THE MULTI-QUERY PARADOX IN ZEROTH-ORDER OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Zeroth-order (ZO) optimization provides a powerful framework for problems where explicit gradients are unavailable and have to be approximated using only queries to function values. The prevalent single-query approach is simple, but suffers from high estimation variance, motivating a multi-query paradigm to improve estimation accuracy. This, however, creates a critical trade-off: under a fixed budget of queries (i.e. cost), queries per iteration and the total number of optimization iterations are inversely proportional to one another. How to best allocate this budget is a fundamental, under-explored question.

This work systematically studies this query allocation problem. We analyze two aggregation methods: the de facto simple averaging (ZO-Avg), and a projection-based rule (ZO-Align), which we derive from local surrogate minimization. By deriving convergence rates for both methods that make the dependence on the number of queries explicit across strongly convex, convex, and non-convex settings, we uncover aggregation-dependent allocation behaviors: For ZO-Avg, we prove that using more than one query per iteration is always query-inefficient, rendering the single-query approach optimal. On the contrary, ZO-Align generally performs better with more queries per iteration, resulting in a full-subspace estimation as the optimal approach. Thus, our work clarifies that the multi-query problem boils down to a choice not about an intermediate query size, but between two classic algorithms, a choice dictated entirely by the aggregation method used. These theoretical findings are also consistently validated by extensive experiments.

1 INTRODUCTION

Zeroth-order (ZO) optimization provides a powerful framework for problems where explicit gradient information is unavailable or prohibitively expensive to compute (Conn et al., 2009). Unlike first-order methods that rely on analytical gradients, ZO algorithms guide their search using only function value queries to approximate gradients. Typically, it is achieved by evaluating the objective function along one or more random directions and observing the change in function value (Spall, 2002; Mania et al., 2018). This paradigm has recently gained significant attention due to its applicability to a wide array of challenging, high-dimensional problems, including fine-tuning of LLMs (Malladi et al., 2023; Zhang et al., 2024), policy optimization in reinforcement learning (Sun et al., 2020), and hyperparameter tuning (Li et al., 2021).

The predominant approach in modern ZO optimization is the single-query method, where a gradient estimate is constructed using one random direction per optimization step (Nesterov & Spokoiny, 2017). While foundational, these estimators are inherently noisy, with an estimation variance that often scales with the problem's dimension (Nesterov & Spokoiny, 2017). A natural extension to mitigate this issue is the multi-query approach, which employs a block of queries at each iteration to construct a more accurate gradient estimate (Duchi et al., 2015).

However, this extension introduces a fundamental trade-off. Using multiple queries per iteration proportionally reduces the total number of iterations that can be performed under a fixed query budget. This tension gives rise to what we call the *multi-query paradox*: increasing the number of queries per iteration improves the quality of each gradient estimate but simultaneously decreases the total number of updates one can afford. A systematic investigation of this paradox requires addressing two basic questions: (1) How does the rule used to aggregate multiple directional queries affect their query

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.

efficiency? (2) Given a fixed total query budget, how should queries be allocated across iterations under different aggregation rules?

This work systematically addresses these questions. First, we study two representative ways of aggregating the same set of directional queries: the canonical simple averaging rule (ZO-Avg) (Duchi et al., 2015) and a projection-based rule, which we refer to as Projection Alignment (ZO-Align). ZO-Align can be derived by minimizing a local quadratic surrogate over the span of the queried directions, yielding a geometry-aware aggregation that accounts for correlations among them. Closely related projection-based estimators have appeared in prior zeroth-order optimization literature, particularly from the perspective of sparse gradient recovery (Cai et al., 2022; Wang et al., 2018). Our local-surrogate derivation provides an optimization-centric interpretation that facilitates a direct comparison with simple averaging. Our estimator-level analysis reveals a substantial difference in their estimation errors, suggesting that the aggregation rule can fundamentally change how effectively multiple directional queries are utilized.

Second, to resolve the query allocation problem, we derive convergence rates for algorithms using both ZO-Align and ZO-Avg, making the dependence on the number of queries used per iteration explicit. We then leverage these rates to formally pose and solve an optimization problem: given a fixed total query budget, determine the allocation of queries per iteration that minimizes the final optimization error. To ensure the generality and robustness of our conclusions, this theoretical framework is systematically applied across a broad spectrum of problem structures, including strongly convex, convex, non-convex, and stochastic settings.

Our analysis reveals a stark dichotomy in optimal allocation strategies, critically dependent on the chosen aggregation method. For ZO-Avg, we prove that it cannot effectively utilize information from multiple queries: under a fixed budget, any attempt to use more than one query per iteration is detrimental, and the optimal strategy is always the single-query approach. In contrast, the ZO-Align estimator combines multi-query information in a geometry-aware way that avoids wasting queries on redundant directions, and in strongly convex and stochastic problems with initialization-dominated regimes, multi-query ZO-Align is more query-efficient than any single-query scheme. We further analyze the regime where the ambient dimension is much larger than the query block size, showing that ZO-Avg continues to follow the same single-query allocation rule, while ZO-Align degenerates into a step direction that is nearly colinear with a suitably rescaled ZO-Avg update due to the near-orthogonality of random directions in high dimensions.

To complement our theoretical analysis, we conducted extensive experiments across a suite of optimization problems. The empirical results consistently validate our theoretical findings, illustrating the practical implications of the query allocation dichotomy. Our work clarifies that the central design choice is not about tuning an intermediate query size, but rather about committing to an aggregation method that dictates whether the optimal algorithm is a sequential single-query method or a full-subspace estimation procedure.

The primary contributions of this work are as follows:

- Our work is the first to formally formulate the fixed-budget query-allocation problem as a distinct research question in zeroth-order optimization, and to provide a systematic theoretical framework for analyzing it.
- We provide rigorous theoretical proofs that under a fixed query budget, the optimal query allocation for ZO-Avg is always single-query, whereas ZO-Align can make better use of multi-query information, resolving the query allocation question for these two classes of estimators.
- We empirically validate our theoretical findings across a range of optimization problems, and also include high-dimensional tasks that illustrate how our query-allocation insights manifest in large-scale problems.

2 Related Work

ZO gradient estimation. ZO methods approximate gradients using finite-difference techniques, a concept explored through various random perturbation strategies. For instance, theoretical convergence has been demonstrated using Gaussian random noise perturbations (Nesterov & Spokoiny,

2017), while other approaches have utilized uniform sampling (Flaxman et al., 2004) and coordinate-wise perturbations (Lian et al., 2016). A distinct line of research focuses on problems where the gradient is assumed to be sparse, framing the estimation task as one of compressed sensing; here, techniques like LASSO (Wang et al., 2018) or other forms of regularized regression (Choromanski et al., 2020) enable the reconstruction of a sparse gradient from a limited number of function queries (Cai et al., 2022).

Query-efficient ZO optimizers. A central challenge in zeroth-order optimization is its high query complexity, which has motivated a broad range of techniques aimed at improving query efficiency. Some methods construct more accurate local models for the update direction by incorporating second-order information (Ye et al., 2025) or by exploiting problem geometry via Riemannian optimization (He et al., 2024). Other approaches focus on optimizing the querying process itself, for example by using compressed sensing to achieve double-logarithmic query complexity in sparse settings (Qiu & Tong, 2024), adaptively reusing past queries to reduce variance (Xiao et al., 2023), or learning query directions through reinforcement learning (Zhai et al., 2024). Despite their differences, these approaches share a common focus: they seek to improve the quality of the gradient estimate at each individual iteration under a fixed number of queries per step.

In contrast, our work takes a complementary perspective. Rather than designing yet another estimator for a given per-iteration query budget, we ask how to optimally allocate a fixed total query budget across iterations and aggregation schemes. Our analysis shows that, even for two canonical multi-query estimators, this allocation problem exhibits a sharp dichotomy: the optimal query-per-iteration choice collapses to opposite extremes depending on the aggregation method used.

3 ZO Estimators: From Single-Query to Multi-Query

This section introduces the core ZO gradient estimators used in our analysis. We begin with the standard single-query estimator before presenting two multi-query approaches: the intuitive Naive Averaging (ZO-Avg) and the principled Projection Alignment (ZO-Align), which we derive from local surrogate minimization. Then we analyze their Mean Squared Error (MSE) to establish a fundamental difference in their accuracy, which underpins the main convergence results in Section 4.

3.1 Preliminaries and the Single-Query Estimator

We consider the unconstrained optimization problem $\min_{x\in\mathbb{R}^d} f(x)$ where direct access to the gradient $\nabla f(x)$ is unavailable. Throughout this work, we assume the objective function f is differentiable and L-smooth:

$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|, \quad \forall x, y \in \mathbb{R}^d.$$

ZO algorithms navigate this problem by iteratively updating a parameter vector x_t using a gradient estimate $\hat{g}(x_t)$ constructed from function value queries: $x_{t+1} = x_t - \eta_t \hat{g}(x_t)$.

The foundational building block for most ZO methods is the single-query estimator. It approximates the gradient by evaluating the function along a single random direction $u \in \mathbb{R}^d$ using a *one-sided* finite-difference scheme. A common form is given by:

$$\hat{g}_u(x) = \frac{f(x + \mu u) - f(x)}{\mu} u,$$

where $\mu > 0$ is a small smoothing parameter. Throughout this work, we focus on an idealized regime in which $\mu \to 0$ and the estimator approximates the projection of the true gradient onto the direction u:

$$\hat{g}_u(x) \overset{\mu\to 0}{\approx} (u^T \nabla f(x))u.$$

All MSE and convergence results below are derived under this idealization. Appendix B studies finite-smoothing effects and empirically characterizes how the validity of this approximation depends on the ambient dimension. The same geometric analysis applies to *two-sided* finite-difference estimators, which differ from the one-sided form only in the higher-order μ-dependent terms Nesterov & Spokoiny (2017).

While fundamental, this single-direction estimate is inherently noisy, limiting its effectiveness. A natural extension is to probe the function along multiple directions to construct a more accurate gradient estimate.

3.2 Simple Averaging Estimator (ZO-Avg)

The most direct method for aggregating information from a block of q queries is to compute q independent single-query estimates and average them. Let $U = [u_1, \ldots, u_q] \in \mathbb{R}^{d\times q}$ be a matrix whose columns are random direction vectors drawn independently from $\mathcal{N}(0, I)$. We let $q \leq d$ and assume the columns of U are linearly independent. The averaging estimator, which we term ZO-Avg, is defined as:

$$\hat{g}_{\text{AVG}}(x) = \frac{1}{q}\sum_{i=1}^{q} \frac{f(x+\mu u_i) - f(x)}{\mu} u_i \overset{\mu\to 0}{\approx} \left(\frac{1}{q}\sum_{i=1}^{q} u_i u_i^T\right)\nabla f(x) = \frac{1}{q} UU^T \nabla f(x). \quad (1)$$

This estimator is appealing due to its simplicity and straightforward parallel implementation.

3.3 Projection Alignment Method (ZO-Align)

Instead of simply averaging, we adopt a more principled, model-based perspective. The q queries provide a snapshot of the function's local geometry within the column space of U. We seek the best update step of the form $-Uy$ for some coefficient vector $y \in \mathbb{R}^q$ by minimizing a local surrogate of the objective function. Given that f is L-smooth, we can form a local quadratic upper bound around the point x:

$$f(x - Uy) \leq f(x) - \nabla f(x)^T (Uy) + \frac{L}{2}\|Uy\|^2.$$

Minimizing this quadratic upper bound with respect to y is a principled way to find the best local step within the column space. This yields the optimal coefficient vector:

$$y^* = \frac{1}{L}(U^T U)^{-1} U^T \nabla f(x).$$

The optimal update direction is therefore proportional to $Uy^* = U(U^TU)^{-1}U^T\nabla f(x)$. This motivates our second estimator, which replaces the true directional derivatives with their finite-difference approximations:

$$\hat{g}_{\text{ALG}}(x) = U(U^TU)^{-1} \begin{bmatrix} \frac{f(x+\mu u_1)-f(x)}{\mu} \\ \vdots \\ \frac{f(x+\mu u_q)-f(x)}{\mu} \end{bmatrix}. \quad (2)$$

The name "Projection Alignment" comes from the powerful geometric property of this estimator. The matrix $P_U = U(U^TU)^{-1}U^T$ is the orthogonal projection operator onto the column space of U. Therefore, assuming $\mu \to 0$, $\hat{g}_{\text{ALG}}$ is the orthogonal projection of the true gradient onto the query subspace. Exact ZO-Align requires forming and solving the $q \times q$ Gram system, incurring $O(dq^2 + q^3)$ arithmetic in a direct implementation. However, at the full-subspace endpoint $q = d$, the projection is independent of the particular full-rank basis and one may choose $U = I_d$, reducing the method to standard coordinate-wise finite differences without a dense Gram solve.

This reveals a crucial difference from ZO-Avg. While ZO-Align is constructed to be geometrically consistent within the sampled subspace, ZO-Avg is not. The projection of the ZO-Avg estimate onto a query direction u_i does not generally match the true directional derivative, leading to a less accurate representation of the local geometry.

3.4 Accuracy Analysis: A Tale of Two Estimators

To formalize this comparison, we analyze the MSE of each estimator, defined as $\mathbb{E}[\|\hat{g}(x) - \nabla f(x)\|^2]$, which captures the total error from both bias and variance.

Proposition 3.1. *Let $g = \nabla f(x)$, and assume the query directions are i.i.d. samples from $\mathcal{N}(0, I)$. Then the idealized estimators satisfy*

$$\text{MSE}(\hat{g}_{\text{AVG}}) = \frac{d+1}{q}\|g\|^2, \qquad \text{MSE}(\hat{g}_{\text{ALG}}) = \frac{d-q}{d}\|g\|^2.$$

Proof. See Appendix C.2. □

For any $q > 1$, it is clear that $\frac{d-q}{d} < \frac{d+1}{q}$, indicating that ZO-Align offers a significantly lower estimation error. This fundamental difference in estimator quality is the primary driver of the performance gap we will establish in the following sections. It highlights that the principled, model-based approach of ZO-Align is far more effective at aggregating information from multiple queries than the naive averaging of ZO-Avg.

4 Query Allocation Dichotomy

This section presents our main theoretical results, establishing a fundamental dichotomy in how multi-query ZO estimators should be used under a fixed total query budget. We focus on the deterministic setting with exact function evaluations, and analyze the convergence behaviour of algorithms based on ZO-Avg and ZO-Align for three standard classes of objectives: strongly convex, convex, and non-convex. Across these settings, we show that simple averaging is an inefficient way to aggregate query information: for ZO-Avg, the optimal strategy is always to use a single query per iteration. In contrast, the projection-based construction of ZO-Align can exploit multi-query information much more effectively, leading to fundamentally different optimal allocation rules. A stochastic convex extension is deferred to Appendix D, where stochastic variance introduces an additional regime-dependent query-allocation trade-off.

Suppose we are given a fixed total query budget K. An algorithm performs a series of updates over T iterations, and at each iteration t it uses a set of q_t function queries as stated in Section 3.2. The total number of queries is therefore constrained by $\sum_{t=0}^{T-1} q_t \leq K$. Our central goal is to determine the optimal query allocation strategy—the sequence of queries per iteration, $\{q_t\}_{t=0}^{T-1}$—that minimizes the final optimization error.

To analyze this allocation problem, we consider the update rule $x_{t+1} = x_t - \eta_t \hat{g}(x_t)$, where $\hat{g}$ is either the ZO-Avg or ZO-Align estimator constructed from q_t query directions at iteration t, as described in Section 3. For a given estimator and query block size q_t, we first choose the step size η_t by maximizing the standard L-smooth guaranteed decrease. The following proposition summarizes the resulting choices, which will be used throughout this section.

Proposition 4.1. *Let f be L-smooth, and consider a single update $x^+ = x - \eta\hat{g}(x)$. In the idealized regime $\mu \to 0$, the step size that maximizes the guaranteed descent bound is*

$$\eta^\star_{\mathrm{AVG}}(q) = \frac{q}{L(q+d+1)} \quad \textit{for } \hat{g} = \hat{g}_{\mathrm{AVG}}, \qquad \eta^\star_{\mathrm{ALG}}(q) = \frac{1}{L} \quad \textit{for } \hat{g} = \hat{g}_{\mathrm{ALG}}.$$

Proof. See Appendix C.3. □

4.1 Strongly Convex Case

We begin with the strongly convex case. A function is defined as γ-strongly convex if there exists a constant γ such that for all x and y in its domain:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{\gamma}{2}\|y - x\|^2.$$

When the objective function $f(x)$ is both L-smooth and γ-strongly convex, we can achieve a linear convergence rate (Boyd & Vandenberghe, 2004). However, the rate's dependence on the number of queries per iteration, q_t, reveals the core dichotomy of our work. We first analyze the ZO-Avg estimator.

Theorem 4.2. *Let f be an L-smooth and γ-strongly convex function. The ZO-Avg algorithm with the optimal step size stated in Prop 4.1 achieves the following convergence guarantee:*

$$\mathbb{E}[f(x_T) - f(x^*)] \leq \left(\prod_{t=0}^{T-1} \left(1 - \frac{\gamma q_t}{L(q_t + d + 1)} \right) \right) (f(x_0) - f(x^*)), \tag{3}$$

where x^* represents the optimum. Under a fixed total query budget K, this convergence rate is optimized by choosing $q_t = 1$ for all $t = 0, \ldots, K-1$, resulting in $T = K$ iterations.

Proof. See Appendix C.4. □

The result in Theorem 4.2 suggests that when using the simple averaging estimator, allocating more than one query per iteration is detrimental to the convergence rate. The core reason lies in the trade-off between variance reduction and the cost of queries. While increasing q_t does reduce the variance of the gradient estimate, the improvement scales sublinearly $1/q_t$, as shown in Section 3.4. However, the cost increases linearly with q_t. The analysis shows that the marginal benefit of reducing variance is not worth the cost of the additional queries. It is always better to spend the query budget on taking more, albeit noisier, steps. This exposes a fundamental flaw in the simple average strategy: it is an inefficient use of a limited query budget.

In stark contrast, the principled structure of the ZO-Align estimator leads to the opposite conclusion.

Theorem 4.3. *Let f be an L-smooth and γ-strongly convex function. The ZO-Align algorithm with the optimal step size $\eta_t = 1/L$ achieves the following convergence guarantee:*

$$\mathbb{E}[f(x_T) - f(x^*)] \leq \left(\prod_{t=0}^{T-1} \left(1 - \frac{\gamma q_t}{Ld}\right) \right) (f(x_0) - f(x^*)), \tag{4}$$

where x^* represents the optimum. Under a fixed total query budget K, this convergence rate is optimized by using as few iterations as possible with the maximum number of queries, i.e., $q_t = d$ for $T = K/d$ iterations.

Proof. See Appendix C.5. □

For ZO-Align, the optimal strategy is to use the maximum number of queries per step. The intuition is that ZO-Align does not simply average away noise; it constructs a principled, low-error projection of the gradient onto the query subspace. Optimization error is reduced multiplicatively at each step, meaning a more accurate gradient provides an exponential advantage over time. It is therefore more effective to spend the budget forming a single high-quality step than to take many uncertain steps.

4.2 Convex Case

We then analyze the case where the objective function $f(x)$ is convex and L-smooth, but not necessarily strongly convex. We demonstrate that the optimal query allocation strategy is fundamentally different for the two estimators, reinforcing the core dichotomy of our work.

Theorem 4.4. *Let f be an L-smooth and convex function. The ZO-Avg algorithm with step size stated in Prop 4.1 achieves the following convergence guarantee after T iterations:*

$$\mathbb{E}[f(x_T) - f(x_*)] \leq \frac{L(\|x_0 - x_*\|^2 + \frac{2}{L}(f(x_0) - f(x_*)))}{\sum_{t=0}^{T-1} \frac{2q_t}{q_t + d + 1}}. \tag{5}$$

Under a fixed total query budget K, this convergence rate is optimized by choosing $q_t = 1$ for all t.

Proof. See Appendix C.6. □

Theorem 4.4 reinforces the conclusion from the strongly convex case: for the ZO-Avg estimator, using multiple queries per step is an inefficient use of the query budget. Next, we show that for the ZO-Align estimator, the conclusion is different.

Theorem 4.5. *Let f be an L-smooth and convex function. The ZO-Align algorithm with step size $\eta_t = 1/L$ achieves the following rate under any allocation $\{q_t\}$ such that $\sum_{t=0}^{T-1} q_t = K$:*

$$\mathbb{E}[f(x_T) - f(x_*)] \leq \frac{d}{2K}(L\|x_0 - x_*\|^2 + 2(f(x_0) - f(x_*))). \tag{6}$$

Proof. See Appendix C.7. □

4.3 NON-CONVEX CASE

Then we proceed to the non-convex setting, where the objective is only assumed to be L-smooth. In this scenario, the goal is typically to find a first-order stationary point (Nocedal & Wright, 2006), i.e., a point x where $\|\nabla f(x)\|^2 \leq \epsilon$. The convergence is measured by the number of queries required to achieve this goal. The analysis reveals the same fundamental dichotomy between the two estimators.

Theorem 4.6. *Let f be an L-smooth function. The ZO-Avg algorithm with step size $\eta_t = \frac{q_t}{L(q_t+d+1)}$ ensures that:*

$$\min_{t=0,\ldots,T-1} \mathbb{E}[\|\nabla f(x_t)\|^2] \leq \frac{2L(f(x_0) - f^*)}{\sum_{t=0}^{T-1} \frac{q_t}{q_t+d+1}}, \tag{7}$$

where f^* is the minimum value of f. Under a fixed total query budget K, this rate is optimized by choosing $q_t = 1$ for all t.

Proof. See Appendix C.8. □

Theorem 4.7. *Let f be an L-smooth function. The ZO-Align algorithm with step size $\eta_t = 1/L$ ensures that:*

$$\min_{t\in\{0,1\ldots,T-1\}} \mathbb{E}[\|\nabla f(x_t)\|^2] \leq \frac{2Ld(f(x_0) - f^*)}{K} \tag{8}$$

under any allocation $\{q_t\}$ that exhausts the budget, $\sum_{t=0}^{T-1} q_t = K$.

Proof. See Appendix C.9. □

4.4 THE MULTI-QUERY PARADOX

The theoretical results reveal an aggregation-dependent allocation pattern under a fixed total query budget. For ZO-Avg, the optimal allocation is to take a single query per iteration, so that the budget is spent on as many steps as possible. For ZO-Align, the picture is the opposite in the strongly convex case, where the best strategy is to use as many queries as possible per step (up to d when feasible), while in the convex and non-convex cases the convergence rate depends only on the total number of queries and is essentially indifferent to how they are distributed across iterations.

This dichotomy has a clear geometric interpretation. ZO-Avg forms its update by averaging directional derivatives along a block of sampled directions, but it never accounts for the geometry of these directions. If several query vectors happen to be close to each other, their information is effectively counted multiple times, while directions that would explore new parts of the space may be underrated. Additional queries therefore provide diminishing returns: they mainly reduce variance through repeated measurements along nearly redundant directions, and this variance reduction scales sublinearly in the block size. Once the linear cost in queries is taken into account, it is always better for ZO-Avg to take more, cheaper steps with a single direction per step than to aggregate many directions at once.

By contrast, ZO-Align uses the entire block of directional measurements to reconstruct the orthogonal projection of the gradient onto the span of the sampled directions. The correlations and unequal norms among the queries are compensated, and the resulting update is the unique vector in the sampled subspace whose projections match the observed directional derivatives. As a consequence, each additional query genuinely enlarges and refines the subspace in which the method can make informed progress, rather than merely re-averaging noisy measurements along similar directions. This difference in how query information is used explains why ZO-Align benefits from larger blocks in the strongly convex case, where every improvement in the quality of the update translates into a better contraction factor and hence compounds multiplicatively over iterations. In the limit where the block size equals the dimension and the query matrix has full rank, ZO-Align recovers the true gradient and reduces to a classical full-subspace block coordinate finite-difference method.

5 HIGH-DIMENSIONAL REGIME

Modern applications such as large-scale neural network fine-tuning often operate in a regime where the ambient dimension d is extremely large, whereas the query batch size q is constrained to be much

smaller than d. In this setting, when the columns of $U \in \mathbb{R}^{d\times q}$ are i.i.d. samples from $\mathcal{N}(0, I)$, the Gram matrix $U^\top U$ concentrates sharply around a scaled identity: $U^\top U \approx d\, I_q$ when $d \gg q$, so that the sampled directions are nearly mutually orthogonal and $(U^\top U)^{-1}$ is, to leading order, a scalar multiple of the identity. A rigorous analysis and empirical verification of the approximation are provided in Appendix A.

This observation clarifies how the benefits of ZO-Align evolve in high dimensions. The main advantage of ZO-Align over simple averaging is that it explicitly accounts for correlations among the sampled directions through the factor $(U^\top U)^{-1}$. When d is moderate and q is a nontrivial fraction of d, the Gram matrix $U^\top U$ can deviate substantially from a scaled identity, and this reweighting meaningfully corrects for the non-orthogonality and unequal norms of the sampled directions. In contrast, when $d \gg q$ and $U^\top U \approx dI_q$, the sampled directions are already almost orthogonal and have comparable lengths, so the correction degenerates to an almost uniform scaling. As a result, the geometric distinction between ZO-Align and ZO-Avg within the sampled subspace becomes much less pronounced in this extreme high-dimensional regime.

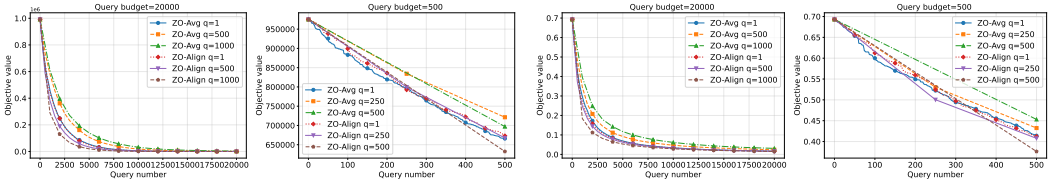

Figure 1: Objective with query used: Strongly convex case.

Figure 2: Objective with query used: Convex case.

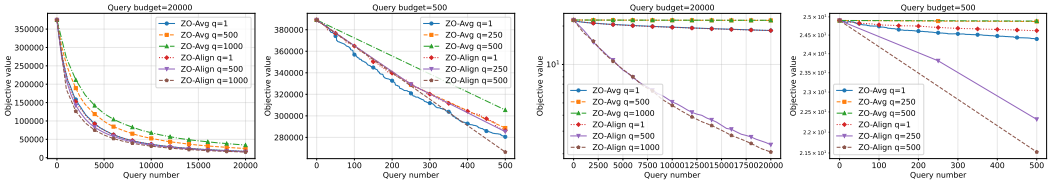

Figure 3: Objective with query used: Nonconvex case.

Figure 4: Objective with query used: Stochastic case.

6 Experiments

To complement our theoretical analysis, we evaluate ZO-Avg and ZO-Align under fixed query budgets on both classical objectives, LQR problem and LLM fine-tuning tasks.

6.1 Classical Optimization Problems

We first consider four standard objectives at dimension $d = 1000$, representing the regimes studied in Section 4: a strongly convex quadratic, convex logistic regression on synthetic data, the non-convex Rosenbrock function, and a stochastic convex logistic regression objective (details in Appendix E). For each problem we run ZO-Avg and ZO-Align with several query sizes q per iteration under two total budgets, $K = 20{,}000$ and $K = 500$ (the latter precluding a single full-subspace step). Each configuration is averaged over multiple random seeds.

Across all four problems, the empirical behaviour closely matches our theoretical allocation results. For ZO-Avg, the best performance under a fixed budget is consistently obtained with $q = 1$, and increasing q monotonically harms query efficiency, in line with our Theorems. For ZO-Align, larger query blocks yield clear gains in the strongly convex and stochastic convex cases, where the convergence factors depend linearly on q, while in the deterministic convex and non-convex problems the curves for different q are much closer, reflecting the allocation-insensitive rates. Overall, these experiments confirm the query allocation dichotomy of Section 4.4: ZO-Avg prefers single-query steps, whereas ZO-Align can exploit larger blocks without the collapse in efficiency seen for ZO-Avg.

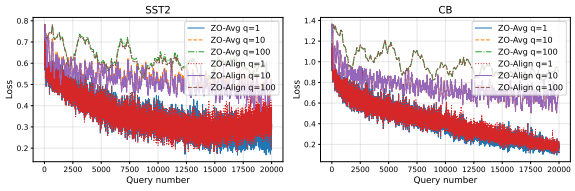

Figure 5: Finetuning Qwen3-0.6b on SST2 and CB.

Table 1: Queries required to reach a target optimality gap on the LQR problem.

Gap	ZO-Avg $q=1$	ZO-Avg $q=3$	ZO-Avg $q=9$	ZO-Align $q=1$	ZO-Align $q=3$	ZO-Align $q=9$
10^{-3}	759	999	1413	745	699	612
10^{-5}	1298	1575	2250	1250	1116	1017
10^{-7}	1870	2211	3105	1819	1566	1422
10^{-9}	2378	2583	3987	2328	1980	1845

6.2 LQR Control Problem

We next evaluate the query-allocation dichotomy on a Linear Quadratic Regulator (LQR) problem, following Section 2 of (Malik et al., 2019). The exact formulation is given in Appendix E. In this experiment, the optimization variable is the feedback controller, and the objective is the closed-loop rollout cost induced by that controller. This makes LQR a natural testbed for zeroth-order methods, since evaluating one query means evaluating one candidate controller through system simulation, rather than reading off a gradient from a static objective. Although the parameter dimension is modest ($d = 9$), the problem is still nontrivial from a ZO perspective. In our implementation, evaluating a single controller requires rolling out the closed-loop dynamics for 200 time steps. The main challenge is therefore not high dimensionality, but the trade-off between spending more queries within one iteration to obtain a better update direction and saving queries for more optimization steps.

Table 1 reports the number of queries required to reach several target optimality-gap thresholds. The pattern is fully consistent with our theory. For ZO-Avg, larger query blocks are harmful. In contrast, ZO-Align benefits from larger blocks.

6.3 High-Dimensional Language Model Fine-Tuning

We next examine a high-dimensional setting where the parameter dimension d is extremely large ($q \ll d$). We fine-tune the Qwen3-0.6B model on two sentence classification benchmarks, SST-2 and CB. Further protocol details are given in Appendix E.

We compare ZO-Avg with an approximate variant of ZO-Align adapted to the high-dimensional regime. For ZO-Align, forming and inverting the full Gram matrix $U^\top U$ is prohibitively expensive at this scale. Motivated by the concentration result in Section 5, we approximate the Gram matrix by its diagonal. We fix a total query budget of $K = 20{,}000$ function evaluations and vary the per-iteration query count over $q \in \{1, 10, 100\}$. The step size is tuned separately for each setting by grid search.

As shown on Figure 5, on both SST-2 and CB, ZO-Avg continues to favour small query blocks: the best optimization performance per query is achieved for $q = 1$, and increasing q consistently slows progress under the fixed budget. The diagonal ZO-Align variant closely tracks ZO-Avg across all choices of q, with the two sets of curves almost indistinguishable at the scale of our plots. This behaviour matches the high-dimensional analysis in Section 5: when $d \gg q$, the Gram matrix concentrates near a scaled identity, so the projection step in ZO-Align reduces essentially to a rescaled version of the ZO-Avg update. In large-scale LM fine-tuning, the query allocation dichotomy for ZO-Avg therefore remains clearly visible, while the additional geometric structure of ZO-Align provides little practical gain beyond what can be achieved by a suitably tuned ZO-Avg scheme.

7 Conclusions

This work addresses two core questions in zeroth-order optimization: how to aggregate information from multiple queries and how to allocate a fixed query budget. By analyzing simple averaging ZO-Avg and the projection-based ZO-Align estimator across different settings, we show that aggregation and allocation are tightly coupled and lead to a sharp dichotomy. Under a fixed budget, the optimal strategy for ZO-Avg is always to use single-query steps, whereas ZO-Align can benefit from larger query blocks. Experiments validate the theoretical analysis. Our results resolve the multi-query paradox by clarifying that the central design choice is not to tune an intermediate query size, but to select an aggregation rule that determines whether the canonical algorithm is a sequential single-query scheme or a projection-based multi-query method.

AI DISCLOSURE

Generative AI tools were used to assist with software implementation, to edit the manuscript for clarity and readability, and to check mathematical derivations and proofs for potential errors and internal consistency. AI-assisted code was reviewed, tested, and verified by the authors before being used in the experiments, and all AI-assisted text and proof checks were independently reviewed by the authors. The authors take full responsibility for the final content of this work, including the code, experimental results, mathematical claims, and written presentation.

REFERENCES

Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

HanQin Cai, Daniel McKenzie, Wotao Yin, and Zhenliang Zhang. Zeroth-order regularized optimization (zoro): Approximately sparse gradients and adaptive sampling. *SIAM Journal on Optimization*, 32(2):687–714, 2022.

Krzysztof Choromanski, Aldo Pacchiano, Jack Parker-Holder, Yunhao Tang, Deepali Jain, Yuxiang Yang, Atil Iscen, Jasmine Hsu, and Vikas Sindhwani. Provably robust blackbox optimization for reinforcement learning. In *Conference on robot learning*, pp. 683–696. PMLR, 2020.

Andrew R Conn, Katya Scheinberg, and Luis N Vicente. *Introduction to derivative-free optimization*. SIAM, 2009.

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.

Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*, 2004.

Chang He, Zhaoye Pan, Xiao Wang, and Bo Jiang. Riemannian accelerated zeroth-order algorithm: improved robustness and lower query complexity. *arXiv preprint arXiv:2405.05713*, 2024.

Beatrice Laurent and Pascal Massart. Adaptive estimation of a quadratic functional by model selection. *Annals of statistics*, pp. 1302–1338, 2000.

Yanan Li, Xuebin Ren, Fangyuan Zhao, and Shusen Yang. A zeroth-order adaptive learning rate method to reduce cost of hyperparameter tuning for deep learning. *Applied Sciences*, 11(21):10184, 2021.

Xiangru Lian, Huan Zhang, Cho-Jui Hsieh, Yijun Huang, and Ji Liu. A comprehensive linear speedup analysis for asynchronous stochastic parallel optimization from zeroth-order to first-order. *Advances in neural information processing systems*, 29, 2016.

Dhruv Malik, Ashwin Pananjady, Kush Bhatia, Koulik Khamaru, Peter Bartlett, and Martin Wainwright. Derivative-free methods for policy optimization: Guarantees for linear quadratic systems. In *The 22nd international conference on artificial intelligence and statistics*, pp. 2916–2925. PMLR, 2019.

Sadhika Malladi, Tianyu Gao, Eshaan Nichani, Alex Damian, Jason D Lee, Danqi Chen, and Sanjeev Arora. Fine-tuning language models with just forward passes. *Advances in Neural Information Processing Systems*, 36:53038–53075, 2023.

Horia Mania, Aurelia Guy, and Benjamin Recht. Simple random search of static linear policies is competitive for reinforcement learning. *Advances in neural information processing systems*, 31, 2018.

Yurii Nesterov and Vladimir Spokoiny. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017.

Jorge Nocedal and Stephen J Wright. *Numerical optimization*. Springer, 2006.

Ruizhong Qiu and Hanghang Tong. Gradient compressed sensing: A query-efficient gradient estimator for high-dimensional zeroth-order optimization. *arXiv preprint arXiv:2405.16805*, 2024.

James C Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE transactions on automatic control*, 37(3):332–341, 2002.

Hao Sun, Ziping Xu, Yuhang Song, Meng Fang, Jiechao Xiong, Bo Dai, and Bolei Zhou. Zeroth-order supervised policy improvement. *arXiv preprint arXiv:2006.06600*, 2020.

Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

Yining Wang, Simon Du, Sivaraman Balakrishnan, and Aarti Singh. Stochastic zeroth-order optimization in high dimensions. In *International conference on artificial intelligence and statistics*, pp. 1356–1365. PMLR, 2018.

Quan Xiao, Qing Ling, and Tianyi Chen. Lazy queries can reduce variance in zeroth-order optimization. *IEEE Transactions on Signal Processing*, 71:3695–3709, 2023.

Haishan Ye, Zhichao Huang, Cong Fang, Chris Junchi Li, and Tong Zhang. Hessian-aware zeroth-order optimization. *IEEE transactions on pattern analysis and machine intelligence*, 2025.

Zhou Zhai, Wanli Shi, Heng Huang, Yi Chang, and Bin Gu. Learning sampling policy to achieve fewer queries for zeroth-order optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 1162–1170. PMLR, 2024.

Yihua Zhang, Pingzhi Li, Junyuan Hong, Jiaxiang Li, Yimeng Zhang, Wenqing Zheng, Pin-Yu Chen, Jason D Lee, Wotao Yin, Mingyi Hong, et al. Revisiting zeroth-order optimization for memory-efficient llm fine-tuning: A benchmark. *arXiv preprint arXiv:2402.11592*, 2024.

Appendix

A High-Dimensional Behaviour of ZO-Align

In this appendix we analyze the behaviour of ZO-Align in the regime where the dimension d is very large while the query batch size q remains much smaller than d. Throughout, we retain the notation from the main text: $U \in \mathbb{R}^{d\times q}$ collects the q query directions as columns, each sampled i.i.d. from $\mathcal{N}(0, I_d)$, and

$$P_U \;=\; U(U^\top U)^{-1}U^\top$$

denotes the orthogonal projection onto the column space of U. In the idealized limit $\mu \to 0$, ZO-Align can be written as

$$\hat{g}_{ALG}(x) \;=\; P_U \nabla f(x),$$

so understanding the high-dimensional behaviour of ZO-Align reduces to understanding the random projector P_U.

A.1 Concentration of the Gram Matrix

It is convenient to view U row-wise. Let $r_1, \dots, r_d \in \mathbb{R}^q$ denote the rows of U. Since the columns of U are i.i.d. $\mathcal{N}(0, I_d)$, each row r_k is an independent sample from $\mathcal{N}(0, I_q)$ and

$$U^\top U = \sum_{k=1}^{d} r_k^\top r_k, \qquad \frac{1}{d}U^\top U = \frac{1}{d}\sum_{k=1}^{d} r_k^\top r_k$$

is the empirical covariance matrix of d i.i.d. Gaussian samples in $\mathbb{R}^q$.

The following standard result quantifies the concentration of this empirical covariance around the identity.

Lemma A.1. *There exist absolute constants $c, C > 0$ such that the following holds. Let $U \in \mathbb{R}^{d\times q}$ have i.i.d. $\mathcal{N}(0,1)$ entries. Then for all $t \in (0,1)$,*

$$\Pr\left(\left\|\frac{1}{d}U^\top U - I_q\right\|_{op} \;\geq\; C\Big(\sqrt{\frac{q}{d}} + t\Big)\right) \;\leq\; 2\exp(-cdt^2).$$

In particular, if $d \gg q$ then with probability at least $1 - 2\exp(-c'q)$,

$$\left\|\frac{1}{d}U^\top U - I_q\right\|_{op} = O\left(\sqrt{\frac{q}{d}}\right).$$

Proof. This is a standard consequence of sample covariance concentration for isotropic sub-Gaussian vectors; see, e.g., the general treatments in Chapter 5 of Vershynin (2018). For completeness, we specialize the argument to the Gaussian case.

Row representation and reduction to one-dimensional marginals. Write U row-wise as

$$U = \begin{bmatrix} r_1^\top \\ \vdots \\ r_d^\top \end{bmatrix}, \qquad r_k \in \mathbb{R}^q, \quad r_k \sim \mathcal{N}(0, I_q) \text{ i.i.d.}$$

Then

$$\frac{1}{d}U^\top U = \frac{1}{d}\sum_{k=1}^{d} r_k r_k^\top$$

is the empirical covariance matrix of d i.i.d. Gaussian samples in $\mathbb{R}^q$.

Let

$$\Sigma_d := \frac{1}{d}U^\top U, \qquad A := \Sigma_d - I_q.$$

We need to control $\|A\|_{op}$. For any $x \in \mathbb{R}^q$ with $\|x\| = 1$, we have

$$x^\top Ax = x^\top \Sigma_d x - x^\top x = \frac{1}{d}\sum_{k=1}^{d} \langle r_k, x\rangle^2 - 1.$$

Define

$$Z_k(x) := \langle r_k, x\rangle, \qquad Y_k(x) := Z_k(x)^2 - 1.$$

For fixed x with $\|x\| = 1$, $Z_k(x) \sim \mathcal{N}(0,1)$ i.i.d., so

$$\sum_{k=1}^{d} Z_k(x)^2 \sim \chi^2_d.$$

Consequently, $\frac{1}{d}\sum_{k=1}^{d} Z_k(x)^2$ concentrates around 1, and we can use a sharp tail inequality for the chi-square distribution due to Laurent and Massart Laurent & Massart (2000).

Let $S(x) := \sum_{k=1}^{d} Z_k(x)^2$. By definition,

$$x^\top Ax = \frac{1}{d}S(x) - 1.$$

Eq. 4.3-4.4 in Laurent & Massart (2000) show that if $S \sim \chi^2_d$, then for all $u > 0$,

$$\Pr\big(S - d \geq 2\sqrt{du} + 2u\big) \leq e^{-u},$$
$$\Pr\big(d - S \geq 2\sqrt{du}\big) \leq e^{-u}.$$

Translating these inequalities to $\frac{1}{d}S(x) - 1$ yields: for all $u > 0$,

$$\Pr\left(\frac{1}{d}S(x) - 1 \geq 2\sqrt{\frac{u}{d}} + 2\frac{u}{d}\right) \leq e^{-u},$$
$$\Pr\left(1 - \frac{1}{d}S(x) \geq 2\sqrt{\frac{u}{d}}\right) \leq e^{-u}.$$

Taking $u \in (0, d)$, we have $\frac{u}{d} \leq 1$, so there exists an absolute constant $C_0 > 2$ such that

$$2\sqrt{\frac{u}{d}} + 2\frac{u}{d} \leq C_0\left(\sqrt{\frac{u}{d}}\right).$$

Hence, for all $u \in (0, d)$,

$$\Pr\left(\left|\frac{1}{d}S(x) - 1\right| \geq C_0\sqrt{\frac{u}{d}}\right) \leq 2e^{-u}.$$

Equivalently, for all $u \in (0, d)$,

$$\Pr\left(\left|x^\top Ax\right| \geq C_0\sqrt{\frac{u}{d}}\right) \leq 2e^{-u}. \tag{9}$$

Now restrict to thresholds $u = c_1 dv^2$ with $v \in (0, 1)$ and $c_1 > 0$ small enough so that $u < d$. Then

$$\sqrt{\frac{u}{d}} = \sqrt{c_1}\, v,$$

and equation 9 becomes

$$\Pr\big(\left|x^\top Ax\right| \geq C_1 v\big) \leq 2\exp(-c_1 dv^2),$$

for constants $C_1 = C_0\sqrt{c_1}$ and $c_1 > 0$. Renaming v to u and c_1 to c_2, we obtain:

$$\Pr\big(\left|x^\top Ax\right| \geq u\big) \leq 2\exp(-c_2 du^2), \qquad \forall\, u \in (0,1),\ \forall\, x \in \mathbb{S}^{q-1}, \tag{10}$$

for some absolute constant $c_2 > 0$ and where $\mathbb{S}^{q-1} = \{x \in \mathbb{R}^q : \|x\| = 1\}$.

Let $\mathbb{S}^{q-1}$ denote the unit sphere in $\mathbb{R}^q$. For $\varepsilon \in (0,1)$, an ε-net $\mathcal{N}_\varepsilon \subset \mathbb{S}^{q-1}$ is a finite set such that for any $x \in \mathbb{S}^{q-1}$ there exists $y \in \mathcal{N}_\varepsilon$ with $\|x-y\| \le \varepsilon$. It is well known that there exists an ε-net with cardinality bounded by

$$|\mathcal{N}_\varepsilon| \le \left(1+\frac{2}{\varepsilon}\right)^q.$$

We fix $\varepsilon = \frac{1}{4}$ and let $\mathcal{N}$ be a $\frac{1}{4}$-net; then

$$|\mathcal{N}| \le 9^q.$$

We also use a standard net approximation for the operator norm of a symmetric matrix.

Claim A.2. Let A be a symmetric $q \times q$ matrix and let $\mathcal{N}$ be a $\frac{1}{4}$-net of $\mathbb{S}^{q-1}$. Then

$$\|A\|_{op} = \sup_{x \in \mathbb{S}^{q-1}} |x^\top A x| \le 2 \sup_{y \in \mathcal{N}} |y^\top A y|.$$

Proof. Proof of Claim A.2

Let

$$M := \sup_{y \in \mathcal{N}} |y^\top A y|.$$

Let $x^\star \in S^{q-1}$ satisfy

$$|x^{\star\top} A x^\star| = \|A\|_{\mathrm{op}}.$$

Choose $y \in \mathcal{N}$ such that

$$\|x^\star - y\| \le \frac{1}{4}.$$

Then

$$\begin{aligned}
\|A\|_{\mathrm{op}} &= |x^{\star\top} A x^\star| \\
&\le |y^\top A y| + |x^{\star\top} A x^\star - y^\top A y| \\
&= |y^\top A y| + |(x^\star - y)^\top A x^\star + y^\top A (x^\star - y)| \\
&\le M + \|x^\star - y\|\,\|A\|_{\mathrm{op}}\|x^\star\| + \|y\|\,\|A\|_{\mathrm{op}}\|x^\star - y\| \\
&\le M + \frac{1}{2}\|A\|_{\mathrm{op}},
\end{aligned}$$

where we used $\|x^\star\| = \|y\| = 1$. Rearranging gives

$$\|A\|_{\mathrm{op}} \le 2M = 2 \sup_{y \in \mathcal{N}} |y^\top A y|.$$

This proves the claim. □

Applying Claim A.2 to $A = \Sigma_d - I_q$ yields

$$\left\| \frac{1}{d} U^\top U - I_q \right\|_{op} = \|A\|_{op} \le 2 \max_{y \in \mathcal{N}} |y^\top A y|.$$

Fix $\delta \in (0,1)$, to be chosen later. Using the net bound and equation 10, we have

$$\begin{aligned}
\Pr(\|A\|_{op} \ge \delta) &\le \Pr\left(\exists\, y \in \mathcal{N} \text{ such that } |y^\top A y| \ge \frac{\delta}{2} \right) \\
&\le \sum_{y \in \mathcal{N}} \Pr\left(|y^\top A y| \ge \frac{\delta}{2} \right) \\
&\le |\mathcal{N}| \cdot 2 \exp\left(-c_2 d \frac{\delta^2}{4} \right) \\
&\le 2 \exp\left(q \log 9 - c_3 d \delta^2 \right),
\end{aligned}$$

for some absolute constant $c_3 > 0$.

Now take

$$\delta = C\Big(\sqrt{\frac{q}{d}} + t\Big), \qquad t \in (0,1),$$

for a sufficiently large absolute constant $C > 0$ (to be fixed). Then

$$\delta^2 = C^2\left(\frac{q}{d} + t^2 + 2t\sqrt{\frac{q}{d}}\right) \geq \frac{C^2}{2}\left(\frac{q}{d} + t^2\right),$$

and hence

$$\begin{aligned} q\log 9 - c_3 d\delta^2 &\leq q\log 9 - c_3 d \cdot \frac{C^2}{2}\left(\frac{q}{d} + t^2\right) \\ &= q\left(\log 9 - \frac{c_3 C^2}{2}\right) - \frac{c_3 C^2}{2} dt^2. \end{aligned}$$

Choose C large enough so that $\log 9 - \frac{c_3 C^2}{2} \leq -1$. Then

$$q\left(\log 9 - \frac{c_3 C^2}{2}\right) \leq -q \leq 0,$$

and hence

$$q\log 9 - c_3 d\delta^2 \leq -c_4 dt^2$$

for some absolute constant $c_4 > 0$. Plugging this into the previous bound gives

$$\Pr\left(\left\|\frac{1}{d}U^\top U - I_q\right\|_{op} \geq C\Big(\sqrt{\frac{q}{d}} + t\Big)\right) \leq 2\exp(-c_4 dt^2),$$

which proves the first statement (with $c := c_4$).

The in particular statement. For the second claim, take $t = \sqrt{q/d}$ (note that $t \in (0,1)$ as soon as d is sufficiently large compared to q). Then

$$C\Big(\sqrt{\frac{q}{d}} + t\Big) = 2C\sqrt{\frac{q}{d}},$$

and

$$\Pr\left(\left\|\frac{1}{d}U^\top U - I_q\right\|_{op} \geq 2C\sqrt{\frac{q}{d}}\right) \leq 2\exp\Big(-c_4 d\frac{q}{d}\Big) = 2\exp(-c_4 q).$$

Thus, with probability at least $1 - 2\exp(-c_4 q)$ we have

$$\left\|\frac{1}{d}U^\top U - I_q\right\|_{op} \leq 2C\sqrt{\frac{q}{d}},$$

which is the claimed $O(\sqrt{q/d})$ bound (with $c' := c_4$). □

Thus, in the regime $d \gg q$ the Gram matrix satisfies

$$U^\top U = d\,(I_q + \Delta_d), \qquad \|\Delta_d\|_{op} = O\left(\sqrt{\frac{q}{d}}\right)$$

with overwhelmingly high probability. Geometrically, the columns of U are almost mutually orthogonal and have almost equal norms.

A.2 Approximate Structure of the Projection Operator

We now use Lemma A.1 to approximate the projection operator P_U in the high-dimensional regime. Recall that

$$P_U = U(U^\top U)^{-1}U^\top = U\big(d(I_q + \Delta_d)\big)^{-1}U^\top = \frac{1}{d}U(I_q + \Delta_d)^{-1}U^\top.$$

Lemma A.3 (Approximation of $(U^\top U)^{-1}$)**.** *Suppose* $\|\Delta_d\|_{op} \le \frac{1}{2}$*. Then*

$$(I_q + \Delta_d)^{-1} = I_q + E_d, \qquad \|E_d\|_{op} \le 2\|\Delta_d\|_{op},$$

and hence

$$(U^\top U)^{-1} = \frac{1}{d}(I_q + E_d), \qquad \|(U^\top U)^{-1} - \tfrac{1}{d} I_q\|_{op} \le \frac{2}{d}\|\Delta_d\|_{op}.$$

Proof. If $\|\Delta_d\|_{op} \le \frac{1}{2}$ then the Neumann series

$$(I_q + \Delta_d)^{-1} = \sum_{k=0}^{\infty} (-\Delta_d)^k$$

converges in operator norm, and

$$(I_q + \Delta_d)^{-1} - I_q = \sum_{k=1}^{\infty} (-\Delta_d)^k.$$

Using the triangle inequality and submultiplicativity yields

$$\|E_d\|_{op} = \left\| \sum_{k=1}^{\infty} (-\Delta_d)^k \right\|_{op} \le \sum_{k=1}^{\infty} \|\Delta_d\|_{op}^k = \frac{\|\Delta_d\|_{op}}{1 - \|\Delta_d\|_{op}} \le 2\|\Delta_d\|_{op},$$

where the last inequality uses $\|\Delta_d\|_{op} \le \frac{1}{2}$. The bound for $(U^\top U)^{-1}$ follows by multiplication by $1/d$. □

Combining Lemmas A.1 and A.3, we obtain that for d sufficiently large and $q \ll d$, with high probability,

$$(U^\top U)^{-1} = \frac{1}{d} I_q + O\left(\frac{1}{d}\sqrt{\frac{q}{d}}\right) \quad \text{in operator norm.}$$

We next propagate this approximation through to the projector P_U.

Proposition A.4 (Approximation of P_U by $UU^\top/d$)**.** *Let* $U \in \mathbb{R}^{d\times q}$ *have i.i.d.* $\mathcal{N}(0,1)$ *entries. There exist absolute constants* $c, C > 0$ *such that the following holds. With probability at least* $1 - 2\exp(-cq)$*,*

$$\left\| P_U - \frac{1}{d} UU^\top \right\|_{op} \le C\sqrt{\frac{q}{d}}.$$

Consequently, for any $g \in \mathbb{R}^d$*,*

$$\left\| P_U g - \frac{1}{d} UU^\top g \right\| \le C\sqrt{\frac{q}{d}}\, \|g\|.$$

Proof. Write

$$P_U - \frac{1}{d} UU^\top = U\Big((U^\top U)^{-1} - \tfrac{1}{d} I_q\Big)U^\top = \frac{1}{d} U E_d U^\top,$$

where E_d is as in Lemma A.3. Thus

$$\left\| P_U - \frac{1}{d} UU^\top \right\|_{op} \le \frac{1}{d}\|U\|_{op}^2\, \|E_d\|_{op}.$$

For a Gaussian matrix $U \in \mathbb{R}^{d\times q}$, standard random matrix bounds give

$$\|U\|_{op} \le \sqrt{d} + \sqrt{q} + t$$

with probability at least $1 - 2\exp(-c' t^2)$ for some absolute constant $c' > 0$. Taking t of order $\sqrt{q}$ and using $q \le d$ yields $\|U\|_{op}^2 = O(d)$ with probability at least $1 - 2\exp(-c'' q)$ for some $c'' > 0$.

On the same high-probability event, Lemma A.1 ensures that $\|\Delta_d\|_{op} = O(\sqrt{q/d})$ and in particular $\|\Delta_d\|_{op} \le \frac{1}{2}$ for d sufficiently large. Lemma A.3 then gives

$$\|E_d\|_{op} \le 2\|\Delta_d\|_{op} = O\left(\sqrt{\frac{q}{d}}\right).$$

Putting these bounds together,

$$\left\| P_U - \frac{1}{d} UU^\top \right\|_{op} \le \frac{1}{d} \cdot C_1 d \cdot C_2 \sqrt{\frac{q}{d}} = C\sqrt{\frac{q}{d}}$$

for suitable absolute constants $C_1, C_2, C > 0$. The bound for $P_U g - (1/d)UU^\top g$ follows immediately by multiplying by $\|g\|$. □

Proposition A.4 formalizes the statement that, when $d \gg q$, the action of P_U on any vector g is very close (in norm) to the action of the simpler operator $(1/d)UU^\top$. In other words, within the sampled subspace, the update direction produced by ZO-Align is, up to a small relative error of order $\sqrt{q/d}$, the same as ZO-Avg.

A.3 Implications for ZO-Align in High Dimensions

The main advantage of ZO-Align over ZO-AVG is that it explicitly accounts for correlations among the sampled directions through the factor $(U^\top U)^{-1}$, and thus adjusts the relative contributions of different directions in U according to the spectrum of $U^\top U$. When d is moderate and q is a nontrivial fraction of d, the Gram matrix $U^\top U$ can deviate substantially from a scaled identity, and these spectral corrections can significantly improve how the sampled subspace represents the local geometry of f.

In the high-dimensional regime $d \gg q$, however, Lemma A.1 and Proposition A.4 show that, with high probability,

$$U^\top U \approx dI_q, \qquad P_U \approx \frac{1}{d} UU^\top,$$

with approximation error of order $\sqrt{q/d}$ in operator norm. In this case the sampled directions are already almost mutually orthogonal with nearly equal norms, and the spectral corrections encoded in $(U^\top U)^{-1}$ reduce, to leading order, to a global scaling by $1/d$. As a consequence, the additional structural advantage of ZO-Align over naive aggregation inside the sampled subspace becomes much less pronounced when d is extremely large and $q \ll d$.

It is important to emphasize that this high-dimensional geometric simplification does not alter the exact MSE formulas in Proposition 3.1, which are valid for every finite d. The high-dimensional analysis merely explains why the *geometric* distinction inside the sampled subspace becomes weaker as $d/q \to \infty$.

A.4 Numerical Illustration of $U^\top U \approx dI_q$

To empirically illustrate the concentration of the Gram matrix, we consider Gaussian query matrices $U \in \mathbb{R}^{d\times q}$ with a fixed $q = 10$ and varying dimension $d \in \{10, 100, 1000, 10000\}$. For each value of d, we sample U with i.i.d. $\mathcal{N}(0,1)$ entries and compute the normalized Gram matrix

$$G_d \;=\; \frac{1}{d} U^\top U \in \mathbb{R}^{q\times q}.$$

We then visualize the absolute values $|G_d|$ as heatmaps. When d is small, G_d exhibits noticeable deviations from the identity, with substantial off-diagonal entries. As d increases, the diagonal entries of G_d stabilize near 1, while the off-diagonal entries shrink toward 0, and the heatmaps become visually indistinguishable from those of an identity matrix.

A typical realization is shown in Figure 6. The progressive disappearance of off-diagonal structure as d grows is consistent with Lemma A.1 and supports the approximation $U^\top U \approx dI_q$ used in the high-dimensional discussion above.

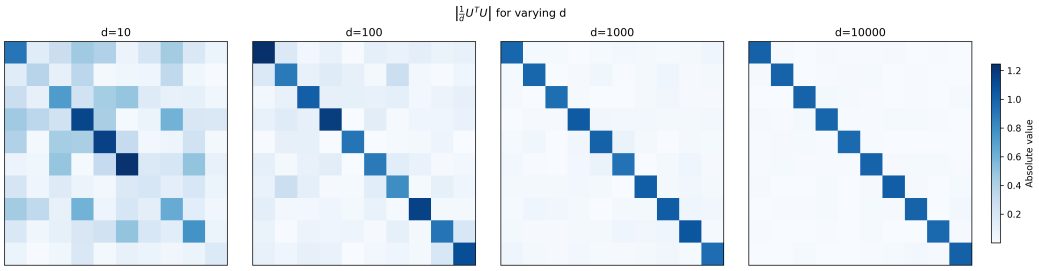

Figure 6: Heatmaps of the absolute value of the normalized Gram matrix $G_d = (1/d)U^\top U$ for $q = 10$ and varying dimensions $d \in \{10, 100, 1000, 10000\}$. Each U has i.i.d. $\mathcal{N}(0, 1)$ entries. As d increases, G_d becomes visually closer to the identity matrix, with diagonal entries near 1 and off-diagonal entries shrinking toward 0.

B Finite-Smoothing Effects

Our main theoretical analysis considers the idealized regime $\mu \to 0$, in which the finite-difference directional derivative converges to its first-order approximation. In this section, we examine the extent to which the resulting query-allocation conclusions remain predictive for finite smoothing parameters. We first quantify the finite-μ remainder, and then empirically study its interaction with the ambient dimension on both controlled quadratic problems and large-scale language-model fine-tuning.

B.1 Finite-μ approximation

For an L-smooth objective, the one-sided finite difference along a direction u satisfies

$$\frac{f(x + \mu u) - f(x)}{\mu} = u^\top \nabla f(x) + r_\mu(u), \qquad |r_\mu(u)| \le \frac{L\mu}{2}\|u\|^2. \tag{11}$$

Therefore, the finite-μ versions of both ZO-Avg and ZO-Align can be viewed as their idealized counterparts plus a remainder that vanishes linearly with μ for any fixed query direction. Consequently, for a fixed problem dimension, the $\mu \to 0$ analysis characterizes the leading-order behavior once μ is sufficiently small.

However, this approximation need not be uniform in the ambient dimension. For Gaussian perturbations $u \sim \mathcal{N}(0, I_d)$, the typical magnitude of $\|u\|^2$ grows with d. Hence, a smoothing parameter that is sufficiently small at moderate dimension can produce a non-negligible finite-difference remainder at a much larger dimension. We therefore interpret our main results as characterizing a small-smoothing regime, whose boundary may depend on the ambient dimension. The following experiments characterize this effect empirically.

B.2 Robustness at moderate dimension

We first examine whether the query-allocation behavior predicted by the idealized analysis is robust to finite μ in a moderate-dimensional setting. We consider a $d = 50$ strongly convex quadratic with condition number 10. For each configuration, we use five random seeds and a fixed budget of $K = 5000$ directional queries. Table 2 reports the median number of queries required to reach an optimality gap of 10^{-2}.

The allocation ordering is stable over a broad range of smoothing parameters. For ZO-Avg, $q = 1$ consistently requires the fewest queries, whereas ZO-Align benefits from larger query blocks. In particular, the choice $\mu = 10^{-6}$ used in our main experiments is empirically indistinguishable from the smaller $\mu = 10^{-8}$ setting in this experiment. Only when μ becomes relatively large do we begin to observe a modest degradation in query efficiency.

Table 2: Finite-smoothing ablation on a $d = 50$ strongly convex quadratic. We report the median number of queries required to reach an optimality gap of 10^{-2} over five random seeds. Lower is better.

Estimator	q	$\mu = 10^{-8}$	$\mu = 10^{-6}$	$\mu = 10^{-4}$	$\mu = 10^{-3}$
ZO-Avg	1	545	545	545	560
ZO-Avg	5	565	565	565	570
ZO-Avg	25	800	800	800	825
ZO-Align	1	545	545	545	560
ZO-Align	5	530	530	530	535
ZO-Align	25	475	475	500	525

Table 3: Interaction between the smoothing parameter μ and ambient dimension d for ZO-Avg. Entries are normalized relative gaps after a fixed budget of $K = 10{,}000$ queries. Lower is better, and the best query block size for each (d, μ) pair is shown in bold.

d	μ	$q = 1$	$q = 10$	$q = 100$
10^3	10^{-2}	0.025100	0.024879	**0.022870**
	10^{-4}	2.510×10^{-6}	4.589×10^{-6}	2.943×10^{-5}
	10^{-6}	2.510×10^{-10}	2.101×10^{-6}	2.714×10^{-5}
	10^{-8}	**0**	2.101×10^{-6}	2.714×10^{-5}
10^5	10^{-2}	2.500100	2.499811	**2.496923**
	10^{-4}	2.500×10^{-4}	2.928×10^{-4}	7.200×10^{-4}
	10^{-6}	2.500×10^{-8}	4.281×10^{-5}	4.704×10^{-4}
	10^{-8}	**0**	4.278×10^{-5}	4.704×10^{-4}
10^7	10^{-2}	250.000100	249.999763	**249.996393**
	10^{-4}	**0.025000010**	0.025000426	0.025004586
	10^{-6}	2.500×10^{-6}	2.950×10^{-6}	7.447×10^{-6}
	10^{-8}	**0**	4.498×10^{-7}	4.947×10^{-6}
10^9	10^{-2}	25000.000725	25000.000387	**24999.997013**
	10^{-4}	2.500000048	2.500000018	**2.499999726**
	10^{-6}	2.499750×10^{-4}	2.499795×10^{-4}	2.500245×10^{-4}
	10^{-8}	**0**	4.500×10^{-9}	4.950×10^{-8}

B.3 Interaction between smoothing and dimension

We next study whether the finite-μ effect becomes more pronounced as the ambient dimension increases. We consider a controlled quadratic objective with dimensions

$$d \in \{10^3, 10^5, 10^7, 10^9\},$$

using the same initialization for all configurations at a given dimension and a fixed budget of $K = 10{,}000$ queries. We vary

$$\mu \in \{10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}\}, \qquad q \in \{1, 10, 100\}.$$

To make the results comparable across dimensions, we report the normalized relative gap

$$\mathrm{RelGap}(d, \mu, q) = \frac{f_{d,\mu,q} - \hat{f}_d^\star}{f(x_0) - \hat{f}_d^\star}, \tag{12}$$

where $f_{d,\mu,q}$ denotes the final objective value and $\hat{f}_d^\star$ denotes the reference optimum used for the corresponding dimension. We report ZO-Avg in this sweep, since in the high-dimensional regime the experiments in the main paper show that ZO-Align behaves similarly to a rescaled ZO-Avg update.

Table 3 reveals a clear interaction between dimension and smoothing. When $\mu \le 10^{-6}$, the theoretically predicted single-query allocation remains optimal across all tested dimensions up to $d = 10^9$. For $\mu = 10^{-4}$, $q = 1$ remains optimal through $d = 10^7$, while the minimizer switches to

Table 4: Finite-smoothing ablation for Qwen3-0.6B on SST2. We report the final optimization loss under the same fixed-budget protocol as the high-dimensional experiment in the main text. Lower is better.

q	$\mu = 10^{-2}$	$\mu = 10^{-4}$	$\mu = 10^{-6}$	$\mu = 10^{-8}$
1	1.059193	**0.157867**	**0.156994**	**0.426127**
10	1.159638	0.463681	0.461286	0.466158
100	1.339527	0.564389	0.563375	0.563426

$q = 100$ at $d = 10^9$. With the much larger $\mu = 10^{-2}$, the finite-difference error dominates and the optimization remains far from the reference optimum.

These observations show that the $\mu \to 0$ allocation result is predictive in a sufficiently small-μ regime, but that the required smoothing scale becomes more stringent as the dimension increases. In particular, the limit $\mu \to 0$ should not be interpreted as a dimension-uniform approximation.

B.4 Finite-μ ablation for language-model fine-tuning

Finally, we test whether the preceding transition is relevant at practical smoothing scales in our high-dimensional language-model experiment. We repeat the Qwen3-0.6B experiment using the same fixed query budget and query-block choices as in the main text, while varying $\mu \in \{10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}\}$. Table 4 reports the final loss for ZO-Avg.

For the practically useful range $\mu \in [10^{-6}, 10^{-4}]$, changing the smoothing parameter has little effect on the final loss, and $q = 1$ remains clearly preferred. When $\mu = 10^{-2}$, the finite-difference approximation is no longer sufficiently accurate and performance degrades substantially. At the other extreme, $\mu = 10^{-8}$ degrades the $q = 1$ result, whereas the $q = 10$ and $q = 100$ results remain close to those obtained with $\mu = 10^{-6}$. This behavior is consistent with floating-point cancellation becoming relevant at an excessively small smoothing scale, with averaging partially mitigating its effect.

Taken together, these experiments clarify the scope of our idealized analysis. Finite smoothing can alter the preferred query allocation when μ is too large relative to the ambient dimension, so the $\mu \to 0$ theory should be viewed as characterizing a small-smoothing regime rather than as a dimension-uniform statement. At the same time, both the moderate-dimensional experiment and the Qwen3-0.6B ablation show that the allocation behavior is stable over the smoothing range used in our main experiments.

C Proofs

C.1 Preliminaries

Lemma C.1. *Fix an iteration t. Let $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot \mid x_t]$, where the conditional expectation is taken over the random query directions in $U_t = [u_{t,1}, \ldots, u_{t,q_t}] \in \mathbb{R}^{d \times q_t}$, whose columns are sampled independently from $\mathcal{N}(0, I_d)$. In the idealized regime $\mu \to 0$, the following hold.*

***(ZO-Avg)** If*

$$\hat{g}_{AVG}(x_t) = \frac{1}{q_t} U_t U_t^\top \nabla f(x_t),$$

then

$$\mathbb{E}_t[\hat{g}_{AVG}(x_t)] = \nabla f(x_t),$$

and

$$\mathbb{E}_t\big[\|\hat{g}_{AVG}(x_t)\|^2\big] = \frac{q_t + d + 1}{q_t} \|\nabla f(x_t)\|^2.$$

***(ZO-Align)** If*

$$\hat{g}_{ALG}(x_t) = P_{U_t} \nabla f(x_t), \qquad P_{U_t} := U_t (U_t^\top U_t)^{-1} U_t^\top,$$

then

$$\mathbb{E}_t[\hat{g}_{ALG}(x_t)] = \frac{q_t}{d}\nabla f(x_t),$$

and

$$\mathbb{E}_t\left[\|\hat{g}_{ALG}(x_t)\|^2\right] = \frac{q_t}{d}\|\nabla f(x_t)\|^2.$$

Proof. Throughout the proof, condition on x_t. Then $\nabla f(x_t)$ is deterministic, and all expectations are taken only over the fresh random query matrix U_t.

For brevity, let $g_t := \nabla f(x_t)$. We first prove the statements for ZO-Avg. In the idealized regime $\mu \to 0$, we have

$$\hat{g}_{AVG}(x_t) = \frac{1}{q_t}U_t U_t^\top g_t = \frac{1}{q_t}\sum_{i=1}^{q_t} u_{t,i}u_{t,i}^\top g_t.$$

Taking conditional expectation and using $\mathbb{E}_t[u_{t,i}u_{t,i}^\top] = I_d$, we obtain

$$\mathbb{E}_t[\hat{g}_{AVG}(x_t)] = \frac{1}{q_t}\sum_{i=1}^{q_t}\mathbb{E}_t[u_{t,i}u_{t,i}^\top]g_t = \frac{1}{q_t}\sum_{i=1}^{q_t} I_d g_t = g_t.$$

Hence $\mathbb{E}_t[\hat{g}_{AVG}(x_t)] = \nabla f(x_t)$.

Next we compute the conditional second moment. By Proposition 3.1 applied conditionally on x_t, the mean squared error of $\hat{g}_{AVG}(x_t)$ is $\mathbb{E}_t\left[\|\hat{g}_{AVG}(x_t) - g_t\|^2\right] = \frac{d+1}{q_t}\|g_t\|^2$. Since $\mathbb{E}_t[\hat{g}_{AVG}(x_t)] = g_t$, the bias is zero, and therefore

$$\mathbb{E}_t\left[\|\hat{g}_{AVG}(x_t)\|^2\right] = \mathbb{E}_t\left[\|\hat{g}_{AVG}(x_t) - g_t\|^2\right] + \|g_t\|^2 = \frac{q_t + d + 1}{q_t}\|g_t\|^2 + \|g_t\|^2.$$

We now turn to ZO-Align. In the idealized regime, $\hat{g}_{ALG}(x_t) = P_{U_t}g_t$. Taking conditional expectation gives $\mathbb{E}_t[\hat{g}_{ALG}(x_t)] = \mathbb{E}_t[P_{U_t}]\, g_t$. By rotational symmetry, the random subspace $\mathrm{col}(U_t)$ is uniformly distributed over the Grassmannian of q_t-dimensional subspaces of $\mathbb{R}^d$, so $\mathbb{E}_t[P_{U_t}]$ must be a scalar multiple of the identity: $\mathbb{E}_t[P_{U_t}] = cI_d$ for some scalar c.

Taking traces on both sides yields $cd = \mathrm{tr}(\mathbb{E}_t[P_{U_t}]) = \mathbb{E}_t[\mathrm{tr}(P_{U_t})]$. Since P_{U_t} is the orthogonal projector onto a q_t-dimensional subspace, $\mathrm{tr}(P_{U_t}) = q_t$. Hence $cd = q_t$, so $c = \frac{q_t}{d}$. Therefore $\mathbb{E}_t[P_{U_t}] = \frac{q_t}{d}I_d$, and consequently

$$\mathbb{E}_t[\hat{g}_{ALG}(x_t)] = \frac{q_t}{d}g_t = \frac{q_t}{d}\nabla f(x_t).$$

Finally, for the conditional second moment, using that P_{U_t} is symmetric and idempotent,

$$P_{U_t}^\top = P_{U_t}, \qquad P_{U_t}^2 = P_{U_t},$$

we obtain

$$\|\hat{g}_{ALG}(x_t)\|^2 = \|P_{U_t}g_t\|^2 = g_t^\top P_{U_t}^\top P_{U_t}g_t = g_t^\top P_{U_t}g_t.$$

Taking conditional expectation,

$$\mathbb{E}_t\left[\|\hat{g}_{ALG}(x_t)\|^2\right] = g_t^\top\mathbb{E}_t[P_{U_t}]g_t = g_t^\top\left(\frac{q_t}{d}I_d\right)g_t = \frac{q_t}{d}\|g_t\|^2.$$

This completes the proof. □

C.2 Proof of Proposition 3.1

Proof. **Analysis of ZO-Avg.** Since ZO-Avg is unbiased, its MSE is the trace of its covariance matrix. Let $g = \nabla f(x)$ and $\hat{g}_i = (g^T u_i)u_i$. The ZO-Avg estimator is $\hat{g}_{AVG} = \frac{1}{q}\sum_{i=1}^q \hat{g}_i$. As the directions $u_i \sim \mathcal{N}(0, I)$ are i.i.d., the terms $\hat{g}_i$ are also i.i.d. random vectors. The variance of the average is thus:

$$\mathrm{Var}(\hat{g}_{AVG}) = \frac{1}{q^2}\sum_{i=1}^{q}\mathrm{Var}(\hat{g}_i) = \frac{1}{q}\mathrm{Var}(\hat{g}_1).$$

To compute $\text{Var}(\hat{g}_1)$, we need the first two moments of $\hat{g}_1$. We have $\mathbb{E}[\hat{g}_1] = \mathbb{E}[u_1 u_1^T]g = Ig = g$. For the second moment, we use the known result for a standard Gaussian vector $u \sim \mathcal{N}(0, I)$ and a fixed vector v: $\mathbb{E}[(v^T u)^2 uu^T] = 2vv^T + \|v\|^2 I$. Thus:

$$\mathbb{E}[\hat{g}_1 \hat{g}_1^T] = \mathbb{E}[(g^T u_1)^2 u_1 u_1^T] = 2gg^T + \|g\|^2 I.$$

The covariance matrix of a single estimator is:

$$\begin{aligned}\text{Var}(\hat{g}_1) &= \mathbb{E}[\hat{g}_1 \hat{g}_1^T] - \mathbb{E}[\hat{g}_1]\mathbb{E}[\hat{g}_1]^T \\ &= (2gg^T + \|g\|^2 I) - gg^T = gg^T + \|g\|^2 I.\end{aligned}$$

The covariance of the ZO-Avg estimator is $\text{Var}(\hat{g}_{AVG}) = \frac{1}{q}(gg^T + \|g\|^2 I)$. The MSE is the trace of this matrix:

$$\begin{aligned}\text{MSE}(\hat{g}_{AVG}) &= \text{tr}(\text{Var}(\hat{g}_{AVG})) = \frac{1}{q}\text{tr}(gg^T + \|g\|^2 I) \\ &= \frac{1}{q}(\text{tr}(gg^T) + \|g\|^2 \text{tr}(I)) = \frac{1}{q}(\|g\|^2 + d\|g\|^2) = \frac{d+1}{q}\|g\|^2.\end{aligned}$$

Analysis of ZO-Align. The ZO-Align estimator, $\hat{g}_{ALGN} = P_U g$, is biased, with $\mathbb{E}[\hat{g}_{ALG}] = \frac{q}{d} g$. Here, we compute the MSE directly. Let $\hat{g}_U = P_U g$.

$$\begin{aligned}\text{MSE}(\hat{g}_{ALG}) &= \mathbb{E}[\|\hat{g}_U - g\|^2] = \mathbb{E}[\|(P_U - I)g\|^2] \\ &= \mathbb{E}[g^T (P_U - I)^T (P_U - I) g].\end{aligned}$$

Since P_U is a projection matrix, it is symmetric ($P_U^T = P_U$) and idempotent ($P_U^2 = P_U$). Therefore, $(P_U - I)^2 = P_U^2 - 2P_U + I = P_U - 2P_U + I = I - P_U$.

$$\begin{aligned}\text{MSE}(\hat{g}_{ALG}) &= \mathbb{E}[g^T (I - P_U) g] = g^T \mathbb{E}[I - P_U] g \\ &= g^T (I - \mathbb{E}[P_U]) g.\end{aligned}$$

Using the known result that the expectation of a projection onto a random q-dimensional subspace is $\mathbb{E}[P_U] = \frac{q}{d} I$, we have:

$$\text{MSE}(\hat{g}_{ALG}) = g^T (I - \frac{q}{d} I) g = \left(1 - \frac{q}{d}\right) \|g\|^2 = \frac{d-q}{d}\|g\|^2.$$

□

C.3 Proof of Proposition 4.1

Proof. Fix an iteration t, and define $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot \mid x_t]$. Throughout the proof, the conditional expectation is taken over the fresh random query directions used at iteration t, while x_t (and hence $\nabla f(x_t)$) is fixed. From the L-smoothness of f, we have the descent lemma

$$f(x_{t+1}) \le f(x_t) + \langle \nabla f(x_t), x_{t+1} - x_t \rangle + \frac{L}{2}\|x_{t+1} - x_t\|^2.$$

ZO-Avg. Substituting the update rule $x_{t+1} - x_t = -\eta_t \hat{g}_{AVG}(x_t)$ into the descent lemma and taking $\mathbb{E}_t[\cdot]$, we obtain

$$\mathbb{E}_t[f(x_{t+1})] \le f(x_t) - \eta_t \big\langle \nabla f(x_t), \mathbb{E}_t[\hat{g}_{AVG}(x_t)] \big\rangle + \frac{L\eta_t^2}{2} \mathbb{E}_t\big[\|\hat{g}_{AVG}(x_t)\|^2\big].$$

Using lemma C.1

$$\mathbb{E}_t[\hat{g}_{AVG}(x_t)] = \nabla f(x_t), \qquad \mathbb{E}_t\big[\|\hat{g}_{AVG}(x_t)\|^2\big] = \frac{q_t + d + 1}{q_t}\|\nabla f(x_t)\|^2,$$

we get

$$\mathbb{E}_t[f(x_{t+1})] \le f(x_t) - \left(\eta_t - \frac{L\eta_t^2}{2} \cdot \frac{q_t + d + 1}{q_t}\right) \|\nabla f(x_t)\|^2.$$

Therefore, the guaranteed decrease is maximized by maximizing

$$\phi_{AVG}(\eta_t) := \eta_t - \frac{L\eta_t^2}{2} \cdot \frac{q_t + d + 1}{q_t},$$

which is a concave quadratic function of η_t. Its derivative is

$$\phi'_{AVG}(\eta_t) = 1 - L\eta_t \frac{q_t + d + 1}{q_t}.$$

Setting $\phi'_{AVG}(\eta_t) = 0$ yields

$$\eta_t^\star = \frac{q_t}{L(q_t + d + 1)}.$$

ZO-Align. Similarly, substituting $x_{t+1} - x_t = -\eta_t \hat{g}_{ALG}(x_t)$ into the descent lemma and taking $\mathbb{E}_t[\cdot]$, we obtain

$$\mathbb{E}_t[f(x_{t+1})] \le f(x_t) - \eta_t \big\langle \nabla f(x_t), \mathbb{E}_t[\hat{g}_{ALG}(x_t)] \big\rangle + \frac{L\eta_t^2}{2} \mathbb{E}_t\big[\|\hat{g}_{ALG}(x_t)\|^2\big].$$

Using lemma C.1

$$\mathbb{E}_t[\hat{g}_{ALG}(x_t)] = \frac{q_t}{d} \nabla f(x_t), \qquad \mathbb{E}_t\big[\|\hat{g}_{ALG}(x_t)\|^2\big] = \frac{q_t}{d} \|\nabla f(x_t)\|^2,$$

we obtain

$$\mathbb{E}_t[f(x_{t+1})] \le f(x_t) - \frac{q_t}{d} \left(\eta_t - \frac{L\eta_t^2}{2} \right) \|\nabla f(x_t)\|^2.$$

Hence, the guaranteed decrease is maximized by maximizing

$$\phi_{ALG}(\eta_t) := \eta_t - \frac{L\eta_t^2}{2},$$

which is again a concave quadratic function. Since

$$\phi'_{ALG}(\eta_t) = 1 - L\eta_t,$$

the maximizer is

$$\eta_t^\star = \frac{1}{L}.$$

This proves the two claims. □

C.4 Proof of Theorem 4.2

Proof. The first part of the proof establishes the one-step contraction factor. From the L-smoothness of f, we have the descent lemma:

$$f(x_{t+1}) \le f(x_t) + \langle \nabla f(x_t), x_{t+1} - x_t \rangle + \frac{L}{2} \|x_{t+1} - x_t\|^2$$

Substituting the update rule $x_{t+1} - x_t = -\eta_t \hat{g}_{AVG}(x_t)$ and taking the expectation conditional on x_t:

$$\mathbb{E}[f(x_{t+1})|x_t] \le f(x_t) - \eta_t \langle \nabla f(x_t), \mathbb{E}[\hat{g}_{AVG}(x_t)] \rangle + \frac{L\eta_t^2}{2} \mathbb{E}[\|\hat{g}_{AVG}(x_t)\|^2]$$

Using the properties $\mathbb{E}[\hat{g}_{AVG}(x_t)] = \nabla f(x_t)$ and $\mathbb{E}[\|\hat{g}_{AVG}(x_t)\|^2] = \frac{q_t + d + 1}{q_t} \|\nabla f(x_t)\|^2$, we get:

$$\mathbb{E}[f(x_{t+1})|x_t] \le f(x_t) - \eta_t \|\nabla f(x_t)\|^2 + \frac{L\eta_t^2 (q_t + d + 1)}{2q_t} \|\nabla f(x_t)\|^2.$$

It is easy to see that the optimal step size $\eta_t^* = \frac{q_t}{L(q_t + d + 1)}$. Plugging in the optimal step size, we obtain the one-step progress:

$$\mathbb{E}[f(x_{t+1})|x_t] \le f(x_t) - \frac{q_t}{2L(q_t + d + 1)} \|\nabla f(x_t)\|^2$$

By γ-strong convexity, $\|\nabla f(x_t)\|^2 \geq 2\gamma(f(x_t) - f(x_*))$. Let $\delta_t = E[f(x_t) - f(x_*)]$. Taking the full expectation and applying the convexity condition gives:

$$\delta_{t+1} \leq \left(1 - \frac{\gamma q_t}{L(q_t + d + 1)}\right) \delta_t$$

Unrolling this recurrence from $t = T - 1$ down to 0:

$$\begin{aligned}\delta_T &\leq \left(1 - \frac{\gamma q_{T-1}}{L(q_{T-1} + d + 1)}\right) \delta_{T-1} \\ &\leq \left(1 - \frac{\gamma q_{T-1}}{L(q_{T-1} + d + 1)}\right)\left(1 - \frac{\gamma q_{T-2}}{L(q_{T-2} + d + 1)}\right) \delta_{T-2} \\ &\cdots \leq \left(\prod_{t=0}^{T-1}\left(1 - \frac{\gamma q_t}{L(q_t + d + 1)}\right)\right) \delta_0.\end{aligned}$$

This establishes the first part of the theorem.

For the second part, we must find the sequence $\{q_t\}_{t=0}^{T-1}$ that minimizes the final error bound under the budget constraint. This is equivalent to minimizing the total contraction factor. Let $\rho_t = \frac{\gamma q_t}{L(q_t+d+1)}$. We solve the following optimization problem:

$$\min_{\{q_t\},T} \prod_{t=0}^{T-1}(1 - \rho_t) \quad \text{s.t.} \sum_{t=0}^{T-1} q_t \leq K, \quad 1 \leq q_t \leq d, \quad q_t \in \mathbb{Z}.$$

Since $0 < 1 - \rho_t < 1$ for every $q_t \geq 1$, any unused budget can only make the product larger, so the optimum must satisfy $\sum_{t=0}^{T-1} q_t = K$. Now define

$$\phi(q) := -\log(1-\rho_t) = -\log\left(1 - \frac{\gamma q}{L(q + d + 1)}\right) = \log(q+d+1) - \log\left(d + 1 + \left(1 - \frac{\gamma}{L}\right) q\right).$$

Then minimizing $\prod_{t=0}^{T-1}(1 - \rho_t)$ is equivalent to maximizing $\sum_{t=0}^{T-1} \phi(q_t)$. Moreover,

$$\phi''(q) = -\frac{1}{(q + d + 1)^2} + \frac{\left(1 - \frac{\gamma}{L}\right)^2}{\left(d + 1 + \left(1 - \frac{\gamma}{L}\right) q\right)^2} < 0,$$

since $\left(1 - \frac{\gamma}{L}\right)(q + d + 1) < d + 1 + \left(1 - \frac{\gamma}{L}\right) q$. Hence ϕ is strictly concave on $q \geq 0$, and $\phi(0) = 0$. Therefore $\phi(q)/q$ is nonincreasing on $q > 0$, which implies $\phi(q) \leq q\,\phi(1), \forall q \geq 1$. Thus

$$\sum_{t=0}^{T-1} \phi(q_t) \leq \phi(1) \sum_{t=0}^{T-1} q_t = K\phi(1),$$

with equality iff $q_t = 1$ for all t. Therefore the optimal allocation is $q_t = 1$ for all t. □

C.5 Proof of Theorem 4.3

Proof. As in C.4, we establish the one-step contraction factor. Using the properties $\mathbb{E}[\hat{g}_{ALG}(x_t)] = \frac{q_t}{d}\nabla f(x_t)$ and $\mathbb{E}[\|\hat{g}_{ALG}(x_t)\|^2] = \frac{q_t}{d}\|\nabla f(x_t)\|^2$, we have:

$$\mathbb{E}[f(x_{t+1})|x_t] \leq f(x_t) - \frac{q_t}{d}\eta_t\|\nabla f(x_t)\|^2 + \frac{L\eta_t^2 q_t}{2d}\|\nabla f(x_t)\|^2.$$

It is easy to see that the optimal step size $\eta_t^* = 1/L$. Plugging in the optimal step size, we have:

$$\mathbb{E}[f(x_{t+1})|x_t] \leq f(x_t) - \frac{q_t}{2Ld}\|\nabla f(x_t)\|^2$$

Applying the γ-strong convexity condition and taking the full expectation yields:

$$\mathbb{E}[f(x_{t+1}) - f(x_*)] \leq \left(1 - \frac{\gamma q_t}{Ld}\right)(E[f(x_t)] - f(x_*))$$

Unrolling this recurrence, just as in the proof of Theorem 4.2, gives the first result.

For the second part, we must find the allocation $\{q_t\}$ that minimizes the final error bound. This is equivalent to solving the optimization problem:

$$\min_{\{q_t\},T} \prod_{t=0}^{T-1} \left(1 - \frac{\gamma q_t}{Ld}\right) \quad \text{s.t.} \sum_{t=0}^{T-1} q_t \leq K, \quad 0 \leq q_t \leq d, \quad q_t \in \mathbb{Z}$$

Let $c = \gamma/(Ld)$. The objective function $F(q_0, \ldots, q_{T-1}) = \prod_{t=0}^{T-1}(1 - cq_t)$ is a product of linear terms. Without loss of generality, it is equivalent to minimize its logarithm, $ln(F(q))$. The logarithm function is concave over the feasible region (the simplex defined by the budget constraint). A fundamental result in convex optimization is that the minimum of a concave function over a compact convex set is achieved at an extreme point (a vertex) of the set.

The vertices of the allocation simplex correspond to solutions where the budget is concentrated as much as possible. Given the constraints, the extreme points are strategies that use the largest possible query count, $q_t = d$, for some number of iterations, and $q_t = 0$ for others. To satisfy the budget $\sum q_t = K$, the optimal strategy is therefore to perform $T = \lfloor K/d \rfloor$ iterations with $q_t = d$ for each iteration plus one residual iteration. This concentrates the entire query budget into the fewest, most powerful steps, thus minimizing the product of contraction factors. □

C.6 Proof of Theorem 4.4

Proof. We analyze the evolution of the squared distance to the optimum, $\|x_t - x_*\|^2$.

$$\|x_{t+1} - x_*\|^2 = \|x_t - x_*\|^2 - 2\eta_t \hat{g}_{AVG}(x_t)^T(x_t - x_*) + \eta_t^2 \|\hat{g}_{AVG}(x_t)\|^2$$

Taking the expectation conditional on x_t and using the properties of the ZO-Avg estimator, we get:

$$E[\|x_{t+1} - x_*\|^2 | x_t] = \|x_t - x_*\|^2 - 2\eta_t \nabla f(x_t)^T(x_t - x_*) + \eta_t^2 \frac{q_t + d + 1}{q_t} \|\nabla f(x_t)\|^2$$

By convexity, we have $\nabla f(x_t)^T(x_t - x_*) \geq f(x_t) - f(x_*)$. Substituting the step size $\eta_t = \frac{q_t}{L(q_t+d+1)}$ yields:

$$E[\|x_{t+1} - x_*\|^2 | x_t] \leq \|x_t - x_*\|^2 - \frac{2q_t}{L(q_t + d + 1)}(f(x_t) - f(x_*)) + \frac{q_t}{L^2(q_t + d + 1)} \|\nabla f(x_t)\|^2$$

From the one-step analysis in Theorem 4.2, we know that $E[f(x_t) - f(x_{t+1})|x_t] \geq \frac{q_t}{2L(q_t+d+1)} \|\nabla f(x_t)\|^2$. This implies $\frac{q_t}{L(q_t+d+1)} \|\nabla f(x_t)\|^2 \leq 2(f(x_t) - E[f(x_{t+1})|x_t])$. Substituting this into the inequality gives:

$$E[\|x_{t+1} - x_*\|^2 | x_t] \leq \|x_t - x_*\|^2 - \frac{2q_t}{L(q_t + d + 1)}(f(x_t) - f(x_*)) + \frac{2}{L}(f(x_t) - E[f(x_{t+1})|x_t])$$

Rearranging and taking the full expectation, we let $\delta_t = E[f(x_t) - f(x_*)]$ and $D_t = E[\|x_t - x_*\|^2]$:

$$\frac{2q_t}{L(q_t + d + 1)} \delta_t \leq D_t - D_{t+1} + \frac{2}{L}(\delta_t - \delta_{t+1})$$

Summing from $t = 0$ to $T - 1$:

$$\sum_{t=0}^{T-1} \frac{2q_t}{L(q_t + d + 1)} \delta_t \leq D_0 - D_T + \frac{2}{L}(\delta_0 - \delta_T) \leq \|x_0 - x_*\|^2 + \frac{2}{L}(f(x_0) - f(x_*))$$

Since δ_t is a non-increasing sequence, $\delta_t \geq \delta_T$. Therefore:

$$\delta_T \left(\sum_{t=0}^{T-1} \frac{2q_t}{L(q_t + d + 1)} \right) \leq \|x_0 - x_*\|^2 + \frac{2}{L}(f(x_0) - f(x_*))$$

Rearranging gives the first result. To obtain the tightest bound on δ_T, we must maximize the sum in the denominator, $\sum_{t=0}^{T-1} \frac{q_t}{q_t+d+1}$, subject to $\sum q_t \leq K$. As established in the proof of Theorem 4.2,

the function $h(q) = q/(q + d + 1)$ is concave, and the value per query, $h(q)/q$, is maximized at $q = 1$. Thus, the optimal strategy is to set $T = K$ and $q_t = 1$ for all t. This yields the second result:

$$E[f(x_K) - f(x_*)] \leq \frac{L(\|x_0 - x_*\|^2 + \frac{2}{L}(f(x_0) - f(x_*)))}{\sum_{t=0}^{K-1} \frac{2}{d+2}}$$
$$= \frac{L(d+2)(\|x_0 - x_*\|^2 + \frac{2}{L}(f(x_0) - f(x_*)))}{2K}$$

This completes the proof. □

C.7 Proof of Theorem 4.5

Proof. Following the same initial steps as in the proof of Theorem 4.4, we start with the squared distance:

$$E[\|x_{t+1} - x_*\|^2 | x_t] = \|x_t - x_*\|^2 - 2\eta_t E[\hat{g}_{ALG}(x_t)]^T (x_t - x_*) + \eta_t^2 E[\|\hat{g}_{ALG}(x_t)\|^2]$$

Using the properties of ZO-Align ($E[\hat{g}_{ALG}] = \frac{q_t}{d}\nabla f(x_t)$, $E[\|\hat{g}_{ALG}\|^2] = \frac{q_t}{d}\|\nabla f(x_t)\|^2$), step size $\eta_t = 1/L$, and the convexity property $\nabla f(x_t)^T (x_t - x_*) \geq f(x_t) - f(x_*)$:

$$E[\|x_{t+1} - x_*\|^2 | x_t] \leq \|x_t - x_*\|^2 - \frac{2q_t}{Ld}(f(x_t) - f(x_*)) + \frac{q_t}{L^2 d}\|\nabla f(x_t)\|^2$$

From the one-step analysis in Theorem 4.3, we have $E[f(x_t) - f(x_{t+1})|x_t] \geq \frac{q_t}{2Ld}\|\nabla f(x_t)\|^2$, which implies $\frac{q_t}{Ld}\|\nabla f(x_t)\|^2 \leq 2(f(x_t) - E[f(x_{t+1})|x_t])$. Substituting this gives:

$$E[\|x_{t+1} - x_*\|^2 | x_t] \leq \|x_t - x_*\|^2 - \frac{2q_t}{Ld}(f(x_t) - f(x_*)) + \frac{2}{L}(f(x_t) - E[f(x_{t+1})|x_t])$$

Rearranging, taking the full expectation, and summing from $t = 0$ to $T - 1$:

$$\sum_{t=0}^{T-1} \frac{2q_t}{Ld} E[f(x_t) - f(x_*)] \leq \|x_0 - x_*\|^2 + \frac{2}{L}(f(x_0) - f(x_*))$$

By convexity of f, the sequence $E[f(x_t) - f(x_*)]$ is non-increasing. Thus, we can lower bound the left side using the final error, $E[f(x_T) - f(x_*)]$:

$$E[f(x_T) - f(x_*)] \left(\sum_{t=0}^{T-1} \frac{2q_t}{Ld} \right) \leq \|x_0 - x_*\|^2 + \frac{2}{L}(f(x_0) - f(x_*))$$

Rearranging gives the final bound. □

C.8 Proof of Theorem 4.6

Proof. From the L-smoothness of f, we have the one-step progress from the proof of Theorem 4.2:

$$E[f(x_{t+1})|x_t] \leq f(x_t) - \frac{q_t}{2L(q_t + d + 1)}\|\nabla f(x_t)\|^2$$

Taking the full expectation and rearranging gives:

$$\frac{q_t}{2L(q_t + d + 1)} \mathbb{E}[\|\nabla f(x_t)\|^2] \leq \mathbb{E}[f(x_t)] - \mathbb{E}[f(x_{t+1})]$$

Summing from $t = 0$ to $T - 1$:

$$\sum_{t=0}^{T-1} \frac{q_t}{2L(q_t + d + 1)} \mathbb{E}[\|\nabla f(x_t)\|^2] \leq \sum_{t=0}^{T-1} (\mathbb{E}[f(x_t)] - \mathbb{E}[f(x_{t+1})])$$

The right-hand side is a telescoping sum, which simplifies to $\mathbb{E}[f(x_0)] - \mathbb{E}[f(x_T)] \leq f(x_0) - f^*$. For the left-hand side, we have:

$$\left(\min_{t=0,\ldots,T-1} \mathbb{E}[\|\nabla f(x_t)\|^2] \right) \left(\sum_{t=0}^{T-1} \frac{q_t}{2L(q_t + d + 1)} \right) \leq \sum_{t=0}^{T-1} \frac{q_t}{2L(q_t + d + 1)} \mathbb{E}[\|\nabla f(x_t)\|^2]$$

Combining these gives the first result. To optimize the bound under a fixed budget K, we must maximize the denominator $\sum_{t=0}^{T-1} \frac{q_t}{q_t + d + 1}$. As established in Theorem 4.2, this sum is maximized by choosing $q_t = 1$ for $T = K$ iterations. Substituting this into the bound yields the final rate. □

C.9 Proof of Theorem 4.7

Proof. The proof follows the same structure as for Theorem 4.6. From the one-step analysis in Theorem 4.3, we have:

$$\mathbb{E}[f(x_{t+1})|x_t] \leq f(x_t) - \frac{q_t}{2Ld}\|\nabla f(x_t)\|^2$$

Taking the full expectation, rearranging, and summing from $t = 0$ to $T - 1$:

$$\sum_{t=0}^{T-1} q_t \mathbb{E}[\|\nabla f(x_t)\|^2] \leq 2Ld \sum_{t=0}^{T-1} (\mathbb{E}[f(x_t)] - \mathbb{E}[f(x_{t+1})]) \leq 2Ld(f(x_0) - f^*)$$

Dividing by $\sum q_t$ and using the fact that $\min_t \mathbb{E}[\|\nabla f(x_t)\|^2] \leq \frac{1}{\sum_{t=0}^{T-1} q_t} \sum q_t \mathbb{E}[\|\nabla f(x_t)\|^2]$ gives the result. The rate depends directly on the total number of queries, $\sum q_t$, which is maximized by using the full budget K. □

D Stochastic Case

We now extend the analysis to stochastic convex optimization, where $f(x) = \mathbb{E}_\xi[F(x, \xi)]$. Unlike the deterministic setting, stochastic gradient noise introduces an additional trade-off between rapid reduction of the initialization error and variance accumulation. Consequently, the preferred query allocation becomes regime-dependent. We derive representative guarantees for ZO-Avg and ZO-Align under diminishing step sizes, and then examine when the initialization-dominated regime underlying the ZO-Align allocation result is relevant in practice.

For this analysis, we adopt standard assumptions: $f(x)$ is convex, each component function $F(x, \xi)$ is convex and L-smooth, and the stochastic gradients have bounded variance at the optimum, $x^* = \arg\min_x f(x)$.

Assumption D.1. The variance of the stochastic gradient at the optimum is bounded, i.e.,

$$\mathbb{E}_\xi[\|\nabla F(x^*, \xi) - \nabla f(x^*)\|^2] = \mathbb{E}_\xi[\|\nabla F(x^*, \xi)\|^2] \leq \sigma^2.$$

This leads to the following standard lemma which bounds the expected squared norm of the stochastic gradient.

Lemma D.2. *For a convex and L-smooth function $f(x)$ satisfying Assumption D.1, we have* $\mathbb{E}_\xi[\|\nabla F(x, \xi)\|^2] \leq 4L(f(x) - f(x^*)) + 2\sigma^2$.

Proof. By the triangle inequality and the property $\|a + b\|^2 \leq 2\|a\|^2 + 2\|b\|^2$, we have:

$$\mathbb{E}_\xi[\|\nabla F(x, \xi)\|^2] \leq 2\mathbb{E}_\xi[\|\nabla F(x, \xi) - \nabla F(x^*, \xi)\|^2] + 2\mathbb{E}_\xi[\|\nabla F(x^*, \xi)\|^2].$$

The second term is bounded by $2\sigma^2$ (or $2\mathbb{E}[\|\nabla F(x^*, \xi)\|^2]$) based on assumption D.1. For the first term, we utilize the property of convex and L-smooth functions. Specifically, for any function g that is convex and L-smooth, we have $\|\nabla g(x) - \nabla g(y)\|^2 \leq 2L(g(x) - g(y) - \langle \nabla g(y), x - y\rangle)$. Applying this to $F(\cdot, \xi)$ with $y = x^*$:

$$\|\nabla F(x, \xi) - \nabla F(x^*, \xi)\|^2 \leq 2L\big(F(x, \xi) - F(x^*, \xi) - \langle \nabla F(x^*, \xi), x - x^*\rangle\big).$$

Taking the expectation with respect to ξ:

$$\begin{aligned}\mathbb{E}_\xi[\|\nabla F(x, \xi) - \nabla F(x^*, \xi)\|^2] &\leq 2L\mathbb{E}_\xi\big[F(x, \xi) - F(x^*, \xi) - \langle \nabla F(x^*, \xi), x - x^*\rangle\big] \\ &= 2L\big(f(x) - f(x^*) - \langle \nabla f(x^*), x - x^*\rangle\big).\end{aligned}$$

Since x^* is the minimizer of $f(x)$, we have $\nabla f(x^*) = 0$, causing the inner product term to vanish. Thus:

$$\mathbb{E}_\xi[\|\nabla F(x, \xi) - \nabla F(x^*, \xi)\|^2] \leq 2L(f(x) - f(x^*)).$$

Substituting this back into the first inequality yields the result:

$$\mathbb{E}_\xi[\|\nabla F(x, \xi)\|^2] \leq 2(2L(f(x) - f(x^*))) + 2\sigma^2 = 4L(f(x) - f(x^*)) + 2\sigma^2.$$

□

This lemma provides the key tool for controlling the error introduced by the stochastic oracle, which we now use to derive the convergence guarantee for two estimators.

Theorem D.3. *Let f be an L-smooth and convex function satisfying Assumption D.1. The ZO-Avg algorithm using a step size $\eta_t = \frac{\eta_0}{\sqrt{t+1}}$ with $\eta_0 \leq \frac{1}{4L(d+2)}$ satisfies the stability condition $\eta_t \leq \frac{q_t}{4L(q_t+d+1)}$ for all $t \geq 0$. After T iterations, it achieves the following convergence guarantee:*

$$\mathbb{E}[f(\bar{x}_T) - f(x^*)] \leq \frac{\|x_0 - x^*\|^2}{\eta_0 \sum_{t=0}^{T-1}(t+1)^{-1/2}} + \frac{\sigma^2}{2L}. \tag{13}$$

where $\overline{x}_T$ is the weighted average of the iterates. Under a fixed total query budget K, this convergence bound is optimized by choosing $q_t = 1$ for all t.

Proof. We start by analyzing the evolution of the squared distance to the optimum, $\|x_t - x^*\|^2$. The update rule is $x_{t+1} = x_t - \eta_t \hat{g}_{AVG}(x_t)$.

$$\|x_{t+1} - x^*\|^2 = \|x_t - x^*\|^2 - 2\eta_t \hat{g}_{AVG}(x_t)^T(x_t - x^*) + \eta_t^2 \|\hat{g}_{AVG}(x_t)\|^2$$

Let $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot|x_t]$ denote the expectation over the random direction matrix U_t and the stochastic sample ξ_t. We have $\mathbb{E}_t[\hat{g}_{AVG}(x_t)] = \nabla f(x_t)$. From Proposition 3.1 and Lemma D.2, the expected squared norm of the estimator is bounded as:

$$\mathbb{E}_t[\|\hat{g}_{AVG}(x_t)\|^2] = \frac{q_t + d + 1}{q_t} \mathbb{E}_{\xi_t}[\|\nabla F(x_t, \xi_t)\|^2] \leq \frac{q_t + d + 1}{q_t}(4L(f(x_t) - f(x^*)) + 2\sigma^2)$$

Taking the conditional expectation of the distance expansion and using the convexity of f, i.e., $f(x_t) - f(x^*) \leq \nabla f(x_t)^T(x_t - x^*)$:

$$\begin{aligned}
\mathbb{E}_t[\|x_{t+1} - x^*\|^2] &\leq \|x_t - x^*\|^2 - 2\eta_t \nabla f(x_t)^T(x_t - x^*) + \eta_t^2 \mathbb{E}_t[\|\hat{g}_{AVG}(x_t)\|^2] \\
&\leq \|x_t - x^*\|^2 - 2\eta_t(f(x_t) - f(x^*)) + \eta_t^2 \frac{q_t + d + 1}{q_t}(4L(f(x_t) - f(x^*)) + 2\sigma^2) \\
&= \|x_t - x^*\|^2 + \left(4L\eta_t^2 \frac{q_t + d + 1}{q_t} - 2\eta_t\right)(f(x_t) - f(x^*)) + 2\eta_t^2 \frac{q_t + d + 1}{q_t}\sigma^2
\end{aligned}$$

To ensure convergence, we require the coefficient of the suboptimality gap to be effectively negative. Specifically, we enforce:

$$\begin{aligned}
4L\eta_t^2 \frac{q_t + d + 1}{q_t} - 2\eta_t &\leq -\eta_t \\
4L\eta_t^2 \frac{q_t + d + 1}{q_t} &\leq \eta_t \\
\eta_t &\leq \frac{q_t}{4L(q_t + d + 1)}
\end{aligned}$$

Since $\frac{q_t}{q_t+d+1}$ is increasing in q_t, the tightest constraint occurs at $q_t = 1$, requiring $\eta_t \leq \frac{1}{4L(d+2)}$. Our choice of η_0 ensures this holds for all t. Substituting this back into the inequality:

$$\mathbb{E}_t[||x_{t+1} - x^*||^2] \leq ||x_t - x^*||^2 - \eta_t(f(x_t) - f(x^*)) + 2\eta_t^2 \frac{q_t + d + 1}{q_t}\sigma^2$$

Taking the full expectation, rearranging to isolate the gap $\delta_t = \mathbb{E}[f(x_t) - f(x^*)]$, and summing from $t = 0$ to $T - 1$:

$$\begin{aligned}
\sum_{t=0}^{T-1} \eta_t \delta_t &\leq \mathbb{E}[||x_0 - x^*||^2] - \mathbb{E}[||x_T - x^*||^2] + 2\sigma^2 \sum_{t=0}^{T-1} \eta_t^2 \frac{q_t + d + 1}{q_t} \\
&\leq ||x_0 - x^*||^2 + 2\sigma^2 \sum_{t=0}^{T-1} \eta_t^2 \frac{q_t + d + 1}{q_t}
\end{aligned}$$

By convexity and Jensen's inequality, for the weighted average iterate $\overline{x}_T$ with weights proportional to η_t:

$$\mathbb{E}[f(\overline{x}_T) - f(x^*)] \leq \frac{\sum_{t=0}^{T-1} \eta_t \delta_t}{\sum_{t=0}^{T-1} \eta_t} \leq \frac{||x_0 - x^*||^2 + 2\sigma^2 \sum_{t=0}^{T-1} \eta_t^2 \frac{q_t+d+1}{q_t}}{\sum_{t=0}^{T-1} \eta_t}$$

we use the stability condition once more:

$$\eta_t \leq \frac{q_t}{4L(q_t+d+1)} \quad \Longrightarrow \quad \eta_t^2 \frac{q_t+d+1}{q_t} \leq \frac{\eta_t}{4L}.$$

Therefore

$$2\sigma^2 \sum_{t=0}^{T-1} \eta_t^2 \frac{q_t+d+1}{q_t} \leq \frac{\sigma^2}{2L} \sum_{t=0}^{T-1} \eta_t.$$

Substituting this into the previous inequality yields

$$\mathbb{E}[f(\bar{x}_T) - f(x^*)] \leq \frac{\|x_0 - x^*\|^2}{\sum_{t=0}^{T-1} \eta_t} + \frac{\sigma^2}{2L}.$$

Finally, substituting $\eta_t = \eta_0/\sqrt{t+1}$ gives

$$\mathbb{E}[f(\bar{x}_T) - f(x^*)] \leq \frac{\|x_0 - x^*\|^2}{\eta_0 \sum_{t=0}^{T-1} (t+1)^{-1/2}} + \frac{\sigma^2}{2L}.$$

It remains to analyze the fixed-budget allocation. Suppose the total query budget satisfies$\sum_{t=0}^{T-1} q_t \leq K,\ q_t \geq 1$. Then necessarily $T \leq K$. Since the function

$$T \mapsto \sum_{t=0}^{T-1} (t+1)^{-1/2}$$

is strictly increasing in T, the above upper bound is minimized by maximizing T, namely by taking $T = K$. This is achieved by choosing $q_t = 1$ for all t. Hence, under a fixed total query budget, the above upper bound is minimized by the single-query allocation. □

Theorem D.4. *Let f be an L-smooth and convex function satisfying Assumption D.1. The ZO-Align algorithm using a diminishing step size $\eta_t = \frac{\eta_0}{\sqrt{t+1}}$, with $\eta_t \leq \frac{1}{4L}$, achieves the following convergence guarantee after T iterations:*

$$\mathbb{E}[f(\bar{x}_T) - f(x^*)] \leq \frac{d\|x_0 - x^*\|^2 + 2\eta_0^2\sigma^2 \sum_{t=0}^{T-1} \frac{q_t}{t+1}}{\eta_0 \sum_{t=0}^{T-1} \frac{q_t}{\sqrt{t+1}}}$$

where $\bar{x}_T$ is the weighted average of the iterates. Under the general cases where initialization error dominate the variance, with a fixed total query budget K, this convergence bound is optimized by choosing $q_t = d$ for $T = K/d$ iterations.

Proof. We start by analyzing the evolution of the squared distance to the optimum, $\|x_t - x^*\|^2$. The update rule is $x_{t+1} = x_t - \eta_t \hat{g}_{ALG}(x_t)$.

$$\|x_{t+1} - x^*\|^2 = \|x_t - x^*\|^2 - 2\eta_t \hat{g}_{ALG}(x_t)^T (x_t - x^*) + \eta_t^2 \|\hat{g}_{ALG}(x_t)\|^2$$

Let $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot|x_t]$ denote the expectation over the random direction matrix U_t and the stochastic sample ξ_t. Using the properties of the ZO-Align estimator, we have $\mathbb{E}_t[\hat{g}_{ALG}(x_t)] = \frac{q_t}{d} \nabla f(x_t)$ and $\mathbb{E}_t[\|\hat{g}_{ALG}(x_t)\|^2] = \frac{q_t}{d} \mathbb{E}_{\xi_t}[\|\nabla F(x_t, \xi_t)\|^2]$. Using Lemma D.2, this becomes:

$$\mathbb{E}_t[\|\hat{g}_{ALG}(x_t)\|^2] \leq \frac{q_t}{d}(4L(f(x_t) - f(x^*)) + 2\sigma^2)$$

Taking the conditional expectation of the distance expansion and using the convexity of f, i.e., $f(x_t) - f(x^*) \leq \nabla f(x_t)^T (x_t - x^*)$:

$$\begin{aligned}\mathbb{E}_t[\|x_{t+1} - x^*\|^2] &\leq \|x_t - x^*\|^2 - 2\eta_t \frac{q_t}{d} \nabla f(x_t)^T (x_t - x^*) + \eta_t^2 \mathbb{E}_t[\|\hat{g}_{ALG}(x_t)\|^2] \\ &\leq \|x_t - x^*\|^2 - \frac{2\eta_t q_t}{d}(f(x_t) - f(x^*)) + \eta_t^2 \frac{q_t}{d}(4L(f(x_t) - f(x^*)) + 2\sigma^2) \\ &= \|x_t - x^*\|^2 + \left(\frac{4L\eta_t^2 q_t}{d} - \frac{2\eta_t q_t}{d}\right)(f(x_t) - f(x^*)) + \frac{2\eta_t^2 q_t}{d}\sigma^2\end{aligned}$$

Given the step-size condition $\eta_t \leq \frac{1}{4L}$, we have $4L\eta_t \leq 1$. We can bound the coefficient of the suboptimality gap:

$$\frac{4L\eta_t^2 q_t}{d} - \frac{2\eta_t q_t}{d} = \frac{\eta_t q_t}{d}(4L\eta_t - 2) \leq \frac{\eta_t q_t}{d}(1 - 2) = -\frac{\eta_t q_t}{d}$$

Substituting this bound back into the inequality for the distance:

$$\mathbb{E}_t[\|x_{t+1} - x^*\|^2] \leq \|x_t - x^*\|^2 - \frac{\eta_t q_t}{d}(f(x_t) - f(x^*)) + \frac{2\eta_t^2 q_t}{d}\sigma^2$$

Rearranging, taking the full expectation with $\delta_t = \mathbb{E}[f(x_t) - f(x^*)]$ and $D_t = \mathbb{E}[\|x_t - x^*\|^2]$:

$$\frac{\eta_t q_t}{d}\delta_t \leq D_t - D_{t+1} + \frac{2\eta_t^2 q_t}{d}\sigma^2$$

Summing this inequality from $t = 0$ to $T - 1$:

$$\begin{aligned}\sum_{t=0}^{T-1} \frac{\eta_t q_t}{d}\delta_t &\leq \sum_{t=0}^{T-1}(D_t - D_{t+1}) + \sum_{t=0}^{T-1} \frac{2\eta_t^2 q_t}{d}\sigma^2 \\ &\leq D_0 - D_T + \frac{2\sigma^2}{d}\sum_{t=0}^{T-1} \eta_t^2 q_t \leq \|x_0 - x^*\|^2 + \frac{2\sigma^2}{d}\sum_{t=0}^{T-1} \eta_t^2 q_t\end{aligned}$$

Let $\alpha_t = \frac{\eta_t q_t}{d}$. By convexity of f and Jensen's inequality, the suboptimality of the weighted average iterate $\bar{x}_T = \frac{1}{\sum \alpha_t}\sum \alpha_t x_t$ is bounded by $\mathbb{E}[f(\bar{x}_T) - f(x^*)] \leq \frac{\sum \alpha_t \delta_t}{\sum \alpha_t}$. This gives:

$$\mathbb{E}[f(\bar{x}_T) - f(x^*)] \leq \frac{\|x_0 - x^*\|^2 + \frac{2\sigma^2}{d}\sum_{t=0}^{T-1} \eta_t^2 q_t}{\sum_{t=0}^{T-1} \frac{\eta_t q_t}{d}} = \frac{d\|x_0 - x^*\|^2 + 2\sigma^2 \sum_{t=0}^{T-1} \eta_t^2 q_t}{\sum_{t=0}^{T-1} \eta_t q_t}$$

Substituting the step size $\eta_t = \eta_0/\sqrt{t+1}$ yields the bound stated in the theorem.

To optimize the bound, we consider the case where the first term in the numerator dominates the second term. This is a reasonable regime in practice: η_0 is typically chosen to be small (often ranging from 10^{-2} to 10^{-7}), while d can be very large, from hundreds in classical problems to billions in LLM settings. In this case, minimizing the bound reduces to maximizing $\sum_{t=0}^{T-1} \frac{q_t}{\sqrt{t+1}}$. The term $\frac{1}{\sqrt{t+1}}$ represents the "value" or "effectiveness" of queries used at iteration t. Since this value is a strictly decreasing function of t, we should allocate our query budget to the earliest possible iterations to achieve the maximum sum. This implies a greedy strategy: use the maximum allowed queries, $q_t = d$, for the first $T = K/d$ iterations. This concentrates the entire budget K where it is most effective, thus maximizing the denominator and providing the tightest convergence bound. □

Empirical relevance of the initialization-dominated regime The allocation result for ZO-Align in Theorem D.4 is derived under the condition that the initialization term dominates the variance term in the convergence bound. We examine whether this condition is satisfied in the stochastic optimization regime used in our experiments.

Specifically, we consider the same stochastic regularized logistic-regression setting as in Section 6, and evaluate the relative contributions of the two terms in the numerator of the bound in Theorem D.4. The resulting proportions at different query budgets are reported in Table 5.

Table 5: Relative contributions of the initialization and variance terms in the ZO-Align stochastic convergence bound.

Queries used	Initialization term	Variance term
50	95.1%	4.9%
500	87.0%	13.0%
2,500	81.3%	18.7%
5,000	79.1%	20.9%

The initialization term accounts for the majority of the bound throughout the tested range. Thus, the initialization-dominated approximation underlying Theorem D.4 is consistent with the optimization horizon considered in our experiments. At the same time, the variance contribution increases as the query budget grows, indicating that this approximation becomes progressively less accurate toward the variance-dominated regime. Accordingly, we interpret the large-query allocation result for stochastic ZO-Align as a characterization of the initialization-dominated phase, rather than a uniform statement over the entire optimization trajectory.

E Experiment Setup

Classical objectives. All four classical problems use dimension $d = 1000$. The results are the average of the 10 repeated experiments.

For the strongly convex quadratic, $f(x) = \frac{1}{2}x^\top Ax + b^\top x$, we first sample a matrix $M \in \mathbb{R}^{d\times d}$ with i.i.d. $\mathcal{N}(0,1)$ entries and set $A = M^\top M + \epsilon I_d$ with a small $\epsilon > 0$ to ensure good conditioning. The vector $b \in \mathbb{R}^d$ is sampled from $\mathcal{N}(0, I_d)$.

For the convex logistic regression objective, we generate synthetic data by sampling feature vectors $a_i \in \mathbb{R}^d$ i.i.d. from $\mathcal{N}(0, I_d)$. We draw a ground-truth weight vector $w_{\text{true}} \sim \mathcal{N}(0, I_d)$ and assign binary labels $y_i \in \{-1, 1\}$ according to $\text{sign}(a_i^\top w_{\text{true}})$ (with ties broken arbitrarily). The objective is the empirical logistic loss $\frac{1}{m}\sum_{i=1}^m \log(1 + \exp(-y_i a_i^\top x))$.

The non-convex objective is the standard Rosenbrock function $f(x) = \sum_{i=1}^{d-1}\big[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2\big]$.

For the stochastic convex objective, we consider an ℓ_2-regularized logistic regression loss,

$$f(x) = \mathbb{E}_{(\mathbf{w},y)\sim\mathcal{D}}\big[\log(1 + \exp(-y\,\mathbf{w}^\top x))\big] + \tfrac{\rho}{2}\|x\|^2,$$

where $\mathcal{D}$ is a synthetic distribution generated in the same way as for the convex logistic regression. In the algorithm, this expectation is approximated using mini-batches of examples resampled at each iteration.

Classical experiments: implementation details. For each classical objective we run both ZO-Avg and ZO-Align with several query block sizes q and two total query budgets $K \in \{500, 20{,}000\}$, as described in Section 6. Unless otherwise stated, we use the idealized one-sided finite-difference estimator with a fixed smoothing parameter $\mu = 10^{-6}$. For ZO-Avg and ZO-Align in the deterministic settings, the step sizes are set according to the theoretically optimal choices in Proposition 4.1, namely $\eta_t = \frac{q_t}{L(q_t+d+1)}$ for ZO-Avg and $\eta_t = 1/L$ for ZO-Align, where L is the Lipschitz constant of the gradient (for the quadratic objective, L equals the largest eigenvalue of A). In the stochastic convex case we use diminishing step sizes of the form $\eta_t = \eta_0/\sqrt{t+1}$, with η_0 tuned by a small grid search. All reported curves are averaged over 10 independent runs with different random seeds.

LQR Experimental Details LQR is a standard optimal-control problem that has also been widely studied in the derivative-free and zeroth-order optimization literature. In our setup, the controller is a static state-feedback policy of the form

$$u_t = -Kx_t, \qquad K \in \mathbb{R}^{3\times 3}.$$

We optimize over the controller gain matrix K. After vectorizing K, the optimization variable lies in $\mathbb{R}^d$ with $d = 9$.

Following (Malik et al., 2019), the linear dynamical system and quadratic cost matrices are fixed as

$$A = \begin{bmatrix} 1 & 0 & -10 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -10 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

$$Q = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad R = \begin{bmatrix} 5 & -3 & 0 \\ -3 & 5 & -2 \\ 0 & -2 & 5 \end{bmatrix}.$$

and the state evolves according to

$$x_{t+1} = Ax_t + Bu_t.$$

For a given controller K, the performance objective is defined by a finite-horizon rollout cost. Rather than taking expectation over random initial states, we follow a deterministic protocol and use the three standard basis vectors

$$e_1, e_2, e_3$$

as initial states. For each initial state, we roll out the closed-loop dynamics and compute the cumulative quadratic cost, and then average over the three trajectories. The resulting objective is

$$J(K) = \frac{1}{3} \sum_{i=1}^{3} \sum_{t=0}^{H-1} \left(x_t^\top Q x_t + u_t^\top R u_t \right),$$

where $H = 200$ is the rollout horizon.

To obtain a conservative smoothness estimate for step-size selection, we approximate the Hessian of J at the initialization $K_0 = 0$ using central finite differences. Let e_i denote the i-th standard basis vector in $\mathbb{R}^d$. The (i, j)-th entry of the Hessian approximation is computed as

$$\widehat{H}_{ij} = \frac{J(K + \varepsilon e_i + \varepsilon e_j) - J(K + \varepsilon e_i - \varepsilon e_j) - J(K - \varepsilon e_i + \varepsilon e_j) + J(K - \varepsilon e_i - \varepsilon e_j)}{4\varepsilon^2}.$$

We then define a conservative smoothness constant by

$$L_{\text{safe}} = 2\, \lambda_{\max}(\widehat{H}).$$

This L_{safe} is used as a safe upper bound on the local curvature scale of the objective near initialization, following the standard practice in LQR-based derivative-free optimization experiments.

High-dimensional language model fine-tuning. For the high-dimensional experiments we fine-tune the Qwen3-0.6B model on the SST-2 and CB sentence classification datasets using zeroth-order optimization on the cross-entropy loss. We follow a standard sequence-classification protocol: each input sentence (or sentence pair) is fed to the model, the representation of the special classification token is passed through a linear head to produce class logits, and the cross-entropy with respect to the ground-truth labels is used as the objective. All trainable parameters (including the classification head) are flattened into a single vector in $\mathbb{R}^d$, where d is on the order of the total number of parameters of the fine-tuned model.

At each zeroth-order update step we sample q i.i.d. Gaussian directions and form the direction matrix $U \in \mathbb{R}^{d \times q}$. ZO-Avg uses the standard averaging estimator based on these q directions. For ZO-Align we use the diagonal approximation described in Section 6.3: instead of explicitly forming and inverting the full Gram matrix $U^\top U$, we compute only its diagonal $D = \mathrm{diag}(U^\top U)$ and apply the update $\hat{g}_{\text{ALG-diag}}(x) = UD^{-1}v$, where $v \in \mathbb{R}^q$ collects the one-sided finite-difference directional derivatives. This approximation is motivated by the high-dimensional analysis in Section 5, which shows that $U^\top U$ concentrates near a scaled identity when $d \gg q$.

We fix a total query budget of $K = 20{,}000$ function evaluations and consider query block sizes $q \in \{1, 10, 100\}$. For each estimator (ZO-Avg and diagonal ZO-Align) and each choice of q, we run the optimizer until the budget K is exhausted. A constant step size is used within each run; for fairness, we tune this step size separately for each (estimator, q) pair by a small logarithmic grid search.

Hardware and Software Our experiments were conducted using Python 3.9 and PyTorch 1.12. The underlying system was Ubuntu 18.04, equipped with an Intel Xeon Gold 5320 CPU and two NVIDIA RTX 3090 GPUs.