}L^3\Delta_f/(\delta\varepsilon^4)$	Lin et al. (2022)
values, Lipschitz per sample	smoothing + O2NC	$d\,L^2\Delta_f/(\delta\varepsilon^3)$	Kornowski & Shamir (2023)
comparisons	smoothing + SGD	$\mathbb{E}[N]\, d^{3/2}\big(1 + \frac{d\,C_{\mathrm{link}}}{\kappa^2}\big)L^3\Delta_f/(\delta\varepsilon^4)$	Thm. 4.3
comparisons	smoothing + O2NC	$\mathbb{E}[N]\, d\,\big(1 + \frac{d\,C_{\mathrm{link}}}{\kappa^2}\big)L^2\Delta_f/(\delta\varepsilon^3)$	Thm. 4.2
comparisons	any	$\Omega\big(d^2\tau^2/(\varepsilon^2\delta^2)\big) = \Omega\big(d^2L^2/(\kappa^2\varepsilon^2)\big)$	Thm. 5.1

the gap-to-temperature ratio. Our queries are symmetric pairs $x \pm \rho u$ with $\|u\| = 1$, so $B = 2L\rho$. The smoothing radius ρ is a parameter distinct from the Goldstein radius δ: it must satisfy $\rho \le \delta/2$, and the main algorithm sets $\rho = \min\{\delta/2, \kappa^\star\tau/(2L)\}$ for a constant $\kappa^\star$, so that $\kappa = \min\{\kappa^\star, L\delta/\tau\}$.

Contributions.

- **Unbiased gap estimation from comparisons (Section 3).** We expand the inverse link $g = \sigma^{-1}$ in the polynomials $b_m(p) = p^m - (1-p)^m$, whose natural estimators from one stream of comparisons vanish after the first change of outcome, so the whole series is computable from finitely many comparisons: $\widetilde{\Delta}$ compares the pair until the outcome changes and returns $\pm\tau$ times a partial sum of the coefficients. It is exactly unbiased, uses at most $1 + 1/(1-\alpha)$ comparisons in expectation, and has $\mathbb{E}[\widetilde{\Delta}^2] \le C_{\mathrm{link}}\tau^2$ with C_{link} depending only on the link and κ ($13(1+\kappa)^2$ for the logistic link). The expansion exists when g is analytic on a lens $\{|z| \le r\} \cap \{|z-1| \le r\}$ of the complex plane (logistic, probit and cauchit links), and analyticity near the real interval alone is not sufficient.
- **Algorithms and rates (Section 4).** The gradient estimator $G = \frac{d}{2\rho}\widetilde{\Delta}\,u$ is unbiased for ∇f_ρ with second moment at most $dL^2(c_1 + d\,C_{\mathrm{link}}/\kappa^2)$. Plugged into the online-to-nonconvex conversion, it yields (δ, ε)-Goldstein stationarity with $O\big((1 + \frac{1}{1-\alpha})\, d(c_1 + dC_{\mathrm{link}}/\kappa^2)\, L^2\Delta_f/(\delta\varepsilon^3)\big)$ comparisons when $\Delta_f \ge L\delta$ (Theorem 4.2), where the conversion's insensitivity to the smoothness of f_ρ lets us choose ρ to keep $\kappa \le \kappa^\star$, so the bound is $O(d^2L^2\Delta_f/(\delta\varepsilon^3))$ for every $\tau \le L\delta/\kappa^\star$; plugged into SGD on the smoothed function, where ρ does enter through the smoothness $\sqrt{d}L/\rho$, it yields $O\big((1 + \frac{1}{1-\alpha})\, d^{3/2}(c_1 + dC_{\mathrm{link}}/\kappa^2)L^3\Delta_f/(\delta\varepsilon^4)\big)$ with $\kappa = 2L\delta/\tau$ (Theorem 4.3).
- **A lower bound in d and τ (Section 5).** For any link with $c_\sigma < \infty$ (a shifted Fisher-information bound met by the logistic, probit and cauchit links), any algorithm that outputs a (δ, ε)-Goldstein stationary point in expectation for every convex, L-Lipschitz, (ε/δ)-smooth f with $f(x_0) - \inf f = O(\varepsilon\delta)$ needs $T = \Omega(d^2\tau^2/(\varepsilon^2\delta^2))$ in high dimension (Theorem 5.1); this agrees with the upper bound in d and τ when κ is bounded and leaves a factor $\Delta_f/(\delta\varepsilon)$ open. The proof is an Assouad argument: a comparison is a Bernoulli observation whose Fisher information about the minimizer has trace $O(\varepsilon^2/\tau^2)$, and locating d coordinates from total information $\mathcal{I}$ costs mean squared error $d^2/\mathcal{I}$.
- **Experiments (Section 6).** Simulations illustrate the cost and variance of the estimator, the d^2 growth of $\mathbb{E}\|G\|^2$, the d^2/T estimation rate on the hard family of the lower bound, and the behavior of both algorithms on a nonsmooth nonconvex test problem.

Table 1 summarizes the rates. The upper and lower bounds match in their dependence on d and κ in the regime $\tau \ge L\delta/\kappa^\star$, where $\kappa = L\delta/\tau$ and C_{link} is an absolute constant; the dependence on δ and ε of the lower bound is not tight, and we state the joint question as open in Section 5.

2 Preliminaries

Notation. $\|\cdot\|$ is the Euclidean norm, $\mathbb{B}_d$ the closed unit ball and $\mathbb{S}^{d-1}$ the unit sphere in $\mathbb{R}^d$. Throughout, $f : \mathbb{R}^d \to \mathbb{R}$ is L-Lipschitz and bounded below, x_0 is the initial point and $\Delta_f \ge f(x_0) - \inf f$.

Goldstein stationarity. By Rademacher's theorem f is differentiable almost everywhere, and its Clarke subdifferential $\partial f(x)$ is the convex hull of all limits of $\nabla f(x_n)$ along differentiable points $x_n \to x$ (Clarke, 1990). The Goldstein δ-subdifferential is $\partial_\delta f(x) := \mathrm{conv}\big(\bigcup_{\|y-x\|\le\delta} \partial f(y)\big)$ and $\big\|\bar\partial_\delta f(x)\big\| := \min\{\|g\| : g \in \partial_\delta f(x)\}$. A point x is (δ,ε)-Goldstein stationary if $\big\|\bar\partial_\delta f(x)\big\| \le \varepsilon$ (Goldstein, 1977; Zhang et al., 2020). If f is H-smooth then $\|\nabla f(x)\| \le \big\|\bar\partial_\delta f(x)\big\| + H\delta$, since every element of $\partial_\delta f(x)$ is a convex combination of gradients at points within δ of x.

Randomized smoothing. For $\rho > 0$ let $f_\rho(x) := \mathbb{E}_{z\sim\mathrm{Unif}(\mathbb{B}_d)} f(x+\rho z)$. Then f_ρ is L-Lipschitz, differentiable, $c_2\sqrt{d}\,L/\rho$-smooth for an absolute constant c_2, and $\|f - f_\rho\|_\infty \le L\rho$ (Yousefian et al., 2012). Two facts connect f_ρ to Goldstein stationarity of f: $\nabla f_\rho(x) \in \partial_\rho f(x)$ (Lin et al., 2022), and more generally $\partial_\nu f_\rho(x) \subseteq \partial_{\rho+\nu} f(x)$ for all $\nu \ge 0$ (Kornowski & Shamir, 2023, Lemma 4). The gradient of f_ρ admits the two-point representation

$$\nabla f_\rho(x) = \frac{d}{2\rho}\,\mathbb{E}_{u\sim\mathrm{Unif}(\mathbb{S}^{d-1})}\big[\Delta_u(x)\,u\big], \qquad \Delta_u(x) := f(x+\rho u) - f(x-\rho u), \tag{1}$$

and the directional gap concentrates: $\mathbb{E}_u[\Delta_u(x)^2] \le c_1 L^2\rho^2/d$ for an absolute constant c_1 (Shamir, 2017, Lemma 10); see Kornowski & Shamir (2023, Lemma 7) for the explicit constant $c_1 = 64\sqrt{2\pi}$. By Lipschitz continuity, $|\Delta_u(x)| \le 2L\rho$ for every u.

Comparison oracle. A link is a strictly increasing $\sigma : \mathbb{R} \to (0,1)$ with $\sigma(-t) = 1 - \sigma(t)$; its inverse is $g := \sigma^{-1} : (0,1) \to \mathbb{R}$, which satisfies $g(1-p) = -g(p)$. Given $\tau > 0$, a query on the pair (x, y) returns $Y \in \{0, 1\}$ with

$$\mathbb{P}(Y = 1 \mid x, y) = \sigma(\Delta/\tau), \qquad \Delta := f(y) - f(x),$$

independently of everything else; repeated queries on the same pair are i.i.d. Writing $p := \sigma(\Delta/\tau)$, the gap is $\Delta = \tau g(p)$. All queries made by our algorithms are symmetric pairs $(x-\rho u, x+\rho u)$ with $u \in \mathbb{S}^{d-1}$, for which $|\Delta| \le B := 2L\rho$. Hence every queried success probability lies in the *operated range* $[1-\alpha, \alpha]$ with

$$\kappa := B/\tau, \qquad \alpha := \sigma(\kappa) \in (1/2, 1). \tag{2}$$

The three links we use as running examples are the logistic $\sigma(t) = 1/(1+e^{-t})$, the probit $\sigma(t) = \Phi(t)$ with Φ the standard normal distribution function, and the cauchit $\sigma(t) = \frac{1}{2} + \frac{1}{\pi}\arctan t$.

3 Unbiased gap estimation from comparisons

The task is to estimate $\Delta = \tau g(p)$ without bias from Bernoulli(p) samples. A function of p is unbiasedly estimable from m Bernoulli samples only if it is a polynomial of degree at most m, so an unbiased estimator of g must use a random number of samples: we expand g in polynomials tailored to comparisons and truncate at a random index.

3.1 A comparison-friendly polynomial basis

Definition 3.1 (Bernoulli-product polynomials and prefix estimators)**.** For $m \ge 1$ let $b_m(p) := p^m - (1-p)^m$. Given a stream $Y_1, Y_2, \ldots$ of i.i.d. $\mathrm{Bern}(p)$ outcomes, let $A_m := \prod_{i\le m} Y_i$ and $B_m := \prod_{i\le m}(1-Y_i)$ be the indicators that the first m outcomes are all ones, respectively all zeros.

Then $\mathbb{E}[A_m - B_m] = b_m(p)$, and because A_m and B_n cannot both equal one,

$$(A_m - B_m)(A_n - B_n) = A_{m\vee n} + B_{m\vee n}, \qquad \mathbb{E}\big[(A_m - B_m)(A_n - B_n)\big] = p^{m\vee n} + (1-p)^{m\vee n}, \tag{3}$$

where $m \vee n := \max(m, n)$. Each b_m is odd about $1/2$, as is g. Expanding g in the b_m requires more than analyticity near the real interval $[1-\alpha, \alpha]$: if $\sum_m |c_m|\alpha^m < \infty$ then $\sum_m c_m z^m$ and $\sum_m c_m(1-z)^m$ converge on the discs $\{|z| < \alpha\}$ and $\{|z-1| < \alpha\}$, so any such expansion extends g analytically to the open lens $\{|z| < \alpha\} \cap \{|z-1| < \alpha\}$, which reaches height $\sqrt{\alpha^2 - 1/4}$ above $1/2$. The converse holds with a slightly larger lens.

Assumption 3.2 (Lens-analytic inverse link)**.** For the operated range $[1-\alpha, \alpha]$ of (2) there is $r \in (\sqrt{\alpha}, 1)$ such that $g = \sigma^{-1}$ extends holomorphically to an open neighborhood of the closed lens $\Lambda_r := \{z \in \mathbb{C} : |z| \le r,\ |z-1| \le r\}$.

Proposition 3.3 (Bernoulli-product expansion)**.** *Under Assumption 3.2 there is a function h holomorphic on $\{|z| < r\}$ such that $g(z) = h(z) - h(1-z)$ on the interior of Λ_r. Let $c_m := h^{(m)}(0)/m!$ for $m \ge 1$ and, for $r' \in (\sqrt{\alpha}, r)$, let $C_{r'} := \max_{|z|=r'} |h(z)|$. Then for every $p \in [1-\alpha, \alpha]$,*

$$g(p) = \sum_{m \ge 1} c_m \, b_m(p), \qquad |c_m| \le C_{r'} \, (r')^{-m},$$

and the series converges absolutely and uniformly on $[1-\alpha, \alpha]$.

The proof (Appendix A.1) splits Cauchy's formula on Λ_r into its two boundary arcs, which give functions holomorphic on $\{|z| < r\}$ and on $\{|z-1| < r\}$ respectively, and uses the oddness of g to merge them into one h. The decomposition is not unique (for instance $b_2 = b_1$), which is harmless: any admissible (c_m) can be used below, and c_m is a Cauchy integral of g over an arc.

Example 3.4. *Logistic:* $g(p) = \log \frac{p}{1-p} = h(p) - h(1-p)$ with $h(z) = -\log(1-z)$, so $c_m = 1/m$ and Assumption 3.2 holds for every $r < 1$, hence for every $\alpha < 1$. *Probit:* Φ is entire with nowhere vanishing derivative, and its only asymptotic values are 0 and 1, so Φ^{-1} continues analytically to every simply connected domain avoiding 0 and 1 (Appendix A.1), in particular to a neighborhood of Λ_r for every $r < 1$. *Cauchit:* $g(p) = \tan(\pi(p - 1/2))$ is meromorphic with poles at the integers only. In all three cases the assumption holds for every $\alpha < 1$, with constants $C_{r'}$ that grow as $\alpha \to 1$. A link whose inverse has a pole at $\frac{1}{2} \pm i\eta$ with $\eta < \sqrt{\alpha^2 - 1/4}$ (Appendix A.1 gives one) is strictly increasing and analytic near $[1-\alpha, \alpha]$, yet admits no expansion with $\sum_m |c_m| \alpha^m < \infty$. The step link (noiseless sign feedback) has no analytic inverse at all, consistent with the fact that magnitudes cannot be recovered from signs.

3.2 The run-length estimator

The products A_m and B_m vanish for every m past the first change of outcome. Hence the series $\sum_m c_m(A_m - B_m)$ has finitely many nonzero terms, and it can be computed exactly.

Definition 3.5 (Run-length estimator)**.** Let (c_m) be as in Proposition 3.3. Query the pair (x, y) repeatedly and stop at the first outcome that differs from Y_1; let $N := \min\{n \ge 2 : Y_n \ne Y_1\}$ be the number of comparisons used, and $S_k := \sum_{m=1}^{k} c_m$. Output

$$\widetilde{\Delta}(x, y) := \tau \, (2Y_1 - 1) \, S_{N-1} = \tau \sum_{m \ge 1} c_m \, (A_m - B_m).$$

The two expressions agree because $A_m - B_m = 2Y_1 - 1$ for $m < N$ and $A_m = B_m = 0$ for $m \ge N$. For the logistic link $S_k = H_k$ is the kth harmonic number: a long run of identical outcomes is evidence of a large gap, and $H_k \approx \log k$ matches the logit. We assume exact arithmetic and access to the real coefficients c_m: $1/m$ for the logistic link, and in general the Cauchy integrals of g over the arcs of $\partial\Lambda_r$ (Appendix A.1), computable on demand; a perturbation δc_m biases $\widetilde{\Delta}$ by at most $2\tau \sum_m |\delta c_m| \alpha^m$.

Theorem 3.6 (Run-length estimator)**.** *Let $p \in [1-\alpha, \alpha]$ and suppose $|c_m| \le C \rho_c^m$ for all m with $\alpha\rho_c < 1$ (Proposition 3.3 gives $\rho_c = 1/r'$ and $C = C_{r'}$).*

(i) $\mathbb{E}[\widetilde{\Delta}] = \Delta$.

(ii) $\mathbb{E}[N] = 1 + \frac{p}{1-p} + \frac{1-p}{p} \le 1 + \frac{1}{1-\alpha}$.

(iii) Let $\bar{\rho} := \max\{1, \rho_c\}$ and $\theta := \alpha\bar{\rho}^{\,2}$. If $\theta < 1$, which holds for $\rho_c = 1/r'$ with $r' > \sqrt{\alpha}$, then

$$\mathbb{E}[\widetilde{\Delta}^2] \ \le \ C_{\mathrm{link}} \, \tau^2, \qquad C_{\mathrm{link}} := \frac{2C^2\theta(1+\theta)}{(1-\theta)^3}.$$

For the logistic link, $\mathbb{E}[\widetilde{\Delta}^2] = \tau^2 \sum_{k \ge 1} \big(p^k(1-p) + (1-p)^k p\big) H_k^2$ exactly, and $\mathbb{E}[\widetilde{\Delta}^2] \le 13\,\tau^2(1+\kappa)^2$.

(iv) $\mathbb{E}|\widetilde{\Delta}|^q < \infty$ *whenever* $\alpha\bar{\rho}^q < 1$ *(all moments for the logistic link).*

The proof is in Appendix A.2: unbiasedness is Fubini's theorem applied to the series, the cost follows from $\mathbb{P}(N > n) = p^n + (1-p)^n$, and the moments from $|\widetilde{\Delta}| = \tau|S_{N-1}|$ with $N-1$ geometric and $|S_k| \le Ck\bar{\rho}^k$, or $S_k = H_k \le 1 + \log k$ for the logistic link. The bound $13\tau^2(1+\kappa)^2$ says that at the edge of the operated range, where $\Delta^2 = \kappa^2\tau^2$, the second moment is at most a constant times $\Delta^2 + \tau^2$: nearly deterministic comparisons cost comparisons ($\mathbb{E}[N] \approx e^\kappa$), not accuracy. The exact series agrees with simulations to three digits and the bound is loose by a factor 15 to 20 (Table 2). The randomly truncated (Russian-roulette) series of McLeish (2010); Rhee & Glynn (2015); Lyne et al. (2015) is also unbiased, and $\widetilde{\Delta}$ is its conditional expectation given the comparison stream, so it has no larger variance when that estimator has a finite second moment. Cost goes the other way: nested Russian roulette may stop at the earlier of its truncation time and the first outcome change, so it is cheaper, and at the settings we tried its product of cost and second moment is smaller. What separates them is the tail (all moments finite for $\widetilde{\Delta}$ under the logistic link, the qth moment when $\alpha\bar{\rho}^q < 1$ in general, against divergent high moments under geometric truncation) and the absence of a parameter to tune; Appendix A.3 has the details, and either estimator can be used in Section 4.

Remark 3.7 (Dependence on κ)**.** Both $\mathbb{E}[N]$ and C_{link} are functions of $\alpha = \sigma(\kappa)$ alone. As $\kappa \to 0$ (gaps small relative to the temperature) $\mathbb{E}[N] \to 3$ and C_{link} stays bounded, while $\Delta^2 \le \kappa^2\tau^2$ becomes negligible against the estimator variance of order τ^2. As $\kappa \to \infty$ (comparisons almost deterministic) $\mathbb{E}[N] \approx 1/(1-\alpha)$, which is e^κ for the logistic link, while $\mathbb{E}[\widetilde{\Delta}^2]/\Delta^2$ stays bounded. Heuristically the cost growth is intrinsic: one comparison carries Fisher information $I_\sigma(\Delta/\tau)/\tau^2$ about Δ, with $I_\sigma(t) := \sigma'(t)^2/(\sigma(t)(1-\sigma(t))) \approx e^{-|t|}$ for the logistic link, so estimating a gap at the edge of the operated range to accuracy τ needs of order $1/(1-\alpha)$ comparisons, which is what $\widetilde{\Delta}$ uses. The figure of merit below is $\mathbb{E}[N] \cdot \mathbb{E}[\widetilde{\Delta}^2]/B^2$, minimized at an intermediate κ (about 11 near $\kappa^\star \approx 1.5$ for the logistic link, Table 2). Since $\kappa = 2L\rho/\tau$, the algorithm controls κ through ρ when τ is small but is confined to $\kappa \le L\delta/\tau$ when τ is large.

4 From gap estimates to Goldstein stationarity

Gradient estimator. For $x \in \mathbb{R}^d$, sample $u \sim \text{Unif}(\mathbb{S}^{d-1})$, run Definition 3.5 on the pair $(x - \rho u, x + \rho u)$ with fresh comparisons, and set

$$G(x;u) := \frac{d}{2\rho}\,\widetilde{\Delta}(x-\rho u, x+\rho u)\,u. \tag{4}$$

Lemma 4.1 (Moments of G)**.** *Under Assumption 3.2,* $\mathbb{E}[G(x;u) \mid x] = \nabla f_\rho(x)$ *and*

$$\mathbb{E}\big[\|G(x;u)\|^2 \mid x\big] \;\le\; \sigma_G^2 := dL^2\Big(c_1 + \frac{d\,C_{\text{link}}}{\kappa^2}\Big), \qquad \kappa = \frac{2L\rho}{\tau}.$$

Proof. Unbiasedness follows from Theorem 3.6(i), conditioning on u, and (1). For the second moment, $\mathbb{E}[\widetilde{\Delta}^2 \mid u] = \Delta_u(x)^2 + \text{Var}(\widetilde{\Delta} \mid u) \le \Delta_u(x)^2 + C_{\text{link}}\tau^2$ by Theorem 3.6(iii), so $\mathbb{E}\|G\|^2 \le \frac{d^2}{4\rho^2}\big(\mathbb{E}_u[\Delta_u(x)^2] + C_{\text{link}}\tau^2\big) \le c_1 dL^2 + \frac{d^2 C_{\text{link}}\tau^2}{4\rho^2}$, and $\tau = 2L\rho/\kappa$. □

The first term is the value-oracle term of Kornowski & Shamir (2023), linear in d because Δ_u concentrates over u. The second is comparison noise: $\text{Var}(\widetilde{\Delta} \mid u)$ is of order τ^2 regardless of u and of ρ, and the factor $d^2/(4\rho^2)$ turns it into $d^2L^2C_{\text{link}}/\kappa^2$, which dominates unless $\kappa^2 \gtrsim d\,C_{\text{link}}$; for the logistic link $C_{\text{link}}/\kappa^2 \le 13(1+1/\kappa)^2$, so this term is at most $13d^2L^2(1+1/\kappa)^2$ for every κ. Averaging k copies of G per step reduces the bound to $\sigma_G^2/k + c_1 dL^2$ at k times the cost, which leaves the total complexities below unchanged in order as long as $k \lesssim \sigma_G^2/(c_1 dL^2) = 1 + dC_{\text{link}}/(c_1\kappa^2)$.

Main algorithm. Algorithm 1 is the online-to-nonconvex conversion of Cutkosky et al. (2023) in the form used by Kornowski & Shamir (2023), with G as the gradient oracle for $h := f_\rho$. It maintains a clipped online-gradient-descent direction Δ_t, evaluates the gradient estimator at a uniformly random point z_t of each step, and outputs the average of z_t over a random block of $\lfloor \nu/D \rfloor$ consecutive steps.

Algorithm 1 Comparison-O2NC

Input: x_0, smoothing radius ρ, radius ν, clipping D, step η, horizon T.
$\Delta_1 \leftarrow 0$.
for $t = 1, \dots, T$ **do**
  $s_t \sim \mathrm{Unif}[0,1]$; $x_t \leftarrow x_{t-1} + \Delta_t$; $z_t \leftarrow x_{t-1} + s_t \Delta_t$.
  $u_t \sim \mathrm{Unif}(\mathbb{S}^{d-1})$; obtain $\widetilde{\Delta}_t \leftarrow \widetilde{\Delta}(z_t - \rho u_t, z_t + \rho u_t)$ (Definition 3.5, N_t comparisons); $G_t \leftarrow \frac{d}{2\rho} \widetilde{\Delta}_t u_t$.
  $\Delta_{t+1} \leftarrow \min\{1,\ D/\|\Delta_t - \eta G_t\|\}\,(\Delta_t - \eta G_t)$.
end for
$M \leftarrow \lfloor \nu/D \rfloor$, $K \leftarrow \lfloor T/M \rfloor$, $\bar{x}_k \leftarrow \frac{1}{M} \sum_{m=1}^{M} z_{(k-1)M+m}$ for $k \le K$; **output** $x^{\mathrm{out}} \leftarrow \bar{x}_k$ with $k \sim \mathrm{Unif}\{1, \dots, K\}$.

Theorem 4.2 (Comparison complexity of Algorithm 1)**.** *Let f be L-Lipschitz and bounded below, $\delta, \varepsilon \in (0,1)$, fix $\kappa^\star > 0$, and set $\rho := \min\{\delta/2,\ \kappa^\star \tau/(2L)\}$, so that $\kappa = 2L\rho/\tau = \min\{\kappa^\star, L\delta/\tau\}$, and $\nu := \delta - \rho \ge \delta/2$. Let the link satisfy Assumption 3.2 with $\alpha = \sigma(\kappa)$. Run Algorithm 1 with these ρ, ν, $\eta = \Delta_h/(\sigma_G^2 T)$, $D = \big(\nu^{1/2} \Delta_h/(\sigma_G T)\big)^{2/3}$, where Δ_h is any upper bound on $f_\rho(x_0) - \inf f_\rho$, for instance $\Delta_f + L\rho$, or $\Delta_f + H\rho^2$ when f is H-smooth, and σ_G is from Lemma 4.1. The algorithm uses L, Δ_h, τ and the link as inputs, and T is taken large enough that $D \le \nu$ (absorbed in c). There is an absolute constant c such that for*

$$T \;=\; c\,\frac{\sigma_G^2\,\Delta_h}{\delta\,\varepsilon^3} \;=\; O\!\left(d\Big(c_1 + \frac{d\,C_{\mathrm{link}}}{\kappa^2}\Big)\frac{L^2(\Delta_f + L\delta)}{\delta\varepsilon^3}\right)$$

the output satisfies $\mathbb{E}\big\|\bar{\partial}_\delta f(x^{\mathrm{out}})\big\| \le \varepsilon$, and the expected number of comparisons is at most $T\,(1 + \frac{1}{1-\alpha})$.

The proof, in Appendix A.4, plugs Lemma 4.1 into the guarantee of the online-to-nonconvex conversion for $h = f_\rho$ and uses $\partial_\nu f_\rho(x) \subseteq \partial_{\rho+\nu} f(x) = \partial_\delta f(x)$.

The simpler variant. SGD on f_δ needs no clipping or blocks but pays for the $\sqrt{d}/\delta$ smoothness of the smoothed objective.

Theorem 4.3 (Comparison-SGD)**.** *Under the assumptions of Theorem 4.2 with $\rho = \delta$ (so $\kappa = 2L\delta/\tau$) and $\varepsilon \le L$, let $H_\delta := c_2\sqrt{d}L/\delta$ and $\Delta_h := \Delta_f + L\delta$. The iteration $x_{t+1} = x_t - \eta G(x_t; u_t)$ with $\eta = \min\{1/H_\delta,\ \sqrt{2\Delta_h/(H_\delta \sigma_G^2 T)}\}$ and output x_R, $R \sim \mathrm{Unif}\{0, \dots, T-1\}$, satisfies $\mathbb{E}\big\|\bar{\partial}_\delta f(x_R)\big\| \le \varepsilon$ once*

$$T \;\ge\; \frac{32\,H_\delta\,\Delta_h\,\sigma_G^2}{\varepsilon^4} \;=\; O\!\left(d^{3/2}\Big(c_1 + \frac{d\,C_{\mathrm{link}}}{\kappa^2}\Big)\frac{L^3(\Delta_f + L\delta)}{\delta\varepsilon^4}\right),$$

using at most $T(1 + \frac{1}{1-\alpha})$ comparisons in expectation.

The proof (Appendix A.5) is the standard nonconvex SGD bound $\frac{1}{T}\sum_t \mathbb{E}\|\nabla f_\delta(x_t)\|^2 \le \frac{2\Delta_h}{\eta T} + \eta H_\delta \sigma_G^2$, followed by $\nabla f_\delta(x) \in \partial_\delta f(x)$.

The smoothing radius is a free parameter. The conversion uses only the Lipschitz constant of f_ρ, its initial gap and σ_G^2, not its smoothness, so ρ can be taken small at no cost beyond $\rho + \nu \le \delta$, and it affects the rate only through $\kappa = 2L\rho/\tau$. With $\rho = \min\{\delta/2, \kappa^\star\tau/(2L)\}$: when $\tau \le L\delta/\kappa^\star$, $\kappa = \kappa^\star$ and C_{link}, $\mathbb{E}[N]$ are absolute constants of the link, so Theorem 4.2 gives $O(d^2L^2\Delta_f/(\delta\varepsilon^3))$ comparisons uniformly in τ; when $\tau > L\delta/\kappa^\star$, $\kappa = L\delta/\tau$ and the bound reads $O(d^2L^2\Delta_f/(\kappa^2\delta\varepsilon^3))$. Against $O(dL^2\Delta_f/(\delta\varepsilon^3))$ evaluations for a Lipschitz-per-sample value oracle, comparisons cost a factor d/κ^2, which the next section shows cannot be removed in the second regime. SGD pays for a small ρ through the smoothness $c_2\sqrt{d}L/\rho$ of f_ρ, so Theorem 4.3 is stated with $\rho = \delta$.

5 A LOWER BOUND IN THE DIMENSION AND THE TEMPERATURE

The lower bound uses only two properties of the oracle: each comparison is a single Bernoulli observation, and its success probability is a smooth function of the gap. Define the link constant

$$c_\sigma := \sup_{t\in\mathbb{R}} \sup_{|s-t|\le 1} \frac{\sigma'(s)^2}{\sigma(t)(1-\sigma(t))},$$

which is at most $e^2/4$ for the logistic link and finite for the probit and cauchit links (Appendix A.6). Taking $s = t$ shows $c_\sigma \ge \sup_t I_\sigma(t)$ for the Fisher information of Remark 3.7, and for the logistic link $c_\sigma \le e^2 \sup_t I_\sigma(t)$.

Theorem 5.1 (Lower bound). *Let σ be a differentiable symmetric link with $c_\sigma < \infty$, let $L, \delta, \varepsilon, \tau > 0$ with $\varepsilon \le L/32$, set $\kappa_\varepsilon := \varepsilon\delta/\tau$, and assume $d \ge 2^{22}\kappa_\varepsilon^2$. Consider any algorithm that adaptively issues at most T comparison queries, possibly using internal randomness, and outputs a point $\hat{x}$. If for every convex, L-Lipschitz, (ε/δ)-smooth $f : \mathbb{R}^d \to \mathbb{R}$ with $f(0) - \inf f \le 128\,\varepsilon\delta$ the output satisfies $\mathbb{E}\|\bar{\partial}_\delta f(\hat{x})\| \le \varepsilon$, then*

$$T \;\ge\; \frac{d^2\tau^2}{2^{25}\,c_\sigma\,\varepsilon^2\delta^2} \;=\; \frac{d^2}{2^{25}\,c_\sigma\,\kappa_\varepsilon^2} \;=\; \frac{d^2L^2}{2^{25}\,c_\sigma\,\kappa_\delta^2\varepsilon^2}, \qquad \kappa_\delta := \frac{L\delta}{\tau}.$$

The bound is for a fixed budget of T comparisons; it extends with a larger constant to a random number of comparisons with $\sup_v \mathbb{E}_v[N] \le T$, the budget model of Theorems 4.2 and 4.3 (Appendix A.6). On the larger, nonconvex class of L-Lipschitz (ε/δ)-smooth functions, an exact first-order oracle simulates a comparison with two queries, so the $\Omega(\Delta_f H/\varepsilon^2)$ bound of Carmon et al. (2020), in the globally Lipschitz form of Cutkosky et al. (2023, Appendix F), gives $T = \Omega(\Delta_f/(\delta\varepsilon))$ in high dimension; this is the only dependence on Δ_f we can prove.

Proof idea (details in Appendix A.6). Let ϕ be the Huber function with curvature $H := \varepsilon/\delta$ and slope $G := 32\varepsilon$, and for $v \in \{-a, a\}^d$ with $a = 16\delta/\sqrt{d}$ let $f_v(x) := \phi(\|x - v\|)$. Each f_v is convex, G-Lipschitz (hence L-Lipschitz), H-smooth, with $\|\nabla f_v(x)\| = \min\{H\|x - v\|, G\}$; the slope $\Theta(\varepsilon)$ rather than L yields the factor L^2/ε^2. Three steps.

Stationarity localizes v. Since $\|\nabla f_v(\hat{x})\| \le \|\bar{\partial}_\delta f_v(\hat{x})\| + H\delta$, the guarantee $\mathbb{E}\|\bar{\partial}_\delta f_v(\hat{x})\| \le \varepsilon$ forces $\mathbb{E}\min\{\|\hat{x} - v\|, G/H\} \le 2\delta$.

One comparison carries bounded information about v. For any query (x, y) and any coordinate j, let $v^{(j)}$ denote v with the jth sign flipped. Because $v \mapsto f_v(z)$ is L-Lipschitz and H-smooth, $\sum_{j=1}^d \big(\Delta_v(x,y) - \Delta_{v^{(j)}}(x,y)\big)^2 \le 64a^2L^2$, uniformly over the query. Since $\mathrm{KL}(\mathrm{Bern}(p)\|\mathrm{Bern}(p')) \le (p-p')^2/(p'(1-p'))$, one comparison contributes at most $c_\sigma(\Delta_v - \Delta_{v^{(j)}})^2/\tau^2$ to the divergence between the jth pair of hypotheses, and the chain rule for adaptive queries gives $\sum_j \mathrm{KL}(P_v\|P_{v^{(j)}}) \le 64c_\sigma TG^2a^2/\tau^2$ after T comparisons.

Assouad. With $a^2 := d\tau^2/(128c_\sigma TL^2)$ the average pairwise total variation is at most $1/2$, so any estimator of v has mean squared error at least $\frac{R^2}{2}(1 - s)$, where $R = a\sqrt{d}$ and $s^2 = 32c_\sigma TG^2R^2/(d^2\tau^2)$. The clipped output $\Pi_{[-a,a]^d}(\hat{x})$ has mean squared error at most $4R\delta$ by the first step; comparing the two bounds and optimizing R gives the claimed bound. The cube radius $a\sqrt{d} = 16\delta$ and the slope $G = 32\varepsilon$ optimize the resulting expression (Appendix A.6); the side conditions ($4Ga \le \tau$ and $2a\sqrt{d} \le G/H$) are where $d \ge 2^{22}\kappa_\varepsilon^2$ and $\varepsilon \le L/32$ enter, and $f_v(0) - \inf f_v = H\|v\|^2/2 = 128\varepsilon\delta$.

Relation to existing lower bounds. For the logistic link a comparison is $\mathbf{1}\{f(y)+\Gamma' > f(x)+\Gamma\}$ with Γ, Γ' i.i.d. Gumbel (Yellott, 1977), so Theorem 5.1 is the comparison-feedback analogue of the quadratic-in-d bound of Shamir (2013) for independently noised evaluations, with Goldstein stationarity as the target; Appendix A.7 relates it to Duchi et al. (2015); Shamir (2017); Jamieson et al. (2012). For $\tau \ge L\delta/\kappa^\star$, where Algorithm 1 uses $\rho = \delta/2$ and its κ equals κ_δ, the dominant term of Theorem 4.2 and the bound of Theorem 5.1 read $O(\mathbb{E}[N]C_{\mathrm{link}}d^2\tau^2\Delta_h/(\delta^3\varepsilon^3))$ and $\Omega(d^2\tau^2/(c_\sigma\delta^2\varepsilon^2))$: for a fixed link they agree in d, τ and δ/ε, and the remaining polynomial gap is $\Delta_h/(\delta\varepsilon)$. On the hard instances themselves ($\Delta_f = 128\varepsilon\delta$, (ε/δ)-smooth, so $\Delta_h = O(\varepsilon\delta)$) the

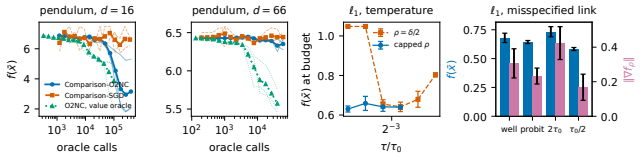

Figure 1: Left: pendulum swing-up ($\delta = 0.05$, $\kappa^\star = 1.5$), cost of the output policy against oracle calls for $d = 16$ ($1.2 \cdot 10^5$ iterations for comparisons, $5 \cdot 10^4$ for the value oracle) and $d = 66$ ($2 \cdot 10^4$ iterations); thin lines are the 3 seeds. Right: ℓ_1 regression, $d = 49$, Comparison-O2NC with the step sizes selected at τ_0, loss after $2 \cdot 10^5$ comparisons against the temperature with $\rho = \delta/2$ (nominal κ from 384 to 0.375) and with the capped rule of Theorem 4.2; and loss and gradient norm under a misspecified link (matched-slope probit, true temperature $2\tau_0$ or $\tau_0/2$) while the estimator assumes the logistic link with τ_0.

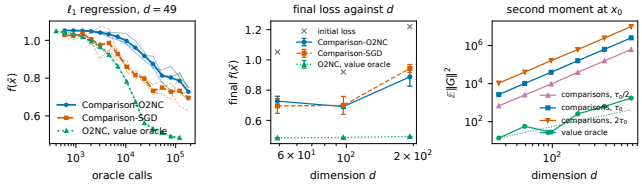

Figure 2: ℓ_1 regression with a ReLU model on the diabetes data ($\delta = 0.05$, $\kappa^\star = 1.5$, 3 seeds). Left: loss against oracle calls for $d = 49$. Middle: final loss after $6 \cdot 10^4$ iterations against d. Right: $\mathbb{E}\|G\|^2$ at the initial point against d with the temperature fixed across d (dotted lines $\propto d$ and $\propto d^2$).

two agree up to a link-dependent factor for bounded κ, under the high-dimensional side condition. Appendix A.6 records why a chain of such instances does not close the gap for larger Δ_f, and why value oracles escape the bound.

6 Experiments

We test the estimator, Lemma 4.1, the d^2/T rate behind Theorem 5.1, and the two algorithms on synthetic functions (Appendix B) and on two small tasks: preference-based policy search on a pendulum swing-up and ℓ_1 regression with a ReLU model on the diabetes data. Comparisons are simulated from the true objective values through the logistic link, and step sizes are selected by grid search per method (Appendix B). In the synthetic experiments both estimators are unbiased at every κ; the run-length estimator has the smaller second moment and the early-stopped nested estimator the smaller cost and cost–second-moment product, both minimized near $\kappa^\star \approx 1.5$ (Table 2); $\mathbb{E}\|G\|^2$ grows as d with values and as d^2 with comparisons (Figure 4); the maximum-likelihood estimate of v on the instance of Theorem 5.1 has mean squared error matching $d^2\tau^2/(TH^2)$ within 10% (Figure 5); and on a Chebyshev–Rosenbrock function Comparison-O2NC reaches a gradient norm of 0.87 ± 0.06 against 0.08 ± 0.03 for the value oracle with ten times fewer calls (Figure 6).

Pendulum (Figure 1). The objective is the average per-step cost $\theta^2 + 0.1\dot{\theta}^2 + 0.001u^2$ of a deterministic pendulum over 8 initial states and 100 steps, as a function of the parameters of a one-hidden-layer ReLU policy with clipped torque: nonsmooth, nonconvex, deterministic, with L replaced by a finite-difference estimate of the local Lipschitz scale ($L \approx 9$ for $d = 16$; the theorems assume a global constant). Two policies are compared through the logistic link applied to their costs, as for trajectory preferences in preference-based reinforcement learning. With $d = 16$, Comparison-O2NC swings the pendulum up from all initial states on two of three seeds ($4 \cdot 10^5$ comparisons; cost $6.9 \to 2.0$ and 2.1, 5.3 on the third), the value oracle on one of three (10^5 evaluations), and Comparison-SGD stays between 6.2 and 6.8. The comparison-based runs use about four times more oracle calls for a comparable outcome, far below the ratio of about 260 between the second moments at the initial policy, because the queried gaps are only about $0.3\,\tau_0$; with $d = 66$ neither oracle solves the task within $2 \cdot 10^4$ iterations.

ℓ_1 **regression (Figure 2).** The objective is the mean absolute error of a one-hidden-layer ReLU regressor on the standardized diabetes data (442 samples, 10 features), a nonsmooth nonconvex function of $d = 12w + 1$ parameters for width w; cheap evaluations allow longer runs and a sweep over d. For $d = 49$ Comparison-O2NC reduces the loss from 1.05 to 0.73 ± 0.03 and Comparison-SGD to 0.70 ± 0.05 with $1.8 \cdot 10^5$ comparisons, against 0.484 ± 0.001 for the value oracle with $1.2 \cdot 10^5$ evaluations. At a fixed number of iterations the comparison-based methods stay above the value oracle for $d \in \{49, 97, 193\}$ (middle panel: 0.73, 0.69, 0.89 against 0.48 to 0.49). The right panel isolates the cause: at fixed temperature $\mathbb{E}\|G\|^2$ for comparisons grows by a factor 951 from $d = 25$ to 769, against 946 for d^2, while the value oracle grows much more slowly. The nominal $\kappa = 1.5$ is a worst case: the largest gap over 4000 random directions is 0.1 to 0.4 of τ_0, so in the linear regime of the link, where $\mathbb{E}[\widetilde{\Delta}^2] \approx 2\tau^2$, a temperature smaller by that factor would reduce the 260-fold second-moment ratio at $d = 49$ by its square; the baseline study uses $\tau_0/32$.

Temperature and misspecification (Figure 1, right). At a budget of $2 \cdot 10^5$ comparisons, $\rho = \delta/2$ works while comparisons stay informative and collapses once they become nearly deterministic: at $\tau_0/64$ an estimate costs about $3 \cdot 10^4$ comparisons, the budget covers seven iterations, and the loss stays at its initial 1.05, whereas the capped rule gives 0.63 to 0.66 for every $\tau \le \tau_0$. For $\tau > \tau_0$ the rules coincide and the loss degrades to 0.80 at $4\tau_0$. Misspecifying the link shape or the temperature by a factor two moves the loss by at most 0.1 from its well-specified value 0.68 (Appendix B).

Baselines (Appendix C). We compare against eight methods on the same oracle and budget, each tuned by grid search: sign-SGD on the two-point comparison (the update of Chayti et al., 2025; Saha et al., 2021), direct search with votes (Bergou et al., 2020; El Bakkali et al., 2026), the one-bit direction estimator of Cai et al. (2022), the ranking estimator of Tang et al. (2024), a biased plug-in in place of $\widetilde{\Delta}$, and O2NC with a value oracle. At $\tau = \tau_0/32$ the methods using a magnitude (ours, the plug-in, sign-SGD) reach $f \approx 0.35$ and $\|\nabla f_\rho(\bar{x})\| \in [0.23, 0.39]$ on Chebyshev–Rosenbrock, the direct searches stall at 0.43 with gradient norms above 1.1, and the direction-recovery methods do not converge. On ℓ_1 regression all reach 0.49 to 0.55, sign-SGD being best (0.493 against 0.503) with a radius $\rho = 2\delta$: a sign costs the same at large and small gaps, whereas the cost of $\widetilde{\Delta}$ grows as $e^{2L\rho/\tau}$. The plug-in is within 0.01 of $\widetilde{\Delta}$ on both tasks (Table 3): at these budgets the experiments suggest that magnitude information matters on Chebyshev–Rosenbrock while sign methods win on ℓ_1 regression, and show no practical advantage of unbiasedness, which is what the analysis requires.

7 Related work

For convex objectives, Jamieson et al. (2012) treat a gap-dependent reliability, Saha et al. (2021; 2025) online convex optimization with a link function, Cai et al. (2022) one-bit recovery of sparse gradient directions. For smooth nonconvex objectives, Bergou et al. (2020); El Bakkali & Saadi (2025) analyze direct search with exact comparisons, Tang et al. (2024) ranking oracles, Lobanov et al. (2024) accelerated methods, El Bakkali et al. (2026) tie-aware comparisons, Chen et al. (2025) preference alignment, and Chayti et al. (2025) stochastic random search, which under δ'-inexact comparisons converges up to a floor $O(\sqrt{d}\,\delta')$; these use smoothness to control the linearization error, while ours recovers the gap without bias, hence without a floor. For nonsmooth Lipschitz functions, (δ, ε)-Goldstein stationarity is the tractable target (Zhang et al., 2020; Kornowski & Shamir, 2022; Jordan et al., 2023); with value oracles Kornowski & Shamir (2023) reach $O(d\delta^{-1}\varepsilon^{-3})$ (see also Lin et al., 2022; Chen et al., 2023; Cutkosky et al., 2023), and lower bounds are per parameter (Duchi et al., 2015; Arjevani et al., 2023; Carmon et al., 2020), not joint.

8 Conclusion and limitations

We gave nonasymptotic guarantees for (δ, ε)-Goldstein stationarity from noisy comparisons: an $O(d(1 + dC_{\mathrm{link}}/\kappa^2)\delta^{-1}\varepsilon^{-3})$ upper bound and a lower bound $\Omega(d^2\tau^2/(\varepsilon^2\delta^2))$ tailored to the comparison model, matching the upper bound in d and τ on a small-initial-gap family and leaving the joint dependence on Δ_f open. Limitations: the link must be known and lens-analytic; the lower bound is informative only when $\tau \gtrsim L\delta$; the constants of Theorem 3.6 are loose; and sign-based updates, which carry no such guarantee, are empirically competitive at the budgets we could afford.

AI USE STATEMENT

We used a large language model as an assistant for restructuring and editing the manuscript, checking and tightening proofs, drafting parts of the experiment code, and suggesting related work. The authors verified all such output. The research ideas, the method and the main results are the authors' own, and the authors take full responsibility for the content of this work.

REPRODUCIBILITY STATEMENT

All assumptions are stated in Sections 2 and 3; complete proofs are in Appendix A. The experiments use only NumPy and are fully specified in Appendix B; the experiments are fairly simple and the code will be made public.

REFERENCES

Yossi Arjevani, Yair Carmon, John C. Duchi, Dylan J. Foster, Nathan Srebro, and Blake Woodworth. Lower bounds for non-convex stochastic optimization. *Mathematical Programming*, 199(1-2): 165–214, 2023.

Alex Beatson and Ryan P. Adams. Efficient optimization of loops and limits with randomized telescoping sums. In *International Conference on Machine Learning*. PMLR 97, 2019.

El Houcine Bergou, Eduard Gorbunov, and Peter Richtárik. Stochastic three points method for unconstrained smooth minimization. *SIAM Journal on Optimization*, 30(4):2726–2749, 2020.

Walter Bergweiler. Iteration of meromorphic functions. *Bulletin of the American Mathematical Society*, 29(2):151–188, 1993.

Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

HanQin Cai, Daniel McKenzie, Wotao Yin, and Zhenliang Zhang. A one-bit, comparison-based gradient estimator. *Applied and Computational Harmonic Analysis*, 60:242–266, 2022.

Yair Carmon, John C Duchi, Oliver Hinder, and Aaron Sidford. Lower bounds for finding stationary points I. *Mathematical Programming*, 184:71–120, 2020.

El Mahdi Chayti, Taha El Bakkali El Kadi, Omar Saadi, and Martin Jaggi. Stochastic optimization with random search. *arXiv preprint arXiv:2510.15610*, 2025.

Lesi Chen, Jing Xu, and Luo Luo. Faster gradient-free algorithms for nonsmooth nonconvex stochastic optimization. In *International Conference on Machine Learning*, pp. 5219–5233. PMLR, 2023.

Peter Chen, Xi Chen, Wotao Yin, and Tianyi Lin. ComPO: Preference alignment via comparison oracles. *arXiv preprint arXiv:2505.05465*, 2025.

Frank H. Clarke. *Optimization and Nonsmooth Analysis*. SIAM, 1990.

Ashok Cutkosky, Harsh Mehta, and Francesco Orabona. Optimal stochastic non-smooth non-convex optimization through online-to-non-convex conversion. In *International Conference on Machine Learning*, pp. 6643–6670. PMLR, 2023.

John C. Duchi, Michael I. Jordan, Martin J. Wainwright, and Andre Wibisono. Optimal rates for zero-order convex optimization: The power of two function evaluations. *IEEE Transactions on Information Theory*, 61(5):2788–2806, 2015.

Taha El Bakkali and Omar Saadi. On the almost sure convergence of the stochastic three points algorithm. In *The Thirteenth International Conference on Learning Representations*, 2025.

Taha El Bakkali, Rayane Bouftini, Qiuyi Zhang, and Omar Saadi. Noisy pairwise-comparison random search for smooth nonconvex optimization. *arXiv preprint arXiv:2601.21166*, 2026.

A. A. Goldstein. Optimization of Lipschitz continuous functions. *Mathematical Programming*, 13 (1):14–22, 1977.

Kevin G. Jamieson, Robert Nowak, and Ben Recht. Query complexity of derivative-free optimization. In *Advances in Neural Information Processing Systems*, volume 25, 2012.

Michael Jordan, Guy Kornowski, Tianyi Lin, Ohad Shamir, and Manolis Zampetakis. Deterministic nonsmooth nonconvex optimization. In *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pp. 4570–4597. PMLR, 2023.

Guy Kornowski and Ohad Shamir. Oracle complexity in nonsmooth nonconvex optimization. *Journal of Machine Learning Research*, 23(314):1–44, 2022.

Guy Kornowski and Ohad Shamir. An algorithm with optimal dimension-dependence for zero-order nonsmooth nonconvex stochastic optimization. In *Advances in Neural Information Processing Systems*, volume 36, 2023. arXiv:2307.04504.

Tianyi Lin, Zeyu Zheng, and Michael Jordan. Gradient-free methods for deterministic and stochastic nonsmooth nonconvex optimization. In *Advances in Neural Information Processing Systems*, volume 35, pp. 26160–26175, 2022.

Aleksandr Lobanov, Alexander Gasnikov, and Andrei Krasnov. Acceleration exists! Optimization problems when oracle can only compare objective function values. In *Advances in Neural Information Processing Systems*, volume 37, pp. 21428–21460, 2024.

Anne-Marie Lyne, Mark Girolami, Yves Atchadé, Heiko Strathmann, and Daniel Simpson. On Russian roulette estimates for Bayesian inference with doubly-intractable likelihoods. *Statistical Science*, 30(4):443–467, 2015.

Don McLeish. A general method for debiasing a Monte Carlo estimator. *arXiv:1005.2228*, 2010.

Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark (eds.). *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, 2010.

Chang-Han Rhee and Peter W. Glynn. Unbiased estimation with square root convergence for SDE models. *Operations Research*, 63(5):1026–1043, 2015.

Aadirupa Saha, Tomer Koren, and Yishay Mansour. Dueling convex optimization. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 9245–9254. PMLR, 2021.

Aadirupa Saha, Tomer Koren, and Yishay Mansour. Dueling convex optimization with general preferences. In *Forty-second International Conference on Machine Learning*, 2025.

Ohad Shamir. On the complexity of bandit and derivative-free stochastic convex optimization. In *Proceedings of the 26th Annual Conference on Learning Theory*, volume 30 of *Proceedings of Machine Learning Research*, pp. 3–24. PMLR, 2013.

Ohad Shamir. An optimal algorithm for bandit and zero-order convex optimization with two-point feedback. *Journal of Machine Learning Research*, 18(1):1703–1713, 2017.

Zhiwei Tang, Dmitry Rybin, and Tsung-Hui Chang. Zeroth-order optimization meets human feedback: Provable learning via ranking oracles. In *The Twelfth International Conference on Learning Representations*, 2024.

John I. Yellott. The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, 15: 109–144, 1977.

Farzad Yousefian, Angelia Nedić, and Uday V. Shanbhag. On stochastic gradient and subgradient methods with adaptive steplength sequences. *Automatica*, 48(1):56–67, 2012.

Jingzhao Zhang, Hongzhou Lin, Stefanie Jegelka, Ali Jadbabaie, and Suvrit Sra. Complexity of finding stationary points of nonsmooth nonconvex functions. In *International Conference on Machine Learning*, pp. 11173–11182. PMLR, 2020.

A PROOFS

A.1 PROOF OF PROPOSITION 3.3

Let U be an open neighborhood of Λ_r on which g is holomorphic. Since $r > 1/2$, the circles $\{|z| = r\}$ and $\{|z-1| = r\}$ intersect at $\frac{1}{2} \pm i\sqrt{r^2 - 1/4}$, and the boundary of Λ_r consists of the arc $\Gamma_0 \subset \{|z| = r\}$ with $|z-1| \le r$ and the arc $\Gamma_1 \subset \{|z-1| = r\}$ with $|z| \le r$. For z in the interior Λ_r°, Cauchy's integral formula gives

$$g(z) = \frac{1}{2\pi i} \oint_{\partial\Lambda_r} \frac{g(\zeta)}{\zeta - z}\, d\zeta = I_0(z) + I_1(z), \qquad I_j(z) := \frac{1}{2\pi i} \int_{\Gamma_j} \frac{g(\zeta)}{\zeta - z}\, d\zeta.$$

The function I_0 is holomorphic on $\mathbb{C} \setminus \Gamma_0$, which contains the open disc $D_r := \{|z| < r\}$, and I_1 is holomorphic on $\mathbb{C} \setminus \Gamma_1 \supseteq 1 - D_r$. Set $k(z) := I_1(1-z)$, holomorphic on D_r. The set Λ_r° is connected and invariant under $z \mapsto 1 - z$, and $g(z) = -g(1-z)$ holds on the real segment $(1-r, r) \subset \Lambda_r^\circ$, hence on all of Λ_r° by the identity theorem. For $z \in \Lambda_r^\circ$ we therefore have both $g(z) = I_0(z) + k(1-z)$ and $g(z) = -g(1-z) = -I_0(1-z) - k(z)$. Averaging the two identities,

$$g(z) = h(z) - h(1-z), \qquad h := \tfrac{1}{2}(I_0 - k),$$

with h holomorphic on D_r. Write $h(z) = \sum_{m \ge 0} c_m z^m$ on D_r. For $r' < r$ the Cauchy estimates give $|c_m| \le C_{r'}(r')^{-m}$ with $C_{r'} = \max_{|z|=r'} |h(z)|$. For $p \in [1-\alpha, \alpha]$ we have $p, 1-p \in [0, \alpha] \subset [0, r')$, so both $\sum_m c_m p^m$ and $\sum_m c_m (1-p)^m$ converge absolutely and uniformly on the interval, the constant terms cancel in the difference, and $g(p) = \sum_{m \ge 1} c_m \big(p^m - (1-p)^m\big)$. □

The coefficients can be taken real: g is real on the real segment of the lens, so by Schwarz reflection $g(\bar z) = \overline{g(z)}$ on the conjugation-symmetric set Λ_r°, and $h_{\mathbb{R}}(z) := \frac{1}{2}\big(h(z) + \overline{h(\bar z)}\big)$ is holomorphic on D_r with real Taylor coefficients and satisfies $h_{\mathbb{R}}(z) - h_{\mathbb{R}}(1-z) = \frac{1}{2}\big(g(z) + \overline{g(\bar z)}\big) = g(z)$, using $\overline{1 - \bar z} = 1 - z$. We use $h_{\mathbb{R}}$ below, so the c_m and the estimator are real.

A link with an analytic but not lens-analytic inverse: $g(p) = \operatorname{artanh}(2p-1) + \epsilon\,(p - \frac{1}{2})/((p - \frac{1}{2})^2 + \eta^2)$ with ϵ/η^2 small enough that $g' > 0$ on $(0,1)$ and $\eta < \sqrt{\alpha^2 - 1/4}$; g is odd about $1/2$, strictly increasing, and analytic on a neighborhood of $[1-\alpha, \alpha]$, but it has poles at $\frac{1}{2} \pm i\eta$ inside the lens $\{|z| < \alpha\} \cap \{|z-1| < \alpha\}$, so no expansion with $\sum_m |c_m| \alpha^m < \infty$ exists.

The probit link. Example 3.4 asserts that Φ^{-1} continues analytically to any simply connected domain avoiding 0 and 1. Two facts are used. First, for an entire function F let S be the closure of the set of its critical values and of its finite asymptotic values (limits of F along paths to ∞); then $F : \mathbb{C} \setminus F^{-1}(S) \to \mathbb{C} \setminus S$ is a covering map (Bergweiler, 1993, Section 2), so every local inverse of F continues analytically along every path in $\mathbb{C} \setminus S$, and by the monodromy theorem single-valuedly on every simply connected domain in $\mathbb{C} \setminus S$. Second, for $F = \Phi$ one has $S = \{0, 1\}$. There are no critical values since $\Phi' = \varphi \ne 0$. For the asymptotic values write $\Phi(z) = 1 - \frac{1}{2}\operatorname{erfc}(w)$ with $w = z/\sqrt{2}$, and recall $\operatorname{erfc}(w) = \frac{e^{-w^2}}{w\sqrt{\pi}}(1 + O(|w|^{-2}))$ as $|w| \to \infty$ in $|\arg w| \le \frac{3\pi}{4} - \epsilon$ (Olver et al., 2010, §7.12), with $\operatorname{erfc}(-w) = 2 - \operatorname{erfc}(w)$ covering the remaining directions. Let $w(t) \to \infty$ be a path along which Φ converges to a finite value. On the tail of the path $|w|$ is large, so $\operatorname{Re}(w^2) = -|w|^2$ on the imaginary axis makes $|\Phi| \to \infty$ there; the tail therefore stays in one open half-plane, say $\operatorname{Re} w > 0$, where $\operatorname{erfc}(w) = e^{-(w^2 + \log w)}(1 + o(1))/\sqrt{\pi}$. If $\operatorname{Re}(w^2) + \log|w| \to +\infty$ then $\operatorname{erfc}(w) \to 0$ and $\Phi \to 1$. Otherwise $\operatorname{Re}(w^2) + \log|w|$ is bounded above along a subsequence, and on the tail it is also bounded below, since $|\operatorname{erfc}(w)|$ is bounded; then $|w^2| = |w|^2 \to \infty$ forces $|\operatorname{Im}(w^2)| \to \infty$, and $\operatorname{Im}(w^2)$ cannot change sign for large $|w|$ (at a sign change $|w|^2 = |\operatorname{Re}(w^2)| = O(\log|w|)$), so the argument of e^{-w^2} winds without bound while its modulus stays of order one, and $\operatorname{erfc}(w)$ has no limit, a contradiction. In the left half-plane the same argument gives $\Phi \to 0$. Hence the finite asymptotic values are exactly 0 and 1.

For the logistic link, $h(z) = -\log(1-z)$ satisfies $h(p) - h(1-p) = \log\frac{p}{1-p}$, is holomorphic on D_1, and has $c_m = 1/m$; the decomposition of the proof produces an h that may differ from this one by a function k_0 with $k_0(z) = k_0(1-z)$, which changes the coefficients but not the identity $g = \sum_m c_m b_m$ on the operated range.

A.2 Proof of Theorem 3.6 and of Remark A.1

Let $K := N - 1 \geq 1$ be the length of the initial run of equal outcomes, so that $A_m - B_m = (2Y_1 - 1)\mathbf{1}\{m \leq K\}$ and $\widetilde{\Delta} = \tau(2Y_1 - 1)S_K = \tau \sum_{m\geq 1} c_m(A_m - B_m)$. Conditionally on $Y_1 = 1$, K is geometric with $\mathbb{P}(K \geq k \mid Y_1 = 1) = p^{k-1}$; conditionally on $Y_1 = 0$, $\mathbb{P}(K \geq k \mid Y_1 = 0) = (1-p)^{k-1}$. In both cases the continuation probability $q \in \{p, 1-p\}$ satisfies $q \leq \alpha$.

(i) Unbiasedness. $\mathbb{E}|c_m(A_m - B_m)| = |c_m|(p^m + (1-p)^m) \leq 2|c_m|\alpha^m \leq 2C(\alpha\rho_c)^m$, which is summable since $\alpha\rho_c < 1$. By Fubini's theorem, $\mathbb{E}[\widetilde{\Delta}] = \tau \sum_{m\geq 1} c_m \mathbb{E}[A_m - B_m] = \tau \sum_{m\geq 1} c_m b_m(p) = \tau g(p) = \Delta$.

(ii) Cost. For $n \geq 1$, $\{N > n\}$ is the event that the first n outcomes are equal, so $\mathbb{P}(N > n) = p^n + (1-p)^n$, and $\mathbb{P}(N > 0) = 1$. Hence $\mathbb{E}[N] = \sum_{n\geq 0} \mathbb{P}(N > n) = 1 + \sum_{n\geq 1}(p^n + (1-p)^n) = 1 + \frac{p}{1-p} + \frac{1-p}{p}$. On $[1-\alpha, \alpha]$ this is maximized at the endpoints, where it equals $1 + \frac{\alpha}{1-\alpha} + \frac{1-\alpha}{\alpha} \leq 1 + \frac{\alpha}{1-\alpha} + 1 = 1 + \frac{1}{1-\alpha}$.

(iii) Second moment. Since $\mathbb{P}(K = k) = p^k(1-p) + (1-p)^k p$ for $k \geq 1$,

$$\mathbb{E}[\widetilde{\Delta}^2] = \tau^2 \sum_{k\geq 1} \big(p^k(1-p) + (1-p)^k p\big) S_k^2, \tag{5}$$

which for the logistic link ($S_k = H_k$) is the exact identity of the theorem. In general $|S_k| \leq \sum_{m\leq k} |c_m| \leq C \sum_{m\leq k} \rho_c^m \leq Ck\bar{\rho}^k$, and $p^k(1-p) + (1-p)^k p \leq 2\alpha^k$, so

$$\mathbb{E}[\widetilde{\Delta}^2] \leq 2\tau^2 C^2 \sum_{k\geq 1} k^2(\alpha\bar{\rho}^2)^k = \frac{2C^2\tau^2\theta(1+\theta)}{(1-\theta)^3}, \qquad \theta = \alpha\bar{\rho}^2,$$

using $\sum_{k\geq 1} k^2\theta^k = \theta(1+\theta)/(1-\theta)^3$. For $\rho_c = 1/r'$ with $r' \in (\sqrt{\alpha}, 1)$ one has $\bar{\rho} = 1/r'$ and $\theta = \alpha/r'^2 < 1$.

For the logistic link, $S_k = H_k \leq 1 + \log k$, so $\mathbb{E}[\widetilde{\Delta}^2] \leq \tau^2\mathbb{E}[(1 + \log K)^2] \leq \tau^2(2 + 2\mathbb{E}[\log^2 K])$. We bound $\mathbb{E}[\log^2 K]$ for a geometric K with $\mathbb{P}(K \geq k) = q^{k-1}$, $q \leq \alpha$. Let $k_0 := \lceil 1/(1-\alpha) \rceil$. Since $\log\alpha \leq -(1-\alpha)$, $\mathbb{P}(K > jk_0) = q^{jk_0} \leq \alpha^{j/(1-\alpha)} \leq e^{-j}$ for every $j \geq 1$. Splitting $\{K \leq k_0\}$ from the blocks $\{jk_0 < K \leq (j+1)k_0\}$ and using $\log^2((j+1)k_0) \leq 2\log^2(j+1) + 2\log^2 k_0$,

$$\begin{aligned}\mathbb{E}[\log^2 K] &\leq \log^2 k_0 + \sum_{j\geq 1} e^{-j}\big(2\log^2(j+1) + 2\log^2 k_0\big) \\ &\leq \Big(1 + \frac{2}{e-1}\Big)\log^2 k_0 + 1.1 \leq 2.2\log^2 k_0 + 1.1,\end{aligned}$$

where $2\sum_{j\geq 1} e^{-j}\log^2(j+1) < 1.1$ by direct summation. For the logistic link $1/(1-\alpha) = 1 + e^{\kappa} \leq 2e^{\kappa}$, so $k_0 \leq 2/(1-\alpha) \leq 4e^{\kappa}$ and $\log k_0 \leq \kappa + \log 4 \leq 1.39(1+\kappa)$. Therefore $\mathbb{E}[\widetilde{\Delta}^2] \leq \tau^2\big(2 + 2(2.2 \cdot 1.39^2(1+\kappa)^2 + 1.1)\big) \leq \tau^2\big(4.2 + 8.6(1+\kappa)^2\big) \leq 13\,\tau^2(1+\kappa)^2$.

(iv) Moments. $\mathbb{E}|\widetilde{\Delta}|^q = \tau^q \sum_k \mathbb{P}(K = k)|S_k|^q \leq 2\tau^q C^q \sum_k k^q(\alpha\bar{\rho}^q)^k$, finite when $\alpha\bar{\rho}^q < 1$. For the logistic link $|S_k| \leq 1 + \log k$ and $\mathbb{P}(K = k) \leq 2\alpha^k$, so every moment is finite. □

Remark A.1 (Random truncation)**.** The randomly truncated series $\widehat{\Delta}_{\mathrm{RR}} := \tau \sum_{m\geq 1} c_m(A_m - B_m)\mathbf{1}\{M \geq m\}/\mathbb{P}(M \geq m)$, with M an independent geometric variable of parameter $1 - \beta$ (the Russian-roulette device of McLeish, 2010; Rhee & Glynn, 2015; Lyne et al., 2015; Beatson & Adams, 2019), is also unbiased, and $\mathbb{E}[\widehat{\Delta}_{\mathrm{RR}} \mid Y_1, Y_2, \dots] = \widetilde{\Delta}$, so $\mathrm{Var}(\widetilde{\Delta}) \leq \mathrm{Var}(\widehat{\Delta}_{\mathrm{RR}})$ whenever the latter is finite. Its cost is $\min\{M, N\}$ comparisons, since the stream may be stopped at the first outcome change, so it is cheaper than $\widetilde{\Delta}$; the weights $1/\mathbb{P}(M \geq m)$ give it heavy tails. Appendix A.3 works this out; at the settings of Table 2 it uses about 70% of the comparisons for a second moment about 10% larger.

A.3 THE RUSSIAN-ROULETTE ESTIMATOR

This section proves Remark A.1 and records the properties of the randomly truncated alternative to $\widetilde{\Delta}$, which is the estimator a reader familiar with unbiased truncation (McLeish, 2010; Rhee & Glynn, 2015; Lyne et al., 2015; Beatson & Adams, 2019) would write down first.

Definition. Let $M \geq 1$ be independent of the comparison stream with $q_m := \mathbb{P}(M \geq m) > 0$ for every m, and set

$$\widehat{\Delta}_{\mathrm{RR}} := \tau \sum_{m \geq 1} c_m \, (A_m - B_m) \, \frac{\mathbf{1}\{M \geq m\}}{q_m}.$$

The sum is finite: it has at most $\min\{M, N-1\}$ nonzero terms, since $A_m = B_m = 0$ for $m \geq N$. We call $q_m = \beta^{m-1}$, that is M geometric with parameter $1-\beta$, the geometric truncation, and we write $\alpha := \sigma(\kappa)$ for the largest success probability in the operated range.

Unbiasedness and the Rao–Blackwell relation. Conditionally on the stream $Y_1, Y_2, \ldots$, each weight has $\mathbb{E}[\mathbf{1}\{M \geq m\}]/q_m = 1$, so

$$\mathbb{E}\big[\widehat{\Delta}_{\mathrm{RR}} \mid Y_1, Y_2, \ldots\big] = \tau \sum_{m \geq 1} c_m (A_m - B_m) = \widetilde{\Delta}, \tag{6}$$

provided the series is absolutely convergent, which holds under Assumption 3.2 as in the proof of Theorem 3.6(i). Taking expectations gives $\mathbb{E}[\widehat{\Delta}_{\mathrm{RR}}] = \Delta$, and the law of total variance gives

$$\mathrm{Var}(\widehat{\Delta}_{\mathrm{RR}}) = \mathrm{Var}(\widetilde{\Delta}) + \mathbb{E}\big[\mathrm{Var}(\widehat{\Delta}_{\mathrm{RR}} \mid Y_1, Y_2, \ldots)\big] \;\geq\; \mathrm{Var}(\widetilde{\Delta})$$

whenever the left-hand side is finite. So the run-length estimator is the Rao–Blackwellization of the truncated one with respect to the comparison stream, and can only reduce variance. It is not, however, uniformly better, because the two differ in cost.

Cost. Computing $\widehat{\Delta}_{\mathrm{RR}}$ requires the outcomes $Y_1, \ldots, Y_m$ only for those m with $M \geq m$ and $A_m - B_m \neq 0$, so the stream may be stopped at the first outcome change or at the truncation time, whichever comes first:

$$N_{\mathrm{RR}} = \min\{M,\, N\}, \qquad \mathbb{E}[N_{\mathrm{RR}}] \leq \min\{\mathbb{E}[M], \mathbb{E}[N]\} = \min\Big\{\tfrac{1}{1-\beta},\, 1 + \tfrac{p}{1-p} + \tfrac{1-p}{p}\Big\}. \tag{7}$$

Nested Russian roulette is therefore never more expensive than the run-length estimator and is typically cheaper, which is the reverse of the variance ordering. At $\beta = \sqrt{\alpha}$ and $p = \alpha$ the two effects nearly cancel: Table 2 reports a cost ratio of about 0.7 and a second-moment ratio of about 1.1, so the product that governs the complexity in Theorem 4.2 favors the truncated estimator by 20 to 30% over the range $\kappa \in [0.25, 3]$.

Second moment. With one stream for all m (the nested version) and the identity $(A_m - B_m)(A_n - B_n) = A_{m \vee n} + B_{m \vee n}$,

$$\mathbb{E}[\widehat{\Delta}_{\mathrm{RR}}^2] = \tau^2 \sum_{m,n \geq 1} \frac{c_m c_n}{q_{m \wedge n}} \big(p^{m \vee n} + (1-p)^{m \vee n}\big),$$

since $\mathbb{E}[\mathbf{1}\{M \geq m\}\mathbf{1}\{M \geq n\}]/(q_m q_n) = q_{m \vee n}/(q_m q_n) = 1/q_{m \wedge n}$. For the logistic link and geometric truncation this is $\tau^2 \sum_{m,n \geq 1} \frac{\beta^{1-(m \wedge n)}}{mn} (p^{m \vee n} + (1-p)^{m \vee n})$, which converges if and only if $\beta \geq \alpha$ at $p = \alpha$. Indeed, for $\beta < \alpha$ the diagonal terms diverge, whereas at $\beta = \alpha$, writing $n = m + k$, the summands are bounded by $2\alpha^{k+1}/m^2$, which is summable over $m \geq 1$ and $k \geq 0$. Drawing independent blocks of sizes $1, \ldots, M$ instead of reusing one stream gives the same expectation and, for nonnegative coefficients such as the logistic $c_m = 1/m$, a smaller second moment, at a cost of $M(M+1)/2$ comparisons; its product of cost and second moment is larger at every κ we tested.

Tails. The weights $1/q_m$ are what make the truncated estimator heavy-tailed. For the geometric truncation, on the event $\{M = m,\ A_m = 1\}$, of probability $(1-\beta)\beta^{m-1}p^m$, one has $\widehat{\Delta}_{\mathrm{RR}} \geq \tau\beta^{1-m}/m$, so

$$\mathbb{P}\Big(\widehat{\Delta}_{\mathrm{RR}} \geq \frac{\tau\beta^{1-m}}{m}\Big) \ \geq\ (1-\beta)(\beta p)^{m-1}p,$$

and $\mathbb{E}|\widehat{\Delta}_{\mathrm{RR}}|^q = \infty$ as soon as $q > 1 + \log(1/\alpha)/\log(1/\beta)$. In particular only finitely many moments exist, and the number of them is governed by the same ratio that controls convergence of the second moment. By contrast $\widetilde{\Delta} = \tau(2Y_1 - 1)S_{N-1}$ with $N-1$ geometric, so its qth moment is finite whenever $\alpha\bar{\rho}^q < 1$ (Theorem 3.6(iv)); since $\bar{\rho} = \rho_c > 1$ for every lens-analytic link (the lens radius is below 1), this covers only finitely many q in general, while for the logistic link $|S_k| = H_k = O(\log k)$ and all moments are finite. This matters for high-probability guarantees and for any use of clipping, not for Theorem 4.2, which needs only a second moment.

Which to use. Either estimator can be substituted into Lemma 4.1 and hence into Theorems 4.2 and 4.3: the analysis uses only unbiasedness, a bound on the second moment, and the expected number of comparisons. We state the results for $\widetilde{\Delta}$ because it has no truncation parameter to tune, its second moment admits the closed form (5), and, for the logistic link, all its moments are finite; a practitioner willing to tune β and to accept heavy tails can use the truncated version and expect a constant-factor saving of the size reported in Table 2.

A.4 The online-to-nonconvex conversion

We use the guarantee of Cutkosky et al. (2023) in the form stated by Kornowski & Shamir (2023, Theorem 8).

Theorem A.2 (Cutkosky et al., 2023)**.** *Let $\nu, \varepsilon \in (0,1)$ and let $h : \mathbb{R}^d \to \mathbb{R}$ be differentiable and L-Lipschitz with $h(x_0) - \inf h \leq \Delta_h$. Suppose that at every query point the gradient oracle returns an unbiased estimate of ∇h with second moment at most σ^2, independently across calls. Then there is $T = O(\sigma^2\Delta_h/(\nu\varepsilon^3))$ such that Algorithm 1 with $\eta = \Delta_h/(\sigma^2 T)$ and $D = (\nu^{1/2}\Delta_h/(\sigma T))^{2/3}$ satisfies $z_{(k-1)M+m} \in \mathbb{B}(\bar{x}_k, \nu)$ for all $m \leq M$, $k \leq K$, and*

$$\mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K}\Big\|\frac{1}{M}\sum_{m=1}^{M}\nabla h\big(z_{(k-1)M+m}\big)\Big\|\right] \leq \varepsilon.$$

In particular $\mathbb{E}\big\|\bar{\partial}_\nu h(x^{\mathrm{out}})\big\| \leq \varepsilon$ for $x^{\mathrm{out}} \sim \mathrm{Unif}\{\bar{x}_1, \dots, \bar{x}_K\}$.

Proof of Theorem 4.2. By Lemma 4.1, G is an unbiased estimator of ∇f_ρ with second moment at most σ_G^2, and f_ρ is L-Lipschitz with $f_\rho(x_0) - \inf f_\rho \leq \Delta_h$ (for instance $\Delta_h = \Delta_f + L\rho$, since $\|f_\rho - f\|_\infty \leq L\rho$ and $\inf f_\rho \geq \inf f$; or $\Delta_f + H\rho^2$ when f is H-smooth). The guarantee of the online-to-nonconvex conversion (Kornowski & Shamir, 2023, Theorem 8), restated as Theorem A.2 in Appendix A.4, gives $\mathbb{E}\big\|\bar{\partial}_\nu f_\rho(x^{\mathrm{out}})\big\| \leq \varepsilon$ after $T = O(\sigma_G^2\Delta_h/(\nu\varepsilon^3)) = O(\sigma_G^2\Delta_h/(\delta\varepsilon^3))$ steps, using $\nu \geq \delta/2$. Since $\partial_\nu f_\rho(x) \subseteq \partial_{\rho+\nu} f(x) = \partial_\delta f(x)$, the same bound holds for $\big\|\bar{\partial}_\delta f(x^{\mathrm{out}})\big\|$. The expected number of comparisons is bounded as described below. □

In Theorem 4.2 the oracle is G of (4) with $h = f_\rho$, which is differentiable and L-Lipschitz; the randomness of each call consists of u_t and the N_t comparison outcomes, all independent across t and of the past, so conditionally on the past the estimates are independent and unbiased, which is the oracle model of the conversion. We take T large enough that $D \leq \nu$, so that $M \geq 1$ and $K \geq 1$ in Algorithm 1. Conditionally on u_t the count N_t has mean $1 + \frac{p_t}{1-p_t} + \frac{1-p_t}{p_t} \leq 1 + \frac{1}{1-\alpha}$ by Theorem 3.6(ii), and T is deterministic, so $\mathbb{E}\sum_{t\leq T} N_t \leq T(1 + \frac{1}{1-\alpha})$.

A.5 Proof of Theorem 4.3

Let $F := f_\delta$, which is $H_\delta := c_2\sqrt{d}L/\delta$-smooth and satisfies $F(x_0) - \inf F \leq \Delta_f + L\delta =: \Delta_h$ (because $F \leq f + L\delta$ and $\inf F \geq \inf f$). For $\eta \leq 1/H_\delta$ and an unbiased gradient estimator with second moment at most σ_G^2, smoothness gives $\mathbb{E}[F(x_{t+1}) \mid x_t] \leq F(x_t) - \eta\|\nabla F(x_t)\|^2 + \frac{\eta^2 H_\delta}{2}\sigma_G^2$;

summing over $t < T$ and dividing by ηT,

$$\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\|\nabla F(x_t)\|^2 \le \frac{\Delta_h}{\eta T} + \frac{\eta H_\delta \sigma_G^2}{2} \le \frac{2\Delta_h}{\eta T} + \eta H_\delta \sigma_G^2.$$

With $\eta = \min\{1/H_\delta,\ \eta^\star\}$ and $\eta^\star := \sqrt{2\Delta_h/(H_\delta\sigma_G^2 T)}$ the right-hand side is at most $2\sqrt{2H_\delta\Delta_h\sigma_G^2/T} + 2\Delta_h H_\delta/T$. For R uniform on $\{0,\dots,T-1\}$, Jensen's inequality gives $\mathbb{E}\|\nabla F(x_R)\| \le \varepsilon$ as soon as $T \ge 32H_\delta\Delta_h\sigma_G^2/\varepsilon^4$ and $T \ge 4\Delta_h H_\delta/\varepsilon^2$; the first condition implies the second because $\sigma_G^2 \ge c_1 dL^2 \ge L^2 \ge \varepsilon^2$ for $\varepsilon \le L$. Finally $\nabla f_\delta(x) \in \partial_\delta f(x)$ (Lin et al., 2022), so $\|\bar\partial_\delta f(x_R)\| \le \|\nabla f_\delta(x_R)\|$. The count of comparisons is as in Appendix A.4. □

A.6 Proof of Theorem 5.1

Hard instances. Let $H := \varepsilon/\delta$ and $G := 32\varepsilon$, and let $\phi(r) := \frac{H}{2}r^2$ for $0 \le r \le G/H = 32\delta$ and $\phi(r) := Gr - \frac{G^2}{2H}$ for $r > G/H$ (a Huber function of slope G). For $v \in \mathbb{R}^d$ put $f_v(x) := \phi(\|x - v\|)$. Then f_v is convex, G-Lipschitz, continuously differentiable, H-smooth, minimized at v, and $\|\nabla f_v(x)\| = \min\{H\|x - v\|, G\}$. Since $\varepsilon \le L/32$ we have $G \le L$, so every f_v belongs to the class of L-Lipschitz H-smooth convex functions. The parameter set is the cube $\{-a, a\}^d$, whose half-diagonal $R := a\sqrt{d}$ we keep symbolic for the moment and fix at the end:

$$a := \frac{R}{\sqrt{d}}, \qquad 2R \le \frac{G}{H}.$$

Because $\|v\| = R \le G/H$, the initial gap is $f_v(0) - \inf f_v = \phi(R) = \frac{H}{2}R^2$.

Taking the slope G proportional to ε rather than to L is what produces the factor L^2/ε^2 over the bound obtained with $G = L$: the sensitivity of a comparison to the hidden parameter is proportional to the slope, while the localization accuracy forced by Goldstein stationarity is not.

Step 1: stationarity localizes v. For an H-smooth function, every $g \in \partial_\delta f(x)$ is a convex combination of gradients at points within distance δ of x, each within $H\delta$ of $\nabla f(x)$; hence $\|\nabla f(x)\| \le \|\bar\partial_\delta f(x)\| + H\delta$. Applied to f_v at the output $\hat x$ and using $H\delta = \varepsilon$,

$$\min\{H\|\hat x - v\|,\, G\} \ \le\ \|\bar\partial_\delta f_v(\hat x)\| + \varepsilon.$$

Let $\hat v := \Pi_{[-a,a]^d}(\hat x)$ be the coordinatewise projection onto the cube. Since v lies in the cube, $\|\hat v - v\| \le \|\hat x - v\|$, and also $\|\hat v - v\| \le 2R \le G/H$; consequently $H\|\hat v - v\| \le \min\{H\|\hat x - v\|, G\}$. The hypothesis $\mathbb{E}\|\bar\partial_\delta f_v(\hat x)\| \le \varepsilon$ therefore gives $\mathbb{E}\|\hat v - v\| \le 2\varepsilon/H = 2\delta$ and

$$\mathbb{E}\|\hat v - v\|^2 \le 2R\,\mathbb{E}\|\hat v - v\| \le 4R\delta. \tag{8}$$

Step 2: information per comparison. Fix a query (x, y) and write $\psi_z(w) := \phi(\|z - w\|)$, so that $\Delta_v(x, y) = \psi_y(v) - \psi_x(v)$. The map $w \mapsto \psi_z(w)$ is G-Lipschitz, continuously differentiable and H-smooth. For $j \le d$ let $v^{(j)}$ be v with its jth coordinate negated, so $\|v - v^{(j)}\| = 2a$. By the fundamental theorem of calculus and H-smoothness of ψ_z,

$$\big|\psi_z(v^{(j)}) - \psi_z(v)\big| = 2a\Big|\int_0^1 \partial_j\psi_z\big(v + t(v^{(j)} - v)\big)\,dt\Big| \le 2a\big(|\partial_j\psi_z(v)| + 2Ha\big),$$

so that, using $(u + w)^2 \le 2u^2 + 2w^2$ and $\|\nabla\psi_z(v)\| \le G$,

$$\sum_{j=1}^{d}\big(\psi_z(v^{(j)}) - \psi_z(v)\big)^2 \le 8a^2\big(\|\nabla\psi_z(v)\|^2 + 4H^2a^2d\big) \le 16a^2G^2,$$

where the last step uses $2Ha\sqrt{d} = 2HR \le G$. Since $\Delta_v - \Delta_{v^{(j)}} = [\psi_y(v) - \psi_y(v^{(j)})] - [\psi_x(v) - \psi_x(v^{(j)})]$,

$$\sum_{j=1}^{d}\big(\Delta_v(x,y) - \Delta_{v^{(j)}}(x,y)\big)^2 \le 64a^2G^2, \qquad \big|\Delta_v(x,y) - \Delta_{v^{(j)}}(x,y)\big| \le 2G\|v - v^{(j)}\| = 4Ga. \tag{9}$$

Both bounds hold for every query pair, however far apart: they use only that f_v depends on v in a G-Lipschitz and H-smooth way, uniformly in the query point. The outcome of the query is $\mathrm{Bern}(\sigma(t))$ under v and $\mathrm{Bern}(\sigma(t'))$ under $v^{(j)}$, with $t = \Delta_v/\tau$ and $t' = \Delta_{v^{(j)}}/\tau$. Using $\mathrm{KL}(\mathrm{Bern}(p)\|\mathrm{Bern}(p')) \le \chi^2(\mathrm{Bern}(p)\|\mathrm{Bern}(p')) = (p-p')^2/(p'(1-p'))$, the mean value theorem, and the condition

$$4Ga \le \tau \tag{10}$$

(which gives $|t - t'| \le 1$), we obtain $\mathrm{KL}\big(\mathrm{Bern}(\sigma(t)) \,\|\, \mathrm{Bern}(\sigma(t'))\big) \le c_\sigma\,(t-t')^2 = c_\sigma(\Delta_v - \Delta_{v^{(j)}})^2/\tau^2$. Let P_v denote the joint law of the T outcomes and of the algorithm's internal randomness under instance v. Conditionally on the past, the query pair is a measurable function of the history, so the chain rule for the Kullback–Leibler divergence gives

$$\sum_{j=1}^{d} \mathrm{KL}(P_v\|P_{v^{(j)}}) \le T \sup_{x,y} \sum_{j=1}^{d} \frac{c_\sigma(\Delta_v(x,y) - \Delta_{v^{(j)}}(x,y))^2}{\tau^2} \le \frac{64 c_\sigma T G^2 a^2}{\tau^2}. \tag{11}$$

Step 3: Assouad. Since $|\hat v_j| \le a$ and $v_j \in \{\pm a\}$, $(\hat v_j - v_j)^2 \ge a^2 \mathbf{1}\{\mathrm{sign}(\hat v_j) \ne \mathrm{sign}(v_j)\}$ (with any fixed convention for $\hat v_j = 0$), so Assouad's lemma in its standard form gives

$$\begin{aligned}
\frac{1}{2^d}\sum_v \mathbb{E}_v\|\hat v - v\|^2 \;&\ge\; \frac{a^2}{2}\sum_{j=1}^{d}\frac{1}{2^d}\sum_v \big(1 - \mathrm{TV}(P_v, P_{v^{(j)}})\big) \\
&\ge\; \frac{da^2}{2}\left(1 - \sqrt{\frac{1}{2d}\,\frac{1}{2^d}\sum_v\sum_{j=1}^{d}\mathrm{KL}(P_v\|P_{v^{(j)}})}\right),
\end{aligned}$$

where the last step uses Pinsker's inequality $\mathrm{TV} \le \sqrt{\mathrm{KL}/2}$ and concavity of the square root. By (11) the square root is at most $s := \sqrt{32 c_\sigma T G^2 a^2/(d\tau^2)} = \sqrt{32 c_\sigma T G^2 R^2/(d^2\tau^2)}$. Combining with (8) and $da^2/2 = R^2/2$,

$$\frac{R^2}{2}(1-s) \;\le\; 4R\delta \qquad\Longrightarrow\qquad s \ge 1 - \frac{8\delta}{R} \qquad\Longrightarrow\qquad T \;\ge\; \Big(1 - \frac{8\delta}{R}\Big)^2 \frac{d^2\tau^2}{32\, c_\sigma\, G^2 R^2}.$$

It remains to choose R and G. The right-hand side is proportional to $(1 - 8\delta/R)^2/R^2$, which is maximized at $R = 16\delta$, and the constraint $2R \le G/H$ then forces $G \ge 32\varepsilon$; taking $G = 32\varepsilon$ (admissible since $\varepsilon \le L/32$) and $R = 16\delta$ gives $s \ge 1/2$, $a = 16\delta/\sqrt{d}$, $\phi(R) = 128\,\varepsilon\delta$, and

$$T \;\ge\; \frac{d^2\tau^2}{4 \cdot 32\, c_\sigma\, (32\varepsilon)^2 (16\delta)^2} \;=\; \frac{d^2\tau^2}{2^{25}\, c_\sigma\, \varepsilon^2\delta^2}.$$

With these values (10) reads $2048\,\varepsilon\delta \le \tau\sqrt{d}$, that is $d \ge 2^{22}\kappa_\varepsilon^2$, and the initial gap is $128\,\varepsilon\delta$, which are the hypotheses of the theorem. Enlarging the cube weakens the bound as $1/R^2$, so no choice of the free parameters of this family yields a bound that grows with Δ_f. □

Why value oracles escape the bound. A comparison carries bounded Fisher information about v because it is a Bernoulli variable with a smooth parameter; a same-seed two-point difference $F(x+\rho u;\xi) - F(x-\rho u;\xi)$ has no such bound, since its noise contracts with ρ. The content of Theorem 5.1 is $\mathbb{E}\|\hat v - v\|^2 \ge d^2/\operatorname{tr}\mathcal{I}_T$ with $\operatorname{tr}\mathcal{I}_T \le 4c_\sigma T G^2/\tau^2$: one factor d for the coordinates to learn, one for receiving one number per query.

Why the slope, and not the Lipschitz constant, controls the bound. Steps 1 and 3 are unchanged if the slope G is replaced by any $G' \in [2\varepsilon, L]$: the localization accuracy 2δ forced by (δ,ε)-Goldstein stationarity comes from the curvature $H = \varepsilon/\delta$, not from the slope. Step 2, on the other hand, is proportional to G'^2, because a comparison between two points reveals the hidden parameter only through differences of function values, whose sensitivity to v is the slope. Taking G' as small as the argument allows, namely $\Theta(\varepsilon)$, is therefore optimal, and yields a bound larger by $(L/\varepsilon)^2$ than the same argument run with $G' = L$.

Finiteness of c_σ for the running links. For the logistic link $\sigma' = \sigma(1-\sigma)$ and $(\log\sigma')' = 1-2\sigma \in (-1,1)$, so $\sigma'(s) \le e\,\sigma'(t)$ for $|s-t| \le 1$ and $\sigma'(s)^2/(\sigma(t)(1-\sigma(t))) = \sigma'(s)^2/\sigma'(t) \le e\,\sigma'(s) \le e/4$; hence $c_\sigma \le e/4$ (numerically $c_\sigma \approx 0.38$). For the probit link, $\sigma' = \varphi$ and, for $t \ge 1$, $1-\Phi(t) \ge \varphi(t)\,t/(1+t^2)$ (Mills' ratio) while $\varphi(s)^2 \le \varphi(t)^2 e^{2t+1}$ for $|s-t| \le 1$, so the ratio is at most $\varphi(t)e^{2t+1}(1+t^2)/t \to 0$ as $t \to \infty$; by symmetry and continuity $c_\sigma < \infty$. For the cauchit link, $\sigma'(t) = 1/(\pi(1+t^2))$ and $1-\sigma(t) \sim 1/(\pi t)$, so the ratio is $O(t^{-3})$ as $|t| \to \infty$ and again $c_\sigma < \infty$.

Random budgets. Theorem 5.1 assumes at most T comparisons almost surely. Suppose instead that the number of comparisons N is random with $\sup_v \mathbb{E}_v[N] \le T$, as in Theorems 4.2 and 4.3. Run the algorithm but stop it after $32T$ comparisons, outputting $\hat{x}$ if it has finished and 0 otherwise. By Markov's inequality the stopped output differs from $\hat{x}$ with probability at most $1/32$ under every v, and since $\|\hat{v}-v\| \le 2R$ always, the bound (8) becomes $\mathbb{E}\|\hat{v}-v\|^2 \le 4R\delta + 4R^2/32$, which at $R = 16\delta$ gives $96\delta^2$ in place of $64\delta^2$; the Assouad step then yields $s \ge 1/4$ and the conclusion $T \ge d^2\tau^2/(2^{32}c_\sigma\varepsilon^2\delta^2)$. The comparison of Theorems 4.2 and 5.1 in Section 5 is therefore legitimate up to this constant.

Tightness on the hard instances. The instances used above have $\Delta_f = 128\varepsilon\delta$ and are (ε/δ)-smooth, so the initial gap of the smoothed function obeys $f_\rho(x_0) - \inf f_\rho \le \Delta_f + H\rho^2 \le \Delta_f + \varepsilon\delta/4$ for $\rho \le \delta/2$. Theorem 4.2 then uses $T = O(\sigma_G^2\Delta_h/(\delta\varepsilon^3))$ iterations with $\sigma_G^2 = O(C_{\mathrm{link}}d^2\tau^2/\delta^2)$ at $\rho = \delta/2$, hence $O\big(\mathbb{E}[N]C_{\mathrm{link}}\,d^2\tau^2/(\delta^2\varepsilon^2)\big)$ comparisons, against the $\Omega\big(d^2\tau^2/(c_\sigma\,\delta^2\varepsilon^2)\big)$ of Theorem 5.1. For $\kappa = \Theta(1)$, where C_{link} and $\mathbb{E}[N]$ are absolute constants, the two match up to a factor depending only on the link, provided $d \ge 2^{22}\kappa_\varepsilon^2$ (about $4\cdot 10^3$ when $\varepsilon = L/32$ and $\kappa_\delta = 1$). On this family, for a fixed link and bounded κ, the algorithm therefore matches the lower bound in d, τ, δ and ε, and the remaining gap between Theorems 4.2 and 5.1 is the dependence on Δ_h.

A second, complementary term. A comparison is a function of the two queried values and of independent randomness, so an algorithm with an exact first-order oracle simulates a comparison query with two of its own queries. Any lower bound for randomized first-order algorithms therefore transfers, up to a factor 2, to the comparison model. Applying this to the class of H-smooth functions with $H = \varepsilon/\delta$, and using that a (δ,ε)-Goldstein stationary point of an H-smooth function is a 2ε-stationary point, the bound $\Omega(\Delta_f H/\varepsilon^2)$ of Carmon et al. (2020) for finding ε-stationary points of H-smooth functions gives

$$T = \Omega\Big(\frac{\Delta_f}{\delta\varepsilon}\Big), \tag{12}$$

valid for randomized algorithms once d is polynomially large in $\Delta_f/(\delta\varepsilon)$, as required by the random-rotation argument of Carmon et al. (2020). Combining with Theorem 5.1,

$$T = \Omega\Big(\max\Big\{\frac{d^2\tau^2}{c_\sigma\,\varepsilon^2\delta^2},\ \frac{\Delta_f}{\delta\varepsilon}\Big\}\Big).$$

The second term is the only Δ_f-dependence we can prove; it is dominated by the first unless $\Delta_f \gtrsim d^2\tau^2/(c_\sigma\,\varepsilon\delta)$, and it carries no dependence on τ or d, since it holds even for an exact value oracle.

On a joint bound in $(d,\delta,\varepsilon,\Delta_f)$. The bound above uses a single-stage instance and is therefore independent of Δ_f. The natural attempt at a joint bound replaces it by a zero-chain $F_U(x) = F(U^\top x)$ with $U \in \mathbb{R}^{d\times K}$ having orthonormal columns, scaled to be H-smooth with gradients of order ε before completion, which forces $K = \Theta(\Delta_f/(\delta\varepsilon))$ stages and a per-stage decrease $\Theta(\varepsilon\delta)$, and charges the algorithm for localizing each hidden direction u_{i+1}. This does not reproduce the factor K, for a reason worth recording. Completing the chain requires $|\langle x, u_j\rangle| \gtrsim \delta$ for all $j \le K$ simultaneously, hence query points of norm $R = \Theta(\delta\sqrt{K})$; the sensitivity of the observed gap to u_{i+1} is then of order εR rather than $\varepsilon\delta$, so a single comparison carries $O(c_\sigma\varepsilon^2R^2/(d\tau^2)) = O(c_\sigma\kappa_\varepsilon^2K/d)$ nats about u_{i+1}, with $\kappa_\varepsilon := \varepsilon\delta/\tau$. At the same time the alignment needed to activate a stage is only $\langle x/\|x\|, u_{i+1}\rangle \ge \Omega(1/\sqrt{K})$, which costs $O(d/K)$ nats. The two effects give $T_i = \Omega(d^2/(\kappa_\varepsilon^2K^2))$ per stage and $T = \Omega(d^2/(\kappa_\varepsilon^2K))$ in total, which is decreasing in K and is maximized at $K = 1$, that is, at the single-stage instance used above. A joint bound therefore requires a hard family in which the sensitivity to the hidden parameter does not grow with the norm of the query, as it does not for the Huber instances, while retaining strict sequentiality across $\Theta(\Delta_f/(\delta\varepsilon))$ stages. We leave this open.

A.7 Relation to existing lower bounds

For the logistic link there is a reduction to independently noised function values: if Γ, Γ' are i.i.d. Gumbel with scale τ, then $\mathbf{1}\{f(y)+\Gamma' > f(x)+\Gamma\}$ is Bernoulli with parameter $\sigma((f(y)-f(x))/\tau)$ (Bradley & Terry, 1952; Yellott, 1977), so a comparison algorithm is an algorithm with independent per-evaluation noise, and lower bounds for that model transfer whenever their proof accommodates Gumbel noise. Shamir (2013) shows that with independent noise the number of evaluations must scale at least quadratically with d even for strongly convex smooth functions, whereas two evaluations under the same noise realization cost only a factor $\sqrt{d}$ (Duchi et al., 2015; Shamir, 2017); Theorem 5.1 is the comparison-feedback analogue, with the link's Fisher information and κ explicit and with Goldstein stationarity as the target. The comparison-oracle lower bound of Jamieson et al. (2012) concerns strongly convex functions with reliability $\mathbb{P}(\text{correct}) \geq \frac{1}{2} + \min\{\delta_0, \mu|\Delta|^{\kappa_J - 1}\}$; a smooth link has $\kappa_J = 2$ and $\mu \asymp 1/\tau$, and their bound $\mathbb{E}[f(\hat{x}) - \inf f] \gtrsim \sqrt{d/T}$ converts, for μ_{sc}-strongly convex f, through $\|\nabla f\|^2 \geq 2\mu_{\mathrm{sc}}(f - \inf f)$ into a bound of order $d\,\mu_{\mathrm{sc}}^2/\varepsilon^4$ comparisons for the mean-squared criterion $\mathbb{E}\|\nabla f(\hat{x})\|^2 \leq \varepsilon^2$, up to link constants (the expected-norm criterion of this paper does not follow without further moment control, so we use the comparison only qualitatively). That bound is linear in d, which is the dependence produced by needle-in-a-haystack arguments, where each query is informative about one of d candidate locations, so the dimension counts places to search and the accuracy enters through the divergence. In the family of Theorem 5.1, by contrast, the sensitivity of a query to the hidden parameter is spread across all coordinates while its total is bounded, which is where the second factor of d comes from. The upper bound's term $d^2 L^2 \Delta_f/(\kappa^2 \delta \varepsilon^3)$ combines both effects, and a matching joint bound would need one hard family exhibiting both; the natural candidate is the chain construction of Arjevani et al. (2023) in the globally Lipschitz form of Cutkosky et al. (2023, Appendix F), with the variance-throttling step replaced by the Bernoulli divergence bound above and the chain embedded in a hypercube of shifts. Even for value oracles no joint zero-order lower bound in (d, δ, ε) is known (Kornowski & Shamir, 2023); we leave the comparison version as a conjecture.

B Experimental details and additional results

All experiments run on a single CPU core with NumPy; the synthetic experiments take under thirty minutes in total and the two real tasks about forty minutes. Random seeds are fixed in the code. Throughout we use the logistic link and the run-length estimator of Definition 3.5. Where a Lipschitz constant is needed it is estimated as the 95th percentile of finite-difference gradient norms over 100 random points around the initial point (the maximum is dominated by a few near-kink points and is reported alongside).

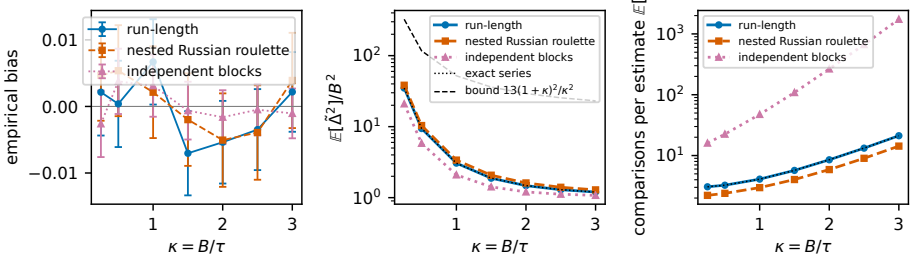

Figure 3: Gap estimators for the logistic link at the edge of the operated range ($\Delta = B$, $\tau = 1$, $2 \cdot 10^5$ samples per point): the run-length estimator of Definition 3.5 and two Russian-roulette variants with geometric truncation $\beta = \sqrt{\alpha}$ (nested single stream, and independent blocks of sizes $1, \ldots, M$). Left: empirical bias with two standard errors. Middle: second moment relative to B^2, with the exact series (5) and the bound $13(1+\kappa)^2/\kappa^2$. Right: comparisons per estimate; the dotted line is $1 + \alpha/(1-\alpha) + (1-\alpha)/\alpha$.

Estimator (Figure 3, Table 2). For each $\kappa \in \{0.25, 0.5, 1, 1.5, 2, 2.5, 3\}$ we set $\tau = 1$, $B = \kappa$ and $\Delta = B$ (the edge of the operated range, where the second moment and the cost are largest) and draw

$2 \cdot 10^5$ independent estimates with the run-length estimator and, for comparison, with two Russian-roulette estimators with geometric truncation $\beta = \sqrt{\alpha}$: the nested version that reuses one stream for all m (cost $\min\{M, N\}$, since the terms vanish after the first outcome change) and the version with independent blocks of sizes $1, \dots, M$ (cost $M(M+1)/2$). All three are unbiased in simulation (sample biases statistically consistent with zero: all within two standard errors, one point at 2.2). The run-length estimator has a smaller second moment than the nested Russian-roulette estimator at every tested κ; the nested Russian-roulette estimator, stopped at the first outcome change as (7) allows, uses about 70% as many comparisons for a second moment about 10% larger, so its product of cost and second moment is smaller by 20 to 30%. The independent-block version is far more expensive than both. The simulated second moments of $\widetilde{\Delta}$ agree with the exact series (5) to three digits, and the bound $13\tau^2(1+\kappa)^2$ is loose by a factor 15 to 20. Both products are minimized near $\kappa = 1.5$ on this grid.

Table 2: Estimator statistics at $\Delta = B$ (logistic link, $\tau = 1$). RR: Russian roulette with $\beta = \sqrt{\alpha}$.

		run-length			nested RR				independent RR	
κ	α	$\mathbb{E}[\widetilde{\Delta}^2]/B^2$	exact (5)	$\mathbb{E}[N]$	$\mathbb{E}[N]\mathbb{E}[\widetilde{\Delta}^2]/B^2$	$\mathbb{E}[\widehat{\Delta}^2]/B^2$	$\mathbb{E}[N]$	product	$\mathbb{E}[\widehat{\Delta}^2]/B^2$	$\mathbb{E}[N]$
0.25	0.562	34.95	34.93	3.06	107.1	38.29	2.21	84.8	20.94	16.0
0.50	0.622	9.41	9.43	3.25	30.6	10.36	2.39	24.8	5.83	22.5
1.00	0.731	3.06	3.06	4.09	12.5	3.38	2.97	10.0	2.09	47.4
1.50	0.818	1.88	1.88	5.69	10.7	2.07	4.03	8.3	1.42	108.6
2.00	0.881	1.48	1.48	8.51	12.6	1.61	5.88	9.5	1.20	265.5
2.50	0.924	1.29	1.30	13.23	17.1	1.40	9.01	12.6	1.12	667.4
3.00	0.953	1.20	1.20	21.17	25.4	1.29	14.29	18.4	1.08	1739.0

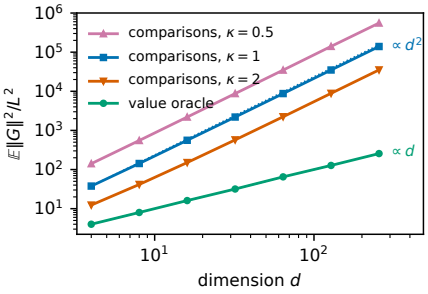

Figure 4: Second moment of the gradient estimator G for $f(x) = \|x\|_\infty$ ($\rho = 0.1$, $4 \cdot 10^4$ directions per point). Values give $\mathbb{E}\|G\|^2 \propto d$; comparisons give $\propto d^2$, with the constant decreasing in κ.

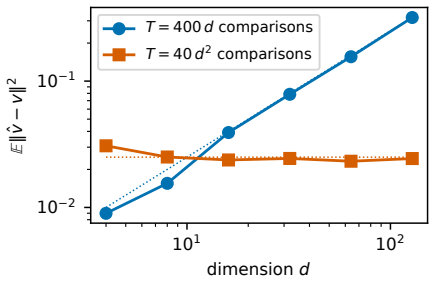

Figure 5: Mean squared error of the logistic maximum-likelihood estimate of v from T comparisons of $\pm u$ on the quadratic instance ($H = \tau = 1$, 30 repetitions). Dotted lines are the Fisher-information prediction $d^2\tau^2/(TH^2)$.

Second moment and estimation rate (Figures 4 and 5). For $f(x) = \|x\|_\infty$ the exact two-point estimator has $\mathbb{E}\|G\|^2 \approx dL^2$ while the comparison-based estimator grows as d^2, which is the mechanism of Lemma 4.1: comparison noise does not benefit from the concentration of Δ_u. On the quadratic instance of Theorem 5.1 the maximum-likelihood estimate of v from T comparisons has mean squared error matching $d^2\tau^2/(TH^2)$ within 10% for d from 4 to 128: doubling d requires quadrupling T.

Figure 4. $f(x) = \|x\|_\infty$ ($L = 1$), $\rho = 0.1$, at a Gaussian random point whose first coordinate is shifted by 3 so that the maximum is attained at a single coordinate. For each $d \in \{4, \dots, 256\}$ we draw $4 \cdot 10^4$ directions u, compute the exact gaps Δ_u, and form $\frac{d^2}{4\rho^2}\Delta_u^2$ (value oracle) and $\frac{d^2}{4\rho^2}\widehat{\Delta}^2$ with one comparison-based estimate per direction at $\kappa \in \{0.5, 1, 2\}$, $\tau = B/\kappa$.

Figure 5. The instance is $f_v(x) = \frac{H}{2}\|x - v\|^2$ (the quadratic region of the Huber function) with $H = 1$, $v \in \{\pm 0.05\}^d$, $\tau = 1$, and queries $(-u, u)$ with $u \sim \mathrm{Unif}(\mathbb{S}^{d-1})$, for which $\Delta_v =$

$-2H\langle u, v\rangle$. We fit v by logistic regression (Newton's method, ridge 10^{-8}) on T outcomes and report the mean squared error over 30 repetitions, for $T = 400d$ and $T = 40d^2$. The prediction $d^2\tau^2/(TH^2)$ is the Cramér–Rao value with $\sigma'(0) = 1/4$ and $\mathbb{E}[uu^\top] = I/d$.

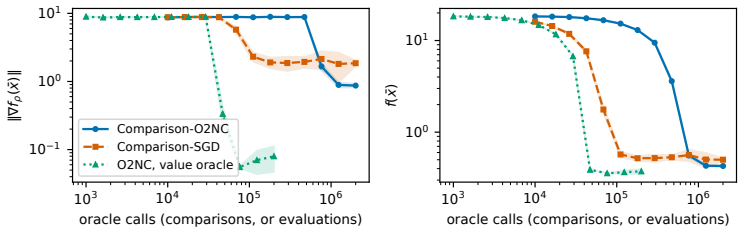

Figure 6: Nonsmooth Chebyshev–Rosenbrock function in $d = 10$ ($\delta = 0.2$, $\rho = \nu = 0.1$, $\kappa = 1$). Comparison-O2NC and Comparison-SGD with a budget of $2 \cdot 10^6$ comparisons and O2NC with an exact two-point value oracle with $2 \cdot 10^5$ evaluations; step sizes tuned per method. Mean and standard deviation over 5 seeds; $\|\nabla f_\rho(\bar{x})\|$ is estimated by Monte Carlo at block averages.

Synthetic optimization (Figure 6). $f(x) = \frac{1}{4}|x_1 - 1| + \sum_{i=1}^{d-1} |x_{i+1} - 2|x_i| + 1|$ with $d = 10$, $x_0 = -\mathbf{1}$ ($f(x_0) = 18.5$), $\delta = 0.2$, $\rho = \nu = 0.1$, $L \approx 8.8$, $\kappa = 1$. Comparison-based methods use $2 \cdot 10^6$ comparisons, the value-oracle reference $2 \cdot 10^5$ evaluations. Step sizes were selected on one seed at half the budget by the final estimated $\|\nabla f_\rho\|$ from the grids $\eta \in \{10^{-6}, 3 \cdot 10^{-6}, 10^{-5}, 3 \cdot 10^{-5}\} \times D \in \{10^{-4}, 3 \cdot 10^{-4}, 10^{-3}\}$ (Comparison-O2NC), $\eta \in \{10^{-5}, 3 \cdot 10^{-5}, 10^{-4}, 3 \cdot 10^{-4}, 10^{-3}\}$ (Comparison-SGD), and $\eta \in \{10^{-5}, 3 \cdot 10^{-5}, 10^{-4}\} \times D \in \{10^{-4}, 3 \cdot 10^{-4}, 10^{-3}\}$ (value-oracle O2NC); selected: $(\eta, D) = (10^{-6}, 10^{-4})$, $\eta = 10^{-5}$, and $(10^{-4}, 3 \cdot 10^{-4})$; the two comparison-based selections lie at the lower edge of their grids. Comparison-O2NC drives f from 18.5 to 0.43 ± 0.01 and $\|\nabla f_\rho(\bar{x})\|$ to 0.87 ± 0.06 in $6.6 \cdot 10^5$ iterations ($\mathbb{E}[N] \approx 3$ comparisons per iteration, since the queried gaps are well inside the operated range), Comparison-SGD to 0.50 ± 0.02 and 1.85 ± 0.18; the value-oracle reference reaches 0.37 ± 0.02 and 0.08 ± 0.03 with ten times fewer oracle calls. The gradient norm $\|\nabla f_\rho(\bar{x})\|$ is estimated with 3000 exact two-point samples (subtracting the sampling variance from the squared norm) at the averages $\bar{x}$ of the last eight blocks of $\lfloor \nu/D \rfloor$ consecutive points z_t for O2NC, and at the last eight iterates for SGD.

Real tasks: common protocol. For each task and dimension, the temperature is $\tau_0 = L\delta/\kappa^\star$ with $\kappa^\star = 1.5$ and $\delta = 0.05$, so that $\rho = \delta/2$ gives $\kappa = \kappa^\star$; $\nu = \delta - \rho$. Step sizes are chosen by grid search on one seed at half the final iteration budget, selecting by the final objective value: for Comparison-O2NC, $\eta \in \eta_{\text{th}} \cdot \{10, 30, \ldots, 3 \cdot 10^3\}$ and $D \in D_{\text{th}} \cdot \{3, 10, 30\}$, where $\eta_{\text{th}} = \Delta_h/(\sigma_G^2 T)$ and $D_{\text{th}} = (\nu^{1/2}\Delta_h/(\sigma_G T))^{2/3}$ are the values of Theorem 4.2 with σ_G^2 computed from Lemma 4.1 using the empirical $\mathbb{E}[\widetilde{\Delta}^2]/B^2 = 1.88$ at $\kappa = 1.5$; for Comparison-SGD $\eta \in \eta_{\text{th}} \cdot \{3 \cdot 10^2, \ldots, 3 \cdot 10^5\}$; for the value oracle $\eta \in \eta_{\text{th}} \cdot \{10^2, \ldots, 10^4\}$ rescaled by the ratio of the two second-moment bounds. Final runs use 3 seeds. All methods use the same number of iterations, so the comparison-based methods use $\mathbb{E}[N]$ comparisons per iteration against 2 evaluations; since the queried gaps are small relative to τ_0, $\mathbb{E}[N] \approx 3$ in practice rather than the nominal 5.7. Selected values: pendulum $d = 16$: Comparison-O2NC $(\eta, D) = (1.1 \cdot 10^{-7}, 4.7 \cdot 10^{-4})$, Comparison-SGD $\eta = 3.3 \cdot 10^{-3}$, value-oracle O2NC $(\eta, D) = (1.9 \cdot 10^{-6}, 1.4 \cdot 10^{-3})$; pendulum $d = 66$: Comparison-O2NC $(\eta, D) = (1.1 \cdot 10^{-8}, 1.4 \cdot 10^{-4})$, Comparison-SGD $\eta = 3.6 \cdot 10^{-5}$, value-oracle O2NC $(\eta, D) = (4.4 \cdot 10^{-7}, 4.3 \cdot 10^{-4})$; ℓ_1 regression $d = 49$: Comparison-O2NC $(\eta, D) = (6.9 \cdot 10^{-8}, 2.0 \cdot 10^{-4})$, Comparison-SGD $\eta = 2.1 \cdot 10^{-4}$, value-oracle O2NC $(\eta, D) = (2.3 \cdot 10^{-5}, 2.0 \cdot 10^{-4})$; $d = 97$: Comparison-O2NC $(\eta, D) = (8.1 \cdot 10^{-10}, 3.1 \cdot 10^{-5})$, Comparison-SGD $\eta = 2.4 \cdot 10^{-5}$, value-oracle O2NC $(\eta, D) = (1.4 \cdot 10^{-6}, 9.4 \cdot 10^{-5})$; $d = 193$: Comparison-O2NC $(\eta, D) = (1.8 \cdot 10^{-9}, 6.3 \cdot 10^{-5})$, Comparison-SGD $\eta = 5.5 \cdot 10^{-6}$, value-oracle O2NC $(\eta, D) = (1.8 \cdot 10^{-7}, 6.3 \cdot 10^{-5})$. The extended pendulum runs and the temperature sweep start from a different random initialization than the corresponding grid search, with the step sizes selected there.

Pendulum. Dynamics are those of the standard pendulum benchmark ($g = 10$, $m = l = 1$, $dt = 0.05$, angular velocity clipped to $[-8, 8]$, torque clipped to $[-2, 2]$), with 8 fixed initial states

drawn once (angle uniform in $[-\pi, \pi]$, velocity uniform in $[-1, 1]$) and a horizon of 100 steps. The policy maps $(\cos\theta, \sin\theta, \dot{\theta}/8)$ through a ReLU layer of width w to a scalar torque, $d = 5w + 1$. Widths 3 and 13 give $d = 16$ and $d = 66$; the initial policy has cost 7.4 for $d = 16$ and 6.4 for $d = 66$ in the grid search, and 6.9 for $d = 16$ in the extended runs, which start from a different initialization. For $d = 16$ the final runs were extended, with the step sizes selected at $2 \cdot 10^4$ iterations, to $5 \cdot 10^4$ iterations for the value oracle and $1.2 \cdot 10^5$ iterations for the two comparison-based methods.

Temperature sweep and misspecified links (ℓ_1 regression, $d = 49$). For $\tau \in \tau_0 \cdot \{1/256, 1/64, 1/16, 1/4, 1, 4\}$, Comparison-O2NC is run with the step sizes selected at τ_0 until a budget of $2 \cdot 10^5$ comparisons is exhausted (at most 10^5 iterations), either with $\rho = \delta/2$ (so the nominal $\kappa = L\delta/\tau$ ranges over $384, 96, 24, 6, 1.5, 0.375$) or with $\rho = \min\{\delta/2, \kappa^\star\tau/(2L)\}$, which keeps $\kappa = 1.5$ for $\tau \le \tau_0$. The number of comparisons per estimate depends on the actual success probabilities: since typical gaps are well below B, it is about 3 for all capped runs and for $\rho = \delta/2$ down to $\tau_0/4$, about 4 at $\tau_0/16$, about $3 \cdot 10^4$ at $\tau_0/64$ and 10^5 at $\tau_0/256$, where the budget covers seven and two iterations. For the misspecified links the outcomes are generated by $\Phi(\Delta/(1.702\tau_0))$ (a probit with the slope of the logistic at zero), by the logistic link with temperature $2\tau_0$, and by the logistic link with temperature $\tau_0/2$, while the estimator uses the logistic link with τ_0; the estimator is then biased, by a factor close to $\tau_0/\tau_{\text{true}}$ near zero gap in the last two cases.

ℓ_1 regression. Inputs and target of the diabetes data are standardized; the model is a ReLU layer of width w followed by a linear output, $d = 12w + 1$; the loss is the mean absolute error over the 442 samples, whose initial value is 1.05 for the width-4 initialization used in Section 6 (0.92 and 1.22 for widths 8 and 16; a least-squares linear fit has mean absolute error 0.56). Widths $4, 8, 16$ ($d = 49, 97, 193$) are run for $6 \cdot 10^4$ iterations with grid search at $3 \cdot 10^4$; the middle panel of Figure 2 reports, per method, the median over seeds of the number of oracle calls at the first checkpoint whose loss is at most 0.65. The second-moment panel uses widths 2 to 64 ($d = 25$ to 769) at the initial point with 4000 directions, $\rho = \delta/2$, and a temperature fixed across d at $\tau_0 = L\delta/\kappa^\star$ with L the estimate for width 4 (and $\tau_0/2$, $2\tau_0$). With the temperature fixed, the comparison term of Lemma 4.1 is exactly $\propto d^2$; the value term grows faster than d here because the gradient norm at a random initialization grows with the width. The largest observed gap over 4000 directions is far below $B = 2L\rho$ (between 0.14 and 0.37 of τ_0), a consequence of the concentration of Δ_u: the nominal κ is a worst case, and typical comparisons are made at gaps well inside the operated range.

C BASELINES

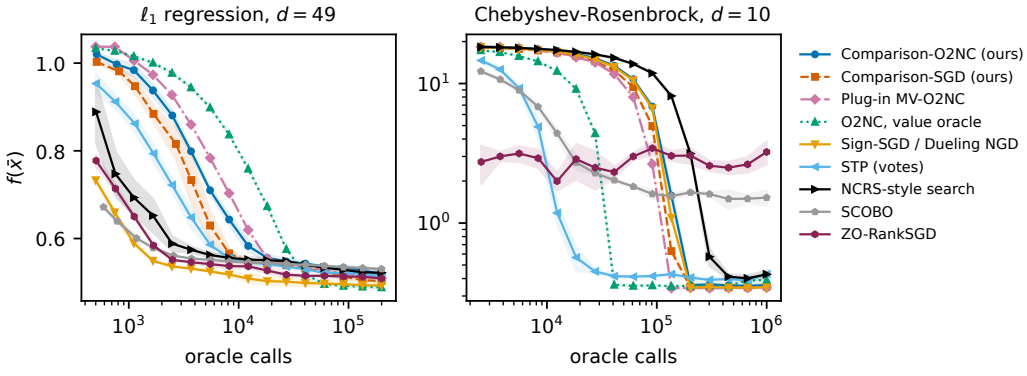

Figure 7: Budget-matched comparison at $\tau = \tau_0/32$ on ℓ_1 regression ($d = 49$, $2 \cdot 10^5$ comparisons) and the nonsmooth Chebyshev–Rosenbrock function ($d = 10$, 10^6 comparisons). Mean and standard deviation over 3 seeds; hyperparameters selected per method by grid search at half budget.

All methods use the same logistic comparison oracle and are charged one unit per comparison (the value oracle, two evaluations per step, is charged two). The objectives are deterministic, so a pair may be compared repeatedly; that is how the vote-based methods spend their budget.

- **Sign-SGD / dueling.** $x \leftarrow x - \eta\, s\, u$, where $s \in \{\pm 1\}$ is the majority outcome of k comparisons of $x \pm \rho u$ (the direction in which f is reported to increase); $k = 1$ is the normalized-gradient update of Saha et al. (2021) and the stochastic random search of Chayti et al. (2025), whose analysis of the sign of a noisy two-point difference also explains why per-comparison signal and noise scale alike, which is what we observe in Section 6. Grid: $\eta \in \{3 \cdot 10^{-4}, \dots, 3 \cdot 10^{-1}\}$, ρ over the radii below and $\{\delta/2, 2\delta\}$, $k \in \{1, 5, 21\}$.
- **STP** (Bergou et al., 2020): move to the better of $x \pm \alpha u$ when a majority of k comparisons of that candidate against x reports it smaller; each candidate is compared with x separately, so an iteration costs $2k$ comparisons. **NCRS-style search** (El Bakkali et al., 2026): one candidate $x + \alpha u$, accepted on a majority of k comparisons. Grids: α as η above, $k \in \{1, 5, 21\}$.
- **SCOBO** (Cai et al., 2022): m comparisons of x against $x + r z_i$, z_i Gaussian, decoded by the Plan–Vershynin program $\max\langle v, g\rangle$ over $\|g\| \le 1$, $\|g\|_1 \le \sqrt{s}$, then a normalized step with $\eta/\sqrt{t}$. Grid: $m \in \{d/2, 2d\}$, $s \in \{d/4, d\}$, $k \in \{1, 5\}$, $r = \delta$. We do not use its comparison-based line search.
- **ZO-RankSGD** (Tang et al., 2024): m perturbed points ranked by their win counts over all $m(m-1)/2$ pairs (k votes each), then the rank-weighted direction. Grid: $m \in \{4, 8\}$, $k \in \{1, 5\}$, $\mu = \delta$.
- **Plug-in** (biased): $\tau \log(\hat{p}/(1-\hat{p}))$ with $\hat{p}$ the fraction of k comparisons clipped to $[1/2k, 1 - 1/2k]$, inside O2NC, $k \in \{5, 21\}$.
- **Ours and the value oracle:** Algorithm 1 and its SGD variant, with the smoothing radius from the capped rule $\rho = \min\{\delta/2, \kappa^\star \tau/(2L)\}$ for $\kappa^\star \in \{1.5, 4, 8, 16, 32, 64, 128\}$, η over ten decades and $D \in \delta \cdot \{0.002, 0.01, 0.05\}$.

Selection is by the final objective on one seed at half budget; final runs use 3 different seeds at full budget, so the selection seed is not among the reported runs. Three seeds and grids whose optimum sometimes lies on an edge (noted where it occurs) limit the resolution of the ranking; differences below about 0.02 in Table 3 should not be read as significant. Table 3 reports the final objective and the smoothed gradient norm $\|\nabla f_\rho(\bar{x})\|$, estimated with 2000 exact two-point samples at $\rho = \delta/2$ for every method, so that the stationarity measure is comparable across methods that use different internal radii.

Table 3: Final objective and smoothed gradient norm at $\tau = \tau_0/32$ (mean $\pm$ std over 3 seeds).

	ℓ_1 regression, $d = 49$		Chebyshev–Rosenbrock, $d = 10$	
method	$f(\bar{x})$	$\|\nabla f_\rho(\bar{x})\|$	$f(\bar{x})$	$\|\nabla f_\rho(\bar{x})\|$
Comparison-O2NC (ours)	0.518 ± 0.010	0.078 ± 0.022	0.359 ± 0.006	0.387 ± 0.035
Comparison-SGD (ours)	0.503 ± 0.003	0.105 ± 0.024	0.348 ± 0.001	0.297 ± 0.020
Plug-in majority vote	0.509 ± 0.008	0.087 ± 0.036	0.344 ± 0.004	0.230 ± 0.023
Sign-SGD / dueling	0.493 ± 0.006	0.062 ± 0.007	0.346 ± 0.004	0.254 ± 0.035
STP (votes)	0.513 ± 0.004	0.144 ± 0.064	0.431 ± 0.009	1.132 ± 0.157
NCRS-style search	0.522 ± 0.010	0.090 ± 0.014	0.429 ± 0.028	1.246 ± 0.357
SCOBO	0.531 ± 0.002	0.073 ± 0.020	1.521 ± 0.194	7.029 ± 0.380
ZO-RankSGD	0.509 ± 0.000	0.117 ± 0.015	3.231 ± 0.608	6.666 ± 0.284
O2NC, value oracle	0.489 ± 0.003	0.018 ± 0.003	0.390 ± 0.023	0.110 ± 0.038

At τ_0 and $\tau_0/8$ the ordering on ℓ_1 regression is the same, with all comparison-based methods within 0.51 to 0.56; at τ_0, where the queried gaps are far below the temperature, the sign of one comparison is already an unbiased estimate of the gap up to a scalar and the sign-based methods are ahead of ours throughout. Our selected radii sit at the largest value the grid allows on ℓ_1 regression ($\rho = \delta/2$ for O2NC at $\tau_0/8$, $\delta/3$ at $\tau_0/32$), so on that problem the binding constraint is $\rho \le \delta/2$ rather than the cost of the estimator; sign-SGD selects $\rho = 2\delta$, outside the range the Goldstein target allows for our smoothing argument.