Proceedings of Machine Learning Research vol 336:1–49, 2026 39th Annual Conference on Learning Theory

Optimal Reconstruction from Linear Queries

Yuval Filmus YUVALFI@CS.TECHNION.AC.IL
Technion – Israel Institute of Technology

Shay Moran SMORAN@TECHNION.AC.IL
Technion – Israel Institute of Technology
Google Research

Elizaveta Nesterova ELIZAVETAN@CAMPUS.TECHNION.AC.IL
Technion – Israel Institute of Technology

Editors: Steve Hanneke and Tor Lattimore

Abstract

We study the problem of reconstructing an unknown point in $\mathbb{R}^d$ from approximate linear queries. This setting arises naturally in applications ranging from low-dimensional remote sensing and signal recovery to high-dimensional data analysis and privacy-sensitive inference. Our main goal is to characterize the optimal *reconstruction error* as a function of the number of queries T, the ambient dimension d, and the noise parameter δ.

We first analyze the limit $T \to \infty$ and show that the optimal reconstruction error converges to the explicit value $\sqrt{2d/(d+1)}\,\delta$, which plays a role analogous to the Bayes optimal error in supervised learning. When the dimension is fixed, we show that the excess error above this limit decays *doubly exponentially* fast as $T \to \infty$, a rate that is significantly faster than those typically encountered in learning curves. When the dimension grows, we show that a number of queries on the order of $\exp(d)$ is necessary and sufficient to achieve vanishing excess error. Finally, we introduce and analyze an improper variant of the reconstruction problem.

From a technical perspective, our main contribution is a generalization of Jung's theorem (1901). The classical theorem bounds the maximum possible radius of a set of diameter 1 and characterizes extremal bodies. Our generalization provides a robust variant that characterizes near-extremal bodies and is proved via geometric and dynamical arguments exploiting symmetry and Lie group actions.

1. Introduction

In this work we study the following question:

> *How accurately can an unknown point $x^* \in \mathbb{R}^d$ be reconstructed using linear measurements?*

We formalize this question via a *linear reconstruction game*. In this game, an adversary holds a secret point $x^* \in \mathbb{R}^d$, and at each round:

1. The reconstructor submits a linear query in the form of a unit vector $v \in \mathbb{R}^d$, $\|v\|_2 = 1$.[1]
2. The adversary replies with an adversarially chosen value r satisfying

$$|r - \langle v, x^* \rangle| \le \delta,$$

where $\delta > 0$ is a fixed additive noise parameter, and $\langle v, x\rangle$ is the standard inner product.

1. Equivalently, we can allow arbitrary vectors v, and require the error to be at most $\delta\|v\|_2$. Without this normalization, the reconstructor can cheat by scaling v by a large amount, effectively reducing the error.

The reconstructor can choose its queries adaptively, based on the answers to previous queries. After a given number of rounds T, it outputs an estimate $\hat{x}$, aiming to minimize the ℓ_2-distance $\|\hat{x} - x^*\|_2$.

The linear reconstruction game captures phenomena that arise naturally in both low- and high-dimensional settings. In low-dimensional geometric problems, it models adaptive sensing and signal acquisition, where measurement directions are chosen sequentially in order to localize an unknown signal or object in space. Related formulations also arise in data analysis and privacy — often, but not exclusively, in high dimensions — where linear queries are used to probe a dataset, and noise is added to limit reconstruction of sensitive information.

A discrete variant of this problem, in which both the hidden point and the queries are restricted to the Boolean cube $\{0, 1\}^n$, was introduced by Dinur and Nissim (2003), and played a foundational role in the development of differential privacy. Subsequent work in private and adaptive data analysis, as well as in the statistical query model, has studied this setting extensively, primarily from the perspective of preventing reconstruction (Bassily et al., 2016; Dwork et al., 2015a,b; Reyzin, 2020). Related questions also appear in the adaptive sensing literature (Arias-Castro et al., 2013; Haupt et al., 2009), though most of that work assumes stochastic noise, whereas we consider adversarial noise.

We study the optimal minimax reconstruction error against an *arbitrary* adversary:

Main Question

What is the optimal reconstruction error that can be guaranteed against an arbitrary (worst-case) adversary, as a function of the error parameter δ, the number of rounds T, and the dimension d?

From a learning-theoretic perspective, this question can be viewed as a natural instance of interactive learning, in which a learner adaptively queries an oracle and seeks to reconstruct an unknown target from noisy feedback. Here, the secret point x^* plays the role of the concept to be learned.

2. Main results

In this section, we outline the key definitions and results. More detailed definitions and statements are deferred to Appendix A.

The *optimal reconstruction error* is the smallest worst-case loss achievable by a reconstructor, i.e., the smallest worst-case loss that the "best" reconstructor can achieve. We denote the optimal reconstruction error for noise level $\delta > 0$, dimension d, and T interaction rounds by $\mathrm{OPT}_d(T, \delta)$. Our first main theorem identifies the reconstruction error achievable (in the limit) given an unlimited number of rounds.

Main Theorem 1 (Asymptotic optimal error)

$$\mathrm{OPT}_d(\infty, \delta) := \lim_{T \to \infty} \mathrm{OPT}_d(T, \delta) = \sqrt{\frac{2d}{d+1}}\delta.$$

From a learning-theoretic viewpoint, the limiting value $\mathrm{OPT}_d(\infty, \delta)$ is analogous to the *Bayes-optimal* error: it is the best performance achievable under the interaction and noise constraints of the model. This motivates studying the *excess reconstruction error*,

$$\mathsf{ExcessErr}_d(T, \delta) \ := \ \mathrm{OPT}_d(T, \delta) - \mathrm{OPT}_d(\infty, \delta).$$

Our second main theorem, which is also our main technical contribution, shows that $\mathsf{ExcessErr}_d(T, \delta)$ decays at a doubly exponential rate in T when $d \geq 2$ (when $d = 1$, there is only one unit vector $v = 1$ up to sign, and so one round suffices).

Main Theorem 2 (Doubly-exponential decay of excess error)
For every dimension $d \geq 2$ there exists T'_d such that for every $T \geq T'_d$:

$$\mathsf{ExcessErr}_d(T, \delta) = 2^{-2^{\Theta_d(T)}}\delta.$$

The linear dependence on δ in Main Theorems 1 and 2 is not accidental: in fact, the reconstruction game is invariant under rescaling δ. We elaborate on this further in the Technical Overview and prove it in Appendix A.

The proof of Main Theorem 2 relies on a reconstruction algorithm with a natural two-phase structure. In the first phase, which can be viewed as a preprocessing step, the reconstructor queries a sufficiently dense (but oblivious) set of directions in order to obtain a coarse geometric localization of the hidden point. In the second phase, the reconstructor exploits this localization to adaptively refine its estimate, leading to the doubly-exponential decay of the excess error. This rapid refinement crucially relies on our robust Jung theorem: a robust variant of Jung's theorem, which identifies a small set of particularly informative directions whose queries sharply reduce the remaining uncertainty. We elaborate on this robust variant of Jung's theorem and explain how it is leveraged algorithmically in the proof overview in the next section.

The cost of the first, preprocessing phase grows rapidly with the dimension d, as it requires querying a dense net of directions on the unit sphere.

This naturally raises the question of whether such a dependence on the dimension is inherent. Our final main theorem answers this question in the affirmative and identifies $T = \exp(d)$ as a threshold for achieving vanishing excess error, in the following sense:

Main Theorem 3 (Dimension-dependent query budgets) *Let $T : \mathbb{N} \to \mathbb{N}$ be a query budget as a function of the dimension d.*

1. *If $T(d) = 2^{o(d)}$ is subexponential in d then*

$$\mathsf{ExcessErr}_d(T(d), \delta) \xrightarrow{d \to \infty} \infty.$$

2. *If $T(d) = 2^{\omega(d)}$ is superexponential in d then*

$$\mathsf{ExcessErr}_d(T(d), \delta) \xrightarrow{d \to \infty} 0.$$

Improper reconstructors. An accurate reconstruction of the point x^* immediately yields accurate answers to future linear queries: indeed, if $\|\hat{x} - x^*\|_2 \leq \varepsilon$, then for every unit vector v,

$$\left|\langle v, \hat{x}\rangle - \langle v, x^*\rangle\right| = \left|\langle v, \hat{x} - x^*\rangle\right| \leq \|\hat{x} - x^*\|_2 \leq \varepsilon.$$

This observation motivates the study of *improper reconstructors*, whose goal is to output a function $\hat{G} \colon S^{d-1} \to \mathbb{R}$ (where S^{d-1}, the unit sphere in $\mathbb{R}^d$, consists of all unit vectors) minimizing

$$\sup_{v \in S^{d-1}} \left|\hat{G}(v) - \langle v, x^*\rangle\right|.$$

We analyze this setting as well; here we briefly summarize the results and highlight the key differences from the proper setting, deferring complete statements and proofs to Appendix C.

In the improper setting, the optimal reconstruction error converges, as $T \to \infty$, to the value δ, which is strictly smaller than the optimal limit in the proper setting, given by $\sqrt{\frac{2d}{d+1}}\,\delta \approx \sqrt{2}\,\delta$. While the excess error in the proper setting decays doubly exponentially fast in T, convergence in the improper setting is only polynomial. Specifically, for any fixed dimension $d \geq 2$,

$$\mathrm{OPT}_d^{\mathrm{improper}}(T, \delta) = \left(1 + \Theta_d(T^{-2/(d-1)})\right)\delta.$$

At first glance, it may seem paradoxical that improper reconstruction converges more slowly, since every proper reconstructor induces an improper one via $\hat{G}(v) = \langle v, \hat{x} \rangle$. The resolution is that improper reconstruction converges to a strictly smaller limiting error, and the slower rate reflects the greater difficulty of approaching this stronger benchmark.

This improvement, however, comes at a cost. While a proper reconstructor outputs a single point in $\mathbb{R}^d$, an improper reconstructor outputs a function on S^{d-1}, which is an infinite object and raises substantial space complexity concerns. In particular, the improper reconstructor underlying our bounds must retain the entire interaction history in order to answer future queries.

Finally, as in the proper setting, the dependence of the hidden constants on the dimension is exponential; in particular, Main Theorem 3 applies to the improper setting as well.

Comparison with prior work. Our work is most closely related to the reconstruction problem of Moran and Nesterova (2025), in which the unknown object is a point $x^\star$ in a metric space (X, dist_X): the reconstructor adaptively queries points $q \in X$, receives approximate values of the distances $\mathrm{dist}_X(q, x^\star)$, and aims to output a point as close as possible to $x^\star$. In our setting, the unknown object is instead a vector in $\mathbb{R}^d$, the queries are linear functionals, and the answers are approximate values of these functionals.

At the level of limiting guarantees, Main Theorem 1 is a direct analogue of Theorem 2 in Moran and Nesterova (2025): both identify the optimal reconstruction error in the limit of infinitely many queries. In the Euclidean setting, the two results yield identical guarantees and rely on the same underlying geometric invariants, as discussed in the Technical Overview.

The main difference is quantitative. Moran and Nesterova (2025) show that, in the Euclidean setting, the limiting error is not attainable in finitely many rounds, but the rate at which it is approached is left largely undetermined there: the proof of their Theorem 6 yields a doubly exponentially small lower bound on the excess error, while the covering argument behind their Theorem 2 implicitly gives only an inverse-polynomial upper bound. Our main contribution is to determine this rate for linear queries: we prove matching upper and lower bounds in fixed dimension (Main Theorem 2) and characterize the dependence on the dimension (Main Theorem 3). The lower bound in Main Theorem 2 adapts the adversary strategy underlying the proof of Theorem 6 in Moran and Nesterova (2025) to the linear-query model. It remains an interesting question whether the upper-bound techniques developed here can be adapted to the distance-query framework of Moran and Nesterova (2025).

Organization. Section 3 provides a proof overview and highlights the key technical contributions of our work. In Section 4 we discuss open directions. Appendix A gives the formal model and basic properties of the reconstruction game. The proofs of the main theorems are given in Appendix B, and the results for improper reconstruction are given in Appendix C. The technical tools underlying the proofs are developed in Appendices D, E, F and G.

3. Technical Overview

We organize the proof overview as follows. We first collect several technical tools that are useful for explaining the key ideas underlying our proofs. We then outline the proofs of the main theorems, each in a separate subsection.

Feasible region. Assume that on query v, the adversary returns an answer $r \in \mathbb{R}$. What information about the secret point is now available to the reconstructor?

The reconstructor can conclude only that the secret x^* lies in $\{x \in \mathbb{R}^d : |\langle x, v\rangle - r| \le \delta\}$, which is a strip orthogonal to v of width 2δ. Therefore, after observing the whole interaction transcript (i.e., the sequence of queries v_i and answers r_i), the reconstructor only learns that the secret lies in the *intersection* of the corresponding strips. Formally, the secret must belong to the set of all points $x \in \mathbb{R}^d$ that are consistent with the transcript up to noise level $\delta > 0$:

$$\Phi_T := \left\{x \in \mathbb{R}^d \;\middle|\; |\langle v_i, x\rangle - r_i| \le \delta \;\text{ for all } i \in [T]\right\}. \tag{1}$$

We call Φ_T the *feasible region*.

What worst-case error can the reconstructor guarantee after observing the transcript $\{(v_i, r_i)\}_{i\in[T]}$, that is, after identifying the feasible set Φ_T? Suppose the reconstructor outputs an estimate $\hat{x}_T$. In the worst case, the true point x^* may be a point in Φ_T that maximizes its distance from $\hat{x}_T$. Consequently, to minimize the worst-case error, the reconstructor should choose an output that minimizes the maximum distance to all points in Φ_T. This motivates the following definition.

For a set $A \subseteq \mathbb{R}^d$, define its *radius* (also known as the *Chebyshev radius*) by

$$\texttt{rad}(A) \;=\; \inf_{x\in\mathbb{R}^d} \sup_{y\in A} \|x - y\|_2.$$

Equivalently, $\texttt{rad}(A)$ is the radius of the minimal enclosing ball of A. See Figure 1 for a visualization.

Remark *The feasible region is a central object in this game. It plays a key role throughout all of our theorems: both the reconstructor's algorithm and the adversary's strategy exploit its geometric properties. In particular, we formulate the guarantees and the convergence rates by tracking how the radius of the feasible region evolves with the transcript.*

Scale invariance. As a final preparatory fact before turning to the proof overview, we record a simple scaling property of the optimal reconstruction error that will be useful throughout the paper. Recall that $\mathrm{OPT}_d(T, \delta)$ denotes the worst-case reconstruction error achievable by the reconstructor under noise level $\delta > 0$. For every $T \in \mathbb{N}$ and every $\delta > 0$, the optimal error satisfies

$$\mathrm{OPT}_d(T, \delta) = \delta\, \mathrm{OPT}_d(T, 1). \tag{2}$$

Intuitively, this is the case since we can obtain an algorithm for error rate δ from an algorithm for error rate 1 by scaling everything by δ. The formal proof appears in Lemma 9.

This scale invariance allows us to simplify several arguments. In particular, in the proof of Main Theorem 2 we work with normalized noise $\delta = 1/2$ without loss of generality. Similarly, in Main Theorem 3, we set $\delta = \Theta(\sqrt{\ln T})$, which simplifies the analysis by allowing us to work with standard Gaussian vectors.

The rest of the technical overview outlines the key ideas behind the three main theorems, in order.

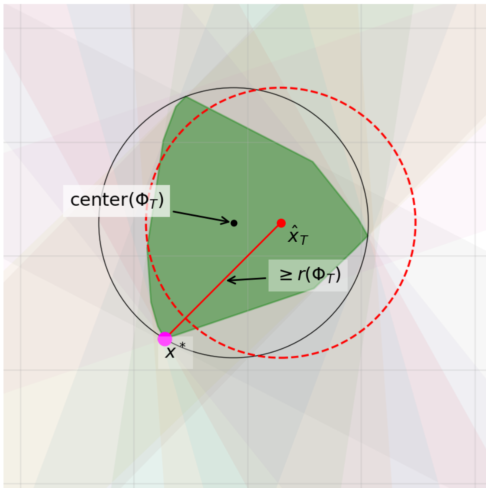

Figure 1: A secret point x^* and an output $\hat{x}_T$ such that $\|\hat{x}_T - x^*\| \geq \mathtt{rad}(\Phi_T)$.

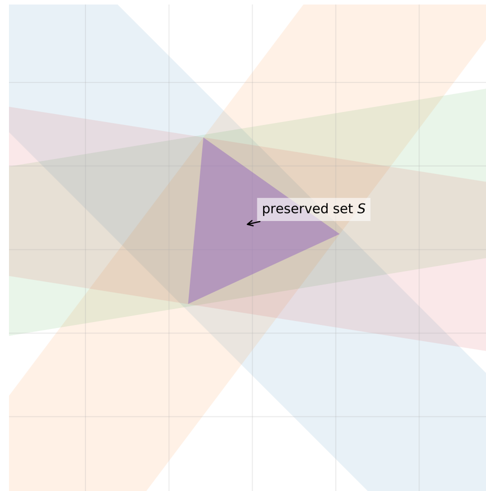

Figure 2: Adversary strategy to maintain S of diam $S \leq 2\delta$ inside Φ_T.

3.1. Asymptotic error: Main Theorem 1

This subsection outlines the key ideas in the proof of Main Theorem 1. The approach used in the proof of Main Theorem 1 is not new; a closely related approach was used in Moran and Nesterova (2025) for the distance reconstruction game on arbitrary metric spaces. In particular, the key step in their proof of Theorem 2 is to bound the Chebyshev radius of the feasible region via a bound on its diameter. We nevertheless present these ideas here, since they form the basis of the intuition behind Main Theorems 2 and 3. Recall that Main Theorem 1 identifies the asymptotic optimal reconstruction error:

$$\mathrm{OPT}_d(\infty, \delta) := \lim_{T \to \infty} \mathrm{OPT}_d(T, \delta) = \sqrt{\frac{2d}{d+1}}\delta.$$

Upper bound. Assume that the reconstructor queries *all* directions $v \in S^{d-1}$; we can simulate this by querying a dense enough net. We first bound the *diameter* of the resulting feasible region Φ:

$$\mathrm{diam}(\Phi) \leq 2\delta.$$

Indeed, for any two points $A, B \in \Phi$, we can consider the direction $v = \frac{A-B}{\|A-B\|_2}$. By definition of the feasible region, $|\langle A - B, v \rangle| \leq 2\delta$, and so $\|A - B\|_2 \leq 2\delta$.

We derive a bound on the Chebyshev radius of the feasible region using Jung's theorem; for a modern proof, see (Gruber, 2007, Theorem 3.3).

Theorem 4 (Jung, 1901) *For every bounded set $S \subseteq \mathbb{R}^d$ with diameter D, its Chebyshev radius satisfies* $\mathtt{rad}(S) \leq \mathrm{Jung}_d\, D$, *where* $\mathrm{Jung}_d = \sqrt{\frac{d}{2(d+1)}}$ *is known as* Jung's constant.

In particular, since the feasible region has diameter at most 2δ, Jung's theorem yields the upper bound

$$\mathtt{rad}(\Phi) \leq \mathrm{Jung}_d \cdot 2\delta = \sqrt{\frac{2d}{d+1}}\delta.$$

Lower bound. We design an adversary which ensures that the feasible region contains some fixed set S whose Chebyshev radius is $\mathrm{Jung}_d \cdot 2\delta$. Motivated by the upper bound, we pick S to be the vertex set of a regular simplex of edge length 2δ, which is tight for Jung's theorem.

The adversary proceeds as follows. Given an arbitrary direction v, it considers the interval $\{\langle s, v\rangle : s \in S\}$. Since S has diameter 2δ, this interval has length at most 2δ, and the adversary reports its midpoint. This way, it ensures that the entire set S lies in the strip corresponding to v. See Figure 2 for a visualization.

3.2. Excess error: Main Theorem 2

The aim of this section is to outline the proof of Main Theorem 2. Recall that this result determines the rate at which the excess error decays:

Main Theorem 2 restatement. *For every dimension $d \geq 2$ there exists T'_d s.t. for every $T \geq T'_d$,*

$$\mathsf{ExcessErr}_d(T, \delta) = 2^{-2^{\Theta_d(T)}}\delta.$$

To prove this result, we first note that by the scaling property we may assume throughout that the noise is normalized to $\delta = \frac{1}{2}$. Under this normalization, it suffices to show that for every dimension $d \geq 2$ there exists T'_d such that for every $T \geq T'_d$,

$$\mathrm{OPT}_d(T, 1/2) = \mathrm{Jung}_d + 2^{-2^{\Theta_d(T)}},$$

where $\mathrm{Jung}_d = \mathrm{OPT}_d(\infty, 1/2) = \sqrt{\frac{d}{2(d+1)}}$ is Jung's constant (see Theorem 4).

The doubly exponential rate in Main Theorem 2 is unusual. Moran and Nesterova (2025) proved a doubly exponential lower bound in a related setting (distance queries), which we repurpose for our setting, but they didn't prove a matching upper bound. The argument underlying the proof of the upper bound in Main Theorem 1 - querying a dense enough net of directions - only yields an upper bound of the form $O(T^{-2/(d-1)}\delta)$. Instead, the proof of the upper bound in Main Theorem 2, which is our main technical contribution, uses a boosting step which reduces an excess error of ϵ to an excess error of $O(\epsilon^2)$ using constantly many queries (the exact number depending on the dimension).

The main idea of the proof of Main Theorem 1 is to reduce the diameter of the feasible region and then apply Jung's theorem to deduce a bound on the Chebyshev radius. This is lossy since there is no guarantee that Jung's theorem is tight for the feasible region. Instead, in the proof of Main Theorem 2 we focus on eliminating subsets of the feasible region with large Chebyshev radius.

Suppose that the feasible region Φ_T at time T has Chebyshev radius at least $\mathrm{Jung}_d + \epsilon$. By Carathéodory's theorem, this is witnessed by some set $S \subseteq \Phi_T$ of size at most $d + 1$ whose Chebyshev radius is at least $\mathrm{Jung}_d + \epsilon$. We can think of S as forming the vertex set of a simplex. If we query the directions of all edges of the simplex, then S can no longer be inside the feasible region, potentially decreasing the Chebyshev radius of the feasible region.

The issue with this strategy is that there could be many such sets, and we cannot afford to query the directions corresponding to all of them. To overcome this difficulty, we prove a robust version of Jung's theorem, which shows that all such sets S are close to a single vertex set Δ. Querying the directions corresponding to the edges of Δ will reduce the Chebyshev radius of the feasible region to $1 + O(\epsilon^2)$.

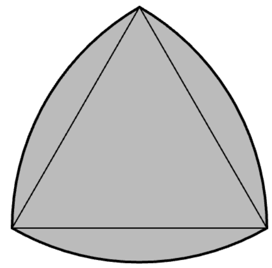

Figure 3: Reuleaux triangle of diameter 1 and Chebyshev radius $\mathrm{Jung}_2 = \sqrt{\frac{1}{3}}$.
The vertices of the highlighted triangle constitute the unique witness.

3.2.1. A ROBUST JUNG'S THEOREM

Jung's theorem states that a set $S \subseteq \mathbb{R}^d$ with diameter 1 has Chebyshev radius at most Jung_d. Furthermore, if the Chebyshev radius is exactly Jung_d then S contains the vertex set of a regular simplex Δ with edge length 1, which is moreover unique. However, S itself could contain points beyond Δ. For example, S could be a Reuleaux triangle (see Figure 3).

The robust version of Jung's theorem describes the structure of sets $S \subseteq \mathbb{R}^d$ with diameter $1+\beta$ whose Chebyshev radius is at least Jung_d. While we cannot say that such sets are close to a regular simplex (this is false even when $\beta = 0$), we are able to characterize the "extremal parts" of S, which we call witnesses. A subset $W \subseteq S$ is a *witness* for S if $|W| = d+1$ and

$$\mathtt{rad}(W) \geq \mathrm{Jung}_d.$$

We would like to say that the witnesses cluster around one another. We formalize this using the following definition: two sets W_1, W_2 are *r-close* if

$$W_1 \subset W_2 + B(r) \qquad \text{and} \qquad W_2 \subset W_1 + B(r),$$

where $B(r)$ is the ball of radius r and addition is Minkowski sum; equivalently, $W + B(r)$ consists of all points which are at distance at most r from some point in W.

We can now state the robust version of Jung's theorem, which describes the structure of sets which are almost tight for Jung's theorem.

Theorem 5 (Robust Jung Theorem) *Let $d \geq 1$ and let $S \subset \mathbb{R}^d$ be a closed set satisfying*

$$\mathrm{diam}(S) \leq 1+\beta \qquad \textit{and} \qquad \mathtt{rad}(S) \geq \mathrm{Jung}_d,$$

for some $0 \leq \beta \leq \beta_d$, where $\beta_d > 0$ depends only on d. Then, there exists a set of points

$$\Delta = \{x_0, x_1, \ldots, x_d\} \subset \mathbb{R}^d$$

forming the vertex set of a regular simplex of edge length 1 such that every witness $W \subseteq S$ is $C_d\beta$-close to Δ, for a constant $C_d > 0$ depending only on d.

We first explain how to use this theorem to prove the upper bound in Main Theorem 2, and then how it suggests the adversary strategy for the lower bound. Finally, we sketch the proof of the theorem itself.

3.2.2. UPPER BOUNDS ON EXCESS ERROR

To obtain a doubly exponential upper bound, it suffices to construct a reconstructor strategy and constants T'_d and $c_d > 0$ such that for all $T \geq T'_d$,

$$\mathtt{rad}(\Phi_T) \leq \mathrm{Jung}_d + 2^{-2^{c_d \cdot (T - T'_d)}}.$$

In particular, for $T \geq 2T'_d$ we have $T - T'_d \geq T/2$, and therefore $\mathtt{rad}(\Phi_T) \leq \mathrm{Jung}_d + 2^{-2^{(c_d/2)\,T}}$, which is a doubly exponential convergence rate in T (up to constants in the exponent). The formal proof of this Theorem is given in Appendix B.3.1

The algorithm proceeds in two phases. In the preprocessing phase of T'_d queries, the reconstructor bounds the diameter of the feasible region so that Robust Jung Theorem (Theorem 5) applies to it. After this point, the algorithm enters a refinement phase. Each refinement step consists of a fixed batch of $\binom{d+1}{2}$ queries, and has the effect of squaring the current excess error (up to dimension-dependent constants). Iterating this refinement yields the desired doubly exponential convergence.

The key geometric object helping to control the Chebyshev diameter of the feasible region is the *witness core*:

$$\mathrm{Witness}_t := \bigcup_{\substack{W \subset \Phi_t: \\ W \text{ is a witness}}} W, \tag{3}$$

that is, the union of all witnesses. Intuitively, $\mathrm{Witness}_t$ is the part of the feasible region that prevents the Chebyshev radius from dropping below $\mathrm{Jung}_d = \mathrm{OPT}_d(\infty, 1/2)$. Indeed, $\mathtt{rad}(\Phi_t) < \mathrm{Jung}_d$ holds if and only if $\mathrm{Witness}_t$ is empty.

By Lemma 22 (via Helly's theorem), there exists a finite set $W \subseteq \Phi_t$ of size $d+1$ such that $\mathtt{rad}(\Phi_t) = \mathtt{rad}(W)$. Assuming $\mathtt{rad}(\Phi_t) \geq \mathrm{Jung}_d$, the set W is a witness for Φ_t, and hence $W \subseteq \mathrm{Witness}_t$ by definition of the witness core. Applying Jung's inequality to W, we obtain

$$\mathtt{rad}(\Phi_t) = \mathtt{rad}(W) \leq \mathrm{Jung}_d \,\mathrm{diam}(W) \leq \mathrm{Jung}_d \,\mathrm{diam}(\mathrm{Witness}_t).$$

Consequently, it suffices to show that $\mathrm{diam}(\mathrm{Witness}_t)$ converges to 1 doubly exponentially fast as t grows. Accordingly, the reconstructor's goal is to reduce the diameter of the witness core.

Preprocessing. At the beginning of the game, the reconstructor obliviously queries a sufficiently dense covering of directions in order to ensure that the diameter of the feasible region satisfies $\mathrm{diam}(\Phi_t) \leq 1 + \beta_d$, where β_d is the constant appearing in the statement of Theorem 5.

This preprocessing step is where we pay to control $\mathrm{diam}(\Phi_t)$. This cost is unavoidable in high dimensions: Main Theorem 3 shows that with a subexponential query budget one cannot uniformly bound the diameter. In particular, an exponential-in-d number of queries is necessary before the refinement phase.

Refinement. Assume that $\mathrm{diam}(\mathrm{Witness}_t) \leq 1 + \beta$ at time t. Theorem 5 then applies to the witness core $\mathrm{Witness}_t$. Hence, there exists a vertex set $\Delta := \{x_0, x_1, \ldots, x_d\}$ of a single regular simplex such that *every* witness W of Φ_t is $O_d(\beta)$-close to Δ. In particular, since $\mathrm{Witness}_t$ is the union of such witnesses, it follows that

$$\mathrm{Witness}_t \subseteq \Delta + B\big(O_d(\beta)\big).$$

We then query the $\binom{d+1}{2}$ directions parallel to the edges of Δ, and denote the resulting feasible region by $\Phi_{t_{\text{new}}}$ and its witness core by $\mathrm{Witness}_{t_{\text{new}}}$. Since $\Phi_{t_{\text{new}}} \subseteq \Phi_t$, we also have $\mathrm{Witness}_{t_{\text{new}}} \subseteq \mathrm{Witness}_t$.

To show that $\mathrm{diam}(\mathrm{Witness}_{t_{\text{new}}}) \leq 1 + O_d(\beta^2)$, take any two points $y_1, y_2 \in \mathrm{Witness}_{t_{\text{new}}}$, and let x_1 and x_2 be the vertices of Δ that are $O_d(\beta)$-close to y_1 and y_2, respectively.

We claim that $\|y_1 - y_2\| \leq 1 + O_d(\beta^2)$. If $x_1 = x_2$, then y_1 and y_2 are already $O_d(\beta)$-close. Otherwise, consider the queried direction $v := \overrightarrow{x_1 x_2}$. Because y_1 and y_2 lie within $O_d(\beta)$ of the endpoints of the edge direction v, a straightforward geometric estimate shows that the segment $y_1 y_2$ is tilted from v by at most an $O_d(\beta)$ angle. Feasibility of y_1, y_2 after querying v implies

$$1 \geq \langle y_2 - y_1, v \rangle = \|y_1 - y_2\| \cdot \cos \angle(\overrightarrow{y_1 y_2}, v) = \|y_1 - y_2\| \cdot \cos O_d(\beta),$$

and hence $\|y_1 - y_2\| \leq 1 + O_d(\beta^2)$, using the Taylor expansion of cosine: $\cos\gamma \approx 1 - \frac{1}{2}\gamma^2$. Since this holds for any two points taken from $\mathrm{Witness}_{t_{\text{new}}}$, the bound on diameter follows. This concludes the refinement step.

Formally, the upper bound is proved in Appendix B.3.1.

3.2.3. Lower bounds on excess error

Although Theorem 5 does not by itself yield the lower bound, its formulation suggests an optimal adversary strategy. By Theorem 5, after the preprocessing step, in the worst-case scenario for the reconstructor (i.e., when $\mathtt{rad}(\Phi_T) > \mathrm{Jung}_d$), all witnesses must cluster near the vertex set of a single regular simplex.

This observation forms the core of the adversary's strategy. At each round t, the adversary ensures that the set $\Delta_t + B(\alpha_t)$ lies inside the feasible region, where Δ_t is the vertex set of some regular simplex and the radii α_t can be chosen to satisfy a recurrence of the form $\alpha_t \geq c\,\alpha_{t-1}^2$ for a constant $c > 0$. In particular, such a strategy forces the excess error at time t to be at least α_t, since

$$\mathtt{rad}(\Phi_t) \geq \mathtt{rad}(\Delta_t + B(\alpha_t)) = \mathrm{Jung}_d + \alpha_t.$$

Iterating this recurrence yields a doubly exponential rate.

This proof follows the same template as in Moran and Nesterova (2025), and relies on a careful rotation of Δ_{t-1} to obtain the new simplex Δ_t. We defer the details to Appendix B.3.2.

3.2.4. Robust Jung Theorem: proof

Finally, we explain the main ideas behind the proof of Robust Jung Theorem. We begin by recalling its statement:

Theorem 5 restatement. *Let $d \geq 1$ and let $S \subset \mathbb{R}^d$ be a compact set satisfying*

$$\mathrm{diam}(S) \leq 1 + \beta \qquad \textit{and} \qquad \mathtt{rad}(S) \geq \mathrm{Jung}_d,$$

for some $0 \leq \beta \leq \beta_d$, where $\beta_d > 0$ depends only on d. Then, there exists a set of points

$$\Delta = \{x_0, x_1, \ldots, x_d\} \subset \mathbb{R}^d$$

forming the vertex set of a regular simplex of edge length 1 such that every witness $W \subseteq S$ is $C_d\beta$-close to Δ, for a constant $C_d > 0$ depending only on d.

In the extremal case $\beta = 0$, the theorem yields essentially *two* conclusions: first, every witness coincides with the vertex set of a regular simplex, and second, this witness is unique. The first conclusion follows directly from the proof of Jung's theorem: after translating so that the center of the minimum enclosing ball is at the origin, equality in Jung's inequality implies $\sum_{i=0}^{d} x_i = 0$, and a straightforward calculation then shows that the points form the vertex set of a regular simplex.

The second statement is more delicate. Observe that any witness must lie on the boundary of the minimum enclosing ball of S. Fix a witness $\{x_0, \ldots, x_d\}$, and suppose that y is another point of S lying on the same sphere. We show that y coincides with one of the x_i. The argument relies on the fact that the vertices of a centered regular simplex are in isotropic position: for every $y \in \mathbb{R}^d$,

$$\|y\|^2 \;=\; \tfrac{1}{2}\sum_{i=0}^{d}\langle y, x_i\rangle^2. \tag{4}$$

This identity captures the rigidity and symmetry of the regular simplex, and we use it repeatedly in the proof of Robust Jung Theorem as well. Combining Equation (4) with the assumption that y lies on the same sphere as the x_i, a short linear-algebraic argument yields the desired conclusion. The formal proof appears in Lemma 27.

The proof of Robust Jung Theorem (Theorem 5) follows the same outline, but in a robust form. As before, the main difficulty lies in the second step. First, we show that any witness is $O_d(\beta)$-close to the vertex set of some regular simplex. Second, we show that if the union of two regular simplices has diameter at most $1 + \beta$ (with β sufficiently small), then the simplices must be $O_d(\beta)$-close.

Step 1: Any witness is close to a regular simplex. We first show, by a mild modification of the proof of Jung's theorem, that every edge of a witness W has length $1 + O_d(\beta)$. We then construct a regular simplex Δ by induction on the dimension d. For $n \le d$, let $\Delta' \subset \mathbb{R}^{n-1}$ be an $(n-1)$-dimensional regular simplex, and suppose a point $y \in \mathbb{R}^n$ satisfies $\mathrm{dist}(y, z) = 1 + O_d(\beta)$ for all vertices $z \in \Delta'$. A quantitative inverse function theorem implies that y must be $O_d(\beta)$-close to some point x that completes Δ' to an n-dimensional regular simplex. Iterating this construction yields the desired simplex Δ and proves that W is $O_d(\beta)$-close to Δ. Step 1 is proved formally in Appendix D.

Step 2: Clustering of witnesses. Step 1 allows us to replace each witness set W by the vertex set of a regular simplex Δ of edge length 1, and to exploit the rigidity of regular simplices. Let $\mathcal{X}$ denote the family of all regular simplices of edge length 1. The notion of r-closeness can be expressed on $\mathcal{X}$ via the Hausdorff distance: for $\Delta, \Delta' \in \mathcal{X}$, the value $\mathrm{dist}_H(\Delta, \Delta')$ is the smallest r such that Δ and Δ' are r-close.

We also introduce a second function on $\mathcal{X} \times \mathcal{X}$, the *excess cross-diameter*

$$D_{\mathrm{diam}}(\Delta, \Delta') \;:=\; \mathrm{diam}(\Delta \cup \Delta') - 1.$$

This quantity is symmetric and vanishes if and only if $\Delta = \Delta'$. We show that, when restricted to pairs of sufficiently close simplices, D_{diam} is bilipschitz equivalent to the Hausdorff distance. However, we do not know whether D_{diam} satisfies the triangle inequality; in particular, local bilipschitz equivalence to a metric does not imply that D_{diam} itself is a metric on $\mathcal{X}$.

The inequality $\mathrm{dist}_H(\Delta, \Delta') \gtrsim D_{\mathrm{diam}}(\Delta, \Delta')$ is immediate from the definitions. Indeed, if the Hausdorff distance between Δ and Δ' is r, then Δ' lies inside the r-neighborhood of Δ. Consequently, the union $\Delta \cup \Delta'$ cannot have diameter bigger than $1 + 2r$. The main content of Step 2 is the reverse inequality. We show that there exists $\beta_d > 0$ such that for all distinct $\Delta, \Delta' \in \mathcal{X}$ with $D_{\mathrm{diam}}(\Delta, \Delta') \leq \beta_d$,

$$c_d \;\leq\; \frac{\mathrm{dist}_H(\Delta, \Delta')}{D_{\mathrm{diam}}(\Delta, \Delta')} \;\leq\; C_d, \tag{5}$$

where constants c_d, C_d depend only on the dimension d. To prove this, we work with the Euclidean motion group $E(d)$, fix a reference simplex Δ, and analyze motions $T \in E(d)$ near the identity. We show that sufficiently small motions, in the sense that they move every point by at most α_d for a dimension-dependent constant α_d, satisfy (5) for Δ and $T(\Delta)$. This is proved in Lemma 32 using Lie group–algebra techniques.

Finally, a compactness argument shows that there exists $\beta_d > 0$ such that if $D_{\mathrm{diam}}(\Delta, \Delta')$ is at most β_d, then there exists a motion $T \in E(d)$ which is α_d-small and satisfies $\Delta' = T(\Delta)$. Applying the local estimate above to this motion concludes the bilipschitz comparison between dist_H and D_{diam}, and hence the clustering of witnesses. See Appendix E.

3.3. Dependence on the dimension: Main Theorem 3

We briefly outline the proof of Main Theorem 3. The theorem shows a sharp transition in the query budget as a function of the dimension: for subexponential budgets $T(d) = 2^{o(d)}$ the excess error diverges as $d \to \infty$, while for superexponential budgets $T(d) = 2^{\omega(d)}$ the excess error vanishes.

The argument in both directions relies on the standard comparison between the diameter and the Chebyshev radius of the feasible region:

$$\tfrac{1}{2}\,\mathrm{diam}(\Phi_T) \;\leq\; \mathrm{rad}(\Phi_T) \;\leq\; \mathrm{Jung}_d \cdot \mathrm{diam}(\Phi_T).$$

Lower bound (subexponential budgets). Assume $T(d) = 2^{o(d)}$. We construct an adversary which ensures that the feasible region has large diameter. Our adversary is a continuous analogue of the one constructed by Dinur and Nissim (2003).

The adversary simply answers 0 to all queries. In order to analyze the feasible region, it would be convenient to take the noise rate to be $\delta = 2\sqrt{\ln T}$. Let X, Y be two independent standard Gaussians. We will show that with positive probability, both X and Y belong to the feasible region, and $|X - Y| = \Omega(\sqrt{d})$.

Consequently, the diameter of the feasible region is $\Omega(\sqrt{d})$, hence so is its Chebyshev radius.

Let v_t be the direction queried at the t'th step. Then $\langle v_t, X\rangle$ is a standard Gaussian, hence $\Pr[\langle v_t, X\rangle \geq \delta] \leq \frac{1}{T^2}$ by the standard Gaussian tail bound. Consequently, X and Y are both feasible with probability at least $1 - \frac{2}{T}$.

On the other hand, $X - Y$ is a d-dimensional Gaussian with variance 2, and so its expected norm is $\Omega(\sqrt{d})$. Standard anti-concentration bounds show that this holds even conditioned on X and Y being both feasible, completing the proof.

Upper bound (superexponential budgets). When $T(d) = 2^{\omega(d)}$, the preprocessing step of querying a dense net of directions already suffices. A superexponential query budget $T(d)$ is enough to query a set S of directions such that any $v \in S^{d-1}$ is $o_d(1)$-close to some direction in S, which forces $\mathrm{diam}(\Phi_T) = 2\delta + o_d(1)$. Applying Jung's theorem then gives $\mathrm{rad}(\Phi_T) \leq \mathrm{Jung}_d \cdot 2\delta + o_d(1)$, so the excess error vanishes.

Both proofs appear in Appendix B.4.

4. Open directions

We end with two variants of the reconstruction problem.

One natural variant is the nonadaptive linear reconstruction game, in which all query directions are chosen before any answers are observed. The proof of the doubly-exponential upper bound in Main Theorem 2 relies on adaptivity: the refinement phase queries directions parallel to the edges of a simplex extracted from the current feasible region. Related reconstruction and query-release problems are studied extensively in the privacy literature (Dinur and Nissim, 2003; Blum et al., 2013), but those works primarily concern discrete databases, subset/counting queries, and private release of answers to query classes. To the best of our knowledge, the corresponding minimax problem for nonadaptive approximate linear queries in $\mathbb{R}^d$ has not been characterized.

A second, closely related question is how the model should be extended when some answers are allowed to violate the δ-accuracy promise. In the present model, every answer is consistent with the unknown point x^*, and hence x^* remains inside the feasible region throughout the interaction; this invariant is central to our analysis, and once it fails the geometric arguments used here do not apply directly. In related settings, several forms of unreliable answers have been studied, including arbitrary corruptions in reconstruction and decoding problems in the privacy literature (Dwork et al., 2007), and missing, faulty, or maliciously corrupted answers in learning from membership queries (Angluin and Slonim, 1994; Angluin et al., 1997). Identifying the right robustness notions for the linear reconstruction game, and determining the corresponding optimal error and rates of convergence, remain open.

Acknowledgments

We thank the anonymous reviewers and the meta-reviewer for their careful reading and helpful comments, which improved the presentation of the paper. We are especially grateful to the meta-reviewer for suggesting the nonadaptive and corrupted-answer variants discussed in Section 4.

Shay Moran is a Robert J. Shillman Fellow and acknowledges support by Israel PBC-VATAT, by the Technion Center for Machine Learning and Intelligent Systems (MLIS), and by the European Union (ERC, GENERALIZATION, 101039692).

Yuval Filmus was supported by ISF grant 507/24.

Elizaveta Nesterova was supported by the Milner Fellowship for Ph.D. Students and by the European Union (ERC, GENERALIZATION, 101039692).

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Dana Angluin and Donna K. Slonim. Randomly fallible teachers: Learning monotone DNF with an incomplete membership oracle. *Mach. Learn.*, 14(1):7–26, 1994. doi: 10.1007/BF00993160. URL https://doi.org/10.1007/BF00993160.

Dana Angluin, Martins Krikis, Robert H. Sloan, and György Turán. Malicious omissions and errors in answers to membership queries. *Mach. Learn.*, 28(2-3):211–255, 1997. doi: 10.1023/A:1007311411259. URL https://doi.org/10.1023/A:1007311411259.

Ery Arias-Castro, Emmanuel J. Candès, and Mark A. Davenport. On the fundamental limits of adaptive sensing. *IEEE Transactions on Information Theory*, 59(1):472–481, 2013. doi: 10.1109/TIT.2012.2213610.

Raef Bassily, Kobbi Nissim, Adam D. Smith, Thomas Steinke, Uri Stemmer, and Jonathan R. Ullman. Algorithmic stability for adaptive data analysis. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 1046–1059. ACM, 2016. doi: 10.1145/2897518.2897566. URL https://doi.org/10.1145/2897518.2897566.

Avrim Blum, Katrina Ligett, and Aaron Roth. A learning theory approach to noninteractive database privacy. *Journal of the ACM (JACM)*, 60(2):1–25, 2013.

Leonard M. Blumenthal. *Theory and Applications of Distance Geometry*. Oxford University Press, Oxford, 2 edition, 1970. Originally published 1953.

Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In Frank Neven, Catriel Beeri, and Tova Milo, editors, *Proceedings of the Twenty-Second ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 9-12, 2003, San Diego, CA, USA*, pages 202–210. ACM, 2003. doi: 10.1145/773153.773173. URL https://doi.org/10.1145/773153.773173.

Cynthia Dwork, Frank McSherry, and Kunal Talwar. The price of privacy and the limits of lp decoding. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, page 85–94, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595936318. doi: 10.1145/1250790.1250804. URL https://doi.org/10.1145/1250790.1250804.

Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Roth. Generalization in adaptive data analysis and holdout reuse. In Corinna Cortes, Neil D. Lawrence, Daniel D. Lee, Masashi Sugiyama, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, pages 2350–2358, 2015a. URL https://proceedings.neurips.cc/paper/2015/hash/bad5f33780c42f2588878a9d07405083-Abstract.html.

Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Leon Roth. Preserving statistical validity in adaptive data analysis. In Rocco A. Servedio and Ronitt Rubinfeld, editors, *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 117–126. ACM, 2015b. doi: 10.1145/2746539.2746580. URL https://doi.org/10.1145/2746539.2746580.

Peter M. Gruber. *Convex and Discrete Geometry*, volume 336 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2007.

Jarvis Haupt, Rui Castro, and Robert D. Nowak. Adaptive sensing for sparse signal recovery. In *IEEE Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE)*, pages 702–707, 2009.

Alejandro Illanes and Sam B. Nadler. *Hyperspaces: Fundamentals and Recent Advances*, volume 216 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, New York, 1999.

Shengqiao Li. Concise formulas for the area and volume of a hyperspherical cap. *Asian Journal of Mathematics & Statistics*, 4(1):66–70, 2011. ISSN 1994-5418. doi: 10.3923/ajms.2011.66.70.

Shay Moran and Elizaveta Nesterova. Reconstruction and secrecy under approximate distance queries. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

L. Reyzin. Statistical queries and statistical algorithms: Foundations and applications. *ArXiv*, abs/2004.00557, 2020. URL https://api.semanticscholar.org/CorpusID:214743339.

Appendix A. Model and Definitions

We now give the formal definition of the linear reconstruction game, the associated minimax optimal reconstruction error, and supporting lemmas. Throughout this section we restrict attention to reconstructors with point-valued outputs, i.e., we do not consider improper reconstructors. The improper model is deferred to Appendix C.

Definition 6 (Linear reconstruction game and loss) *Fix an ambient space $\mathbb{R}^d$ and a noise level $\delta > 0$. The linear reconstruction game is played between an* adversary *and a* reconstructor *as follows.*

1. *At the beginning of the game, the adversary chooses a secret point $x^* \in \mathbb{R}^d$.*
2. *The interaction lasts for T rounds. In round $t \in [T]$:*
   (a) *The reconstructor chooses a query direction $v_t \in \mathbb{R}^d$ with $\|v_t\|_2 = 1$, possibly adaptively as a function of previous answers.*
   (b) *The adversary returns a noisy response $r_t \in \mathbb{R}$ satisfying*
   $$\big|r_t - \langle x^*, v_t\rangle\big| \le \delta.$$
3. *At the end of the interaction, the reconstructor outputs an estimate $\hat{x}_T \in \mathbb{R}^d$.*

We measure the error of the estimate $\hat{x}_T$ with respect to the secret x^* by ℓ_2-distance $\|\hat{x}_T - x^*\|_2$.

Recall that the *optimal reconstruction error* is the smallest worst-case error that a reconstructor can guarantee after T rounds of interaction. Formally,

Definition 7 (Optimal reconstruction error) *Fix a dimension $d \ge 1$ and work in $\mathbb{R}^d$. Let RC be the set of reconstruction strategies for T rounds. We define the* optimal reconstruction error *by*

$$\mathrm{OPT}_d(T,\delta) := \inf_{\mathcal{R}\in\mathsf{RC}} \sup_{x^*\in\mathbb{R}^d} \sup_{\mathcal{A}\in\mathsf{ADV}(x^*)} \|\hat{x}_T - x^*\|_2.$$

Here $\mathsf{ADV}(x^*)$ denotes the set of adversary response strategies consistent with the secret point x^*, and $\hat{x}_T$ is the final output of $\mathcal{R}$. All strategies are assumed deterministic.

The first basic properties of the optimal reconstruction error are nonnegativity and monotonicity in the number of rounds. The argument is standard and does not use any properties of the linear reconstruction game beyond the fact that the loss is nonnegative.

Lemma 8 (Monotonicity and nonnegativity in T) *For every $d \geq 1$ and noise level $\delta > 0$, the sequence*
$$T \mapsto \mathrm{OPT}_d(T, \delta)$$
is nonincreasing and nonnegative.

Proof Nonnegativity is immediate from the definitions since the loss is nonnegative, and taking a supremum over adversaries and an infimum over reconstructors preserves nonnegativity, hence $\mathrm{OPT}_d(T, \delta) \geq 0$ for all T.

We next prove monotonicity. Fix $T \geq 0$ and $\alpha > 0$. By definition of $\mathrm{OPT}_d(T, \delta)$, there exists a reconstructor $\mathcal{R}$ that guarantees worst-case error at most $\mathrm{OPT}_d(T, \delta) + \alpha$ after T rounds.

Define a reconstructor $\mathcal{R}'$ for $T + 1$ rounds as follows. $\mathcal{R}'$ simulates $\mathcal{R}$ for the first T rounds, then issues an arbitrary query at round $T + 1$, ignores the response, and outputs exactly the same final estimate as $\mathcal{R}$ would after T rounds.

For any adversary, the error of $\mathcal{R}'$ coincides with the error of $\mathcal{R}$. Therefore, $\mathcal{R}'$ guarantees worst-case error at most $\mathrm{OPT}_d(T, \delta) + \alpha$ in $T + 1$ rounds, implying
$$\mathrm{OPT}_d(T + 1, \delta) \leq \mathrm{OPT}_d(T, \delta) + \alpha.$$
Since this holds for every $\alpha > 0$, we conclude that $\mathrm{OPT}_d(T + 1, \delta) \leq \mathrm{OPT}_d(T, \delta)$. ■

We are now ready to prove a simple but useful scaling property in the noise level.

Lemma 9 (Scaling in the noise level) *For every $T \in \mathbb{N}$ and $\delta > 0$, the optimal error scales linearly with δ:*
$$\mathrm{OPT}_d(T, \delta) \;=\; \delta\, \mathrm{OPT}_d(T, 1).$$

Proof Fix $T \in \mathbb{N}$ and $\delta > 0$. Since all strategies are deterministic, we can couple a game at noise level 1 with a game at noise level δ by a simple rescaling.

Take any adversary strategy $\mathcal{A}$ for noise level 1. It is determined by a secret point $x^* \in \mathbb{R}^d$ and a deterministic rule for answering queries. Define a new adversary $\mathcal{A}_\delta$ for noise level δ as follows. Its secret is δx^*. On each query v_t, it computes the answer that $\mathcal{A}$ would return to v_t at noise level 1, and then returns δ times this value.

Take any reconstructor strategy $\mathcal{R}$ for noise level 1. Define a new reconstructor $\mathcal{R}_\delta$ for noise level δ as follows. It issues exactly the same queries v_t that $\mathcal{R}$ would issue. Whenever it receives an answer r_t, it rescales it to r_t/δ and feeds this value to $\mathcal{R}$. In the end of the interaction if $\mathcal{R}$ outputs $\hat{x}_{\mathcal{R}}$, then $\mathcal{R}_\delta$ outputs $\delta \hat{x}_{\mathcal{R}}$.

These transformations are invertible by applying the same construction with factor $1/\delta$. Now run $\mathcal{R}$ against $\mathcal{A}$ at noise level 1, and run $\mathcal{R}_\delta$ against $\mathcal{A}_\delta$ at noise level δ. By construction, the query sequence is identical, and the answers in the second interaction are exactly δ times the answers in the first. Therefore the final outputs are also scaled by δ.

This gives
$$\|\delta \hat{x}_{\mathcal{R}} - \delta x^*\| = \delta \|\hat{x}_{\mathcal{R}} - x^*\|.$$

Taking the supremum over adversaries and the infimum over reconstructors yields

$$\mathrm{OPT}_d(T,\delta) = \delta\,\mathrm{OPT}_d(T,1),$$

since the above transformations between noise levels 1 and δ are bijections on both the adversary and reconstructor strategy sets. ■

Finally, we recall the definition of the feasible region (see eq. (1)), the core geometric object of the game.

Definition 10 (Feasible region) *Given a transcript $(v_1, r_1), \ldots, (v_T, r_T)$ of query directions and responses observed over T rounds at noise level δ, the feasible region after T rounds is*

$$\Phi_T := \Big\{ x \in \mathbb{R}^d : |\langle x, v_t\rangle - r_t| \le \delta \text{ for all } t \in [T] \Big\}.$$

The feasible region records exactly the information available to the reconstructor after the interaction. Under this viewpoint, we can simplify the adversary model without changing the optimal reconstruction error.

Lemma 11 (A posteriori adversaries) *In the definition of $\mathrm{OPT}_d(T,\delta)$, we may replace the usual adversary (which fixes x^* in advance) by an adversary that only maintains feasibility: it chooses replies $r_1, \ldots, r_T$ online so that $\Phi_T \neq \varnothing$, and after observing the reconstructor's output it selects any $x^* \in \Phi_T$. For deterministic reconstructors, these two models yield the same minimax value.*

Proof Any *a priori* adversary that fixes a secret $x^* \in \mathbb{R}^d$ in advance and answers within noise δ produces a feasible transcript, hence is a special case of the feasibility-maintaining adversary.

Conversely, fix a deterministic reconstructor $\mathcal{R}$ and consider any feasible transcript $(v_t, r_t)_{t=1}^T$ produced against $\mathcal{R}$, with feasible region $\Phi_T \neq \varnothing$. Choose an arbitrary $x^* \in \Phi_T$. Then, by definition of Φ_T,

$$|r_t - \langle x^*, v_t\rangle| \le \delta \qquad \text{for all } t \in [T].$$

Define an *a priori* adversary with secret x^* as follows: on the query sequence generated by $\mathcal{R}$, it returns the prescribed replies $r_1, \ldots, r_T$. If $\mathcal{R}$ ever deviates from this query sequence, the adversary answers thereafter truthfully with $\langle x^*, v\rangle$ on any query direction $v \in S^{d-1}$. This strategy always respects the noise constraint, and it reproduces the same transcript against $\mathcal{R}$.

Therefore, for deterministic reconstructors, allowing an a priori versus an a posteriori adversary does not change the optimal reconstruction error. ■

Finally we now connect the reconstruction loss to a purely geometric quantity of the feasible region. Recall that the *Chebyshev radius* of a set $S \subset \mathbb{R}^d$ is the radius of its smallest enclosing Euclidean ball,

$$\mathtt{rad}(S) := \inf_{c \in \mathbb{R}^d} \sup_{y \in S} \|c - y\|_2. \tag{6}$$

Any minimizer c is called a *Chebyshev center* of S. (When S is bounded and nonempty, a Chebyshev center exists.)

Lemma 12 (OPT via the feasible region) *Fix a transcript with nonempty feasible region Φ_T. For any reconstructor output $\hat{x} \in \mathbb{R}^d$,*

$$\sup_{x^* \in \Phi_T} \|\hat{x} - x^*\|_2 \;\geq\; \mathtt{rad}(\Phi_T).$$

Moreover, equality holds when $\hat{x}$ is a Chebyshev center of Φ_T. Consequently,

$$\mathrm{OPT}_d(T, \delta) = \inf_{\substack{\mathcal{R} \text{ a reconstructor strategy} \\ \text{at noise level } \delta}} \; \sup_{\substack{\text{feasible transcripts} \\ \text{generated against } \mathcal{R}}} \mathtt{rad}(\Phi_T).$$

Proof By Lemma 11, after the transcript is fixed the adversary may choose any $x^* \in \Phi_T$. Thus the worst-case loss of an output $\hat{x}$ equals $\sup_{y \in \Phi_T} \|\hat{x} - y\|_2$, and minimizing this quantity over $\hat{x}$ is exactly the definition of $\mathtt{rad}(\Phi_T)$ in (6). ■

Appendix B. Proofs

B.1. Querying a Covering of Directions

Before proving our main results, we introduce several technical tools that will be used repeatedly throughout the proofs.

We begin by defining a useful angular metric on the unit sphere. Let $S^{d-1} \subset \mathbb{R}^d$. The *angular distance* on S^{d-1} is defined by

$$\rho(x, y) \;:=\; \angle(x, y) \;=\; \arccos\big(\langle x, y\rangle\big) \in [0, \pi], \qquad x, y \in S^{d-1}. \tag{7}$$

Given $\alpha > 0$, a subset $V \subset S^{d-1}$ is called an α-covering in (S^{d-1}, ρ) if for every $v \in S^{d-1}$ there exists $u \in V$ such that

$$\rho(v, u) \leq \alpha.$$

Definition 13 *The (angular)* covering number $\mathcal{N}_\angle(S^{d-1}, \alpha)$ *is the minimal cardinality of an α-covering in (S^{d-1}, ρ). Since the sphere S^{d-1} is compact, this covering number is finite for every $\alpha > 0$.*

In Appendix F we prove the following useful estimate.

Lemma 14 *For any $0 < \alpha < \frac{\pi}{2}$, the covering number $\mathcal{N}_\angle(S^{d-1}, \alpha)$ satisfies the following bounds:*

$$\sqrt{d}\, \alpha^{-(d-1)} \;<\; \mathcal{N}_\angle(S^{d-1}, \alpha) \;<\; 2^{2d}\, \alpha^{-(d-1)}.$$

The following lemma is a straightforward but useful observation that will be used repeatedly in the proofs of the main theorems.

Lemma 15 *Consider noise scale $\delta > 0$. Assume that the reconstructor queries an α-covering of S^{d-1} in the angular metric, for some $0 < \alpha < \pi/2$, that is, it makes at least $\mathcal{N}_\angle(S^{d-1}, \alpha)$ queries (see Definition 13 for the covering number).*

Then, for any adversary the diameter of the resulting feasible region is bounded by

$$\operatorname{diam} \Phi_T \;\leq\; \frac{2\delta}{\cos\alpha},$$

and hence

$$\sup_{\textit{adversaries}} \operatorname{diam}\Phi_T \;\leq\; \frac{2\delta}{\cos\alpha}.$$

Proof Let $\{v_1, \ldots, v_T\} \subset S^{d-1}$ be an α-covering of the unit sphere, that is, for every $v \in S^{d-1}$ there exists v_i such that

$$\langle v, v_i \rangle \;\geq\; \cos\alpha,$$

equivalently, the angle between v and v_i is at most α. Let Φ_T be the resulting feasible region, which depends on the adversary holding the secret point x^*. Take arbitrary $x, y \in \Phi_T$ and set

$$v := \frac{x-y}{\|x-y\|}.$$

By the α-covering property, there exists v_i with $\langle v, v_i \rangle \geq \cos\alpha$. Since both x and y are consistent with the answer in direction v_i, we have

$$|\langle x, v_i \rangle - r_i| \leq \delta, \qquad |\langle y, v_i \rangle - r_i| \leq \delta,$$

and hence

$$|\langle x, v_i \rangle - \langle y, v_i \rangle| \;\leq\; 2\delta.$$

On the other hand,

$$\langle x, v_i \rangle - \langle y, v_i \rangle = \langle x-y, v_i \rangle = \|x-y\| \, \langle v, v_i \rangle \geq \|x-y\| \, \cos\alpha.$$

Combining the two inequalities yields

$$\|x-y\| \;\leq\; \frac{2\delta}{\cos\alpha}.$$

Since $x, y \in \Phi_T$ were arbitrary, we conclude that

$$\operatorname{diam}(\Phi_T) \;\leq\; \frac{2\delta}{\cos\alpha}. \tag{8}$$

■

B.2. Proof of Main Theorem 1

In this section we determine the limiting optimal reconstruction error in the linear reconstruction game on $\mathbb{R}^d$ with normalized noise, as the number of interaction rounds T tends to infinity. We will also need the following theorem, which links diameter and radius in Euclidean space. Recall Jung's theorem:

Theorem 4 restatement. *For every bounded set $S \subseteq \mathbb{R}^d$ with diameter D, its Chebyshev radius satisfies*

$$\mathtt{rad}(S) \leq \sqrt{\tfrac{d}{2(d+1)}}\, D;$$

see, e.g., Blumenthal (1970). The factor $\sqrt{\frac{d}{2(d+1)}}$ *is known as* Jung's constant*; we denote it by* Jung_d.

Now we are ready to prove Main Theorem 1. Recall it's statement:

Main Theorem 1 restatement. *Consider the linear reconstruction game on* $\mathbb{R}^d$*, where* $d \geq 1$*, with noise level* $\delta > 0$*. Let* $\mathrm{OPT}_d(T, \delta)$ *denote the optimal reconstruction errors after* T *rounds, as defined in Definition 7. Then the limit in* T *exists and satisfies*

$$\mathrm{OPT}_d(\infty, \delta) := \lim_{T \to \infty} \mathrm{OPT}_d(T, \delta) = \sqrt{\frac{2d}{d+1}}\delta.$$

Proof Existence of the limit follows from the monotonicity and nonnegativity of $T \mapsto \mathrm{OPT}_d(T, \delta)$ (Lemma 8). To identify its value, we give an upper bound and a matching lower bound.

Upper bound. The main ingredient is Lemma 15, which provides a reconstruction strategy based on querying a sufficiently fine angular covering of the sphere.

Let $n \in \mathbb{N}$ and let $T_n := \mathcal{N}_\angle(S^{d-1}, 1/n)$ be the size of a $1/n$-angular covering of S^{d-1} (see Definition 13). Applying Lemma 15, after T_n queries the feasible region Φ_{T_n} satisfies

$$\mathrm{diam}(\Phi_{T_d}) \;\leq\; \frac{2}{\cos(1/n)}.$$

Jung's theorem (Theorem 4) states that any subset of $\mathbb{R}^d$ of diameter D has Chebyshev radius at most $\sqrt{\frac{2d}{d+1}} \cdot D/2$. Applying this to Φ_{T_n} gives

$$\mathrm{OPT}_d(T_n, \delta) \;\leq\; \sqrt{\frac{2d}{d+1}} \cdot \frac{\delta}{\cos(1/n)}.$$

Letting $n \to \infty$ yields

$$\mathrm{OPT}_d(\infty) \;\leq\; \sqrt{\frac{2d}{d+1}}\delta.$$

Lower bound. We now prove matching lower bound. Let $S \subset \mathbb{R}^d$ be the vertex set of a regular simplex of diameter 2δ. When the reconstructor issues a query direction $v_t \in S^{d-1}$, the adversary projects S onto the line spanned by v_t and returns the midpoint of the resulting interval. Formally, for $v \in S^{d-1}$ define

$$m_v(S) := \inf_{s \in S} \langle v, s \rangle, \qquad M_v(S) := \sup_{s \in S} \langle v, s \rangle,$$

and set

$$r_t := \frac{M_{v_t}(S) + m_{v_t}(S)}{2}.$$

Then for every $x \in S$,

$$|\langle x, v_t \rangle - r_t| \leq \delta,$$

and hence $S \subseteq \Phi_t$ for all t.

Since the adversary ensures $S \subseteq \Phi_T$ for every transcript, we have

$$\mathtt{rad}(\Phi_T) \geq \mathtt{rad}(S) = 2\mathrm{Jung}_d\,\delta.$$

By Lemma 12, the worst-case loss of a reconstructor for feasible region Φ_T equals $\texttt{rad}(\Phi_T)$. Therefore, for every T,

$$\mathrm{OPT}_d(T,\delta) \geq 2\mathrm{Jung}_d\,\delta = \sqrt{\frac{2d}{d+1}}\,\delta,$$

and hence

$$\mathrm{OPT}_d(\infty,\delta) \geq \sqrt{\tfrac{2d}{d+1}}\,\delta.$$

Combining the upper and lower bounds completes the proof. ∎

B.3. Proof of Main Theorem 2

The goal of this section is to establish the doubly-exponential convergence rate of the excess error. We restate the main theorem with explicit constants (rather than the $\Theta_d(\cdot)$ notation used in the introduction).

Recall that the excess error is defined by

$$\mathsf{ExcessErr}_d(T,\delta) := \mathrm{OPT}_d(T,\delta) - \mathrm{OPT}_d(\infty,\delta), \tag{9}$$

where $\mathrm{OPT}_d(T,\delta)$ is the optimal reconstruction error (see Definition 7) and $\mathrm{OPT}_d(\infty,\delta)$ is the limiting value of the optimal error (see Main Theorem 1).

Main Theorem 2 restatement. *There exist constants $a_d, A_d > 0$ and a threshold T'_d, depending only on the dimension d, such that for all $T > T'_d$,*

$$\delta \cdot 2^{-2^{A_d \cdot T}} \leq \mathsf{ExcessErr}_d(T,\delta) \leq \delta \cdot 2^{-2^{a_d \cdot T}}.$$

For normalization, we assume $\delta = \frac{1}{2}$. By the linear scaling in the noise level (Lemma 9),

$$\mathrm{OPT}_d(T,\delta) = \delta\,\mathrm{OPT}_d(T,1) \qquad \text{and} \qquad \mathrm{OPT}_d(\infty,\delta) = \delta\,\mathrm{OPT}_d(\infty,1).$$

Therefore,

$$\mathrm{OPT}_d(T,\tfrac{1}{2}) - \mathrm{OPT}_d(\infty,\tfrac{1}{2}) = \frac{1}{2\delta}\Big(\mathrm{OPT}_d(T,\delta) - \mathrm{OPT}_d(\infty,\delta)\Big),$$

and it suffices to prove doubly exponential upper and lower bounds on $\mathrm{OPT}_d(T,\frac{1}{2}) - \mathrm{OPT}_d(\infty,\frac{1}{2})$.

We begin with the upper bound and then turn to the lower bound.

B.3.1. UPPER BOUND IN MAIN THEOREM 2

To provide an upper bound on excess error, we recall the core concept introduced in Section 3.

Definition 16 (Witness) *Let $S \subset \mathbb{R}^d$. A* witness *for S is a subset $W \subseteq S$ of cardinality $|W| = d+1$ such that $\texttt{rad}(W) \geq \mathrm{Jung}_d$, where $\mathrm{Jung}_d = \sqrt{\frac{d}{2(d+1)}}$ is the Jung constant (see Theorem 4). If $\texttt{rad}(S) \geq \mathrm{Jung}_d$, then S admits a witness by Lemma 22.*

We can then introduce a witness core of a subset $S \subset \mathbb{R}^d$.

$$\mathrm{Witness}(S) \ := \bigcup_{\substack{W \subset S: \\ W \text{ is a witness}}} W. \tag{10}$$

In Appendix D we prove the main technical tool used in the upper bound, namely the *Robust Jung Theorem*. Using the notion of the witness core, it can be stated in the following form, which is more convenient for our proof.

Theorem 5 restatement. *Let $d \geq 1$ and let $S \subset \mathbb{R}^d$ be a compact set satisfying*

$$\mathrm{diam}(\mathrm{Witness}(S)) \leq 1 + \beta \qquad \textit{and} \qquad \mathtt{rad}(S) \geq \mathrm{Jung}_d,$$

for some $0 \leq \beta \leq \beta_d$, where $\beta_d > 0$ depends only on d. Then there exists a set of points

$$\Delta = \{x_0, x_1, \ldots, x_d\} \subset \mathbb{R}^d$$

forming the vertex set of a regular simplex of edge length 1 such that

$$\mathrm{Witness}(S) \subset \Delta + B(C_d \beta),$$

where $\Delta + B(r)$ denotes the Minkowski sum of Δ with the Euclidean ball of radius $r > 0$.

Now we are ready to present the proof of Main Theorem 2 (Upper bounds).

Main Theorem 2 restatement. *(Upper bound) There exist constants $a_d > 0$ and a threshold T'_d, depending only on the dimension d, such that for all $T > T'_d$,*

$$\mathsf{ExcessErr}_d(T, \tfrac{1}{2}) \leq 2^{-2^{a_d \cdot T}}.$$

Proof Denote by $\beta_d > 0$ the constant from the Robust Jung Theorem (Theorem 5), and define

$$\alpha_d \ := \ \arccos\left(\frac{1}{1+\beta_d}\right).$$

Denote by $T_0 := N(S^{d-1}, \alpha_d)$ the covering number of the sphere S^{d-1} with angular radius α_d (see Definition 13).

We construct, by induction on t, a reconstruction strategy with the following guarantee. After

$$T_0 + \binom{d+1}{2} \cdot t$$

queries, the witness core at time $T_0 + \binom{d+1}{2} t$ (which for convenience we denote by $\mathrm{Witness}_t$) satisfies

$$\mathrm{Witness}_t \subseteq \Delta_t + B(\gamma_t),$$

where Δ_t is the vertex set of a regular simplex of edge length 1, and $(\gamma_t)_{t \geq 0}$ is a sequence obeying the recursion

$$\gamma_0 = C_d \cdot \beta_d, \qquad \gamma_t = A_d \cdot \gamma_{t-1}^2,$$

for a constant $A_d > 0$ depending only on the dimension d such that $C_d \cdot \beta_d < 1/A_d$. Here C_d is the constant from the Robust Jung Theorem (Theorem 5).

Since

$$\mathtt{rad}(\Phi_t) \ = \ \mathtt{rad}(\mathrm{Witness}_t) \ \leq \ \mathtt{rad}\big(\Delta_t + B(\gamma_t)\big) \ = \ \mathtt{rad}(\Delta_t) + \gamma_t \ = \ \mathrm{Jung}_d + \gamma_t,$$

this claim immediately yields a doubly exponential upper bound on the convergence rate of $\mathsf{ExcessErr}_d(T, \delta)$.

Induction base. The base case essentially follows from Lemma 15 which states that whenever one queries an α_d-covering of a sphere S^{d-1} with respect to the angular metric (see Equation (7)), one has

$$\operatorname{diam}\Phi_{T_0} \leq 2/\cos\alpha_d.$$

Since $\mathrm{Witness}_{T_0} \subseteq \Phi_{T_0}$, we have $\operatorname{diam}(\mathrm{Witness}_{T_0}) \leq \operatorname{diam}(\Phi_{T_0})$. Thus Theorem 5 applies to $S = \Phi_{T_0}$, and yields a regular simplex $\Delta_0 := \{x_0, x_1, \dots, x_d\}$ of edge length 1 such that

$$\mathrm{Witness}_{T_0} \subseteq \Delta_0 + B(C_d\,\beta_d).$$

This establishes the base case.

Induction step. Assume the induction hypothesis at step t:

$$\mathrm{Witness}_t \subseteq \Delta_t + B(\gamma_t),$$

where $\Delta_t = \{x_0, x_1, \dots, x_d\}$. We query the directions $v_{ij} := x_i - x_j$ for all $0 \leq i < j \leq d$.

We claim that

$$\operatorname{diam}(\mathrm{Witness}_{t+1}) \leq 1 + 4\gamma_t^2.$$

First note that $\mathrm{Witness}_{t+1} \subseteq \mathrm{Witness}_t$. Hence any two points in the new witness core lie within distance γ_t of vertices of Δ_t. Let $y_i \in B(x_i, \gamma_t)$ and $y_j \in B(x_j, \gamma_t)$, and assume that $x_i \neq x_j$.

Observe that

$$\sin\big(y_i - y_j,\, x_i - x_j\big) \;\leq\; \frac{2\gamma_t}{1 - 2\gamma_t},$$

which follows from the sine theorem applied to the triangle with vertices $A = 0$, $B = x_i - x_j$, and $C = y_i - y_j$. Indeed,

$$|AB| = 1, \qquad |BC| \leq 2\gamma_t.$$

By the sine theorem,

$$\frac{|AB|}{\sin\angle BCA} = \frac{|BC|}{\sin\angle BAC} \quad\Rightarrow\quad \sin\angle BAC = \frac{|BC|}{|AB|}\cdot\sin\angle BCA \;\leq\; 2\gamma_t.$$

Since $\sin\angle BAC = \sin(y_i - y_j,\, x_i - x_j)$, we obtain

$$\sin(y_i - y_j,\, x_i - x_j) \;\leq\; 2\gamma_t.$$

Therefore, since we queried $v_{ij} = x_i - x_j$ and both points y_i, y_j remain feasible,

$$1 \geq \langle y_i - y_j,\, v_{ij}\rangle = \|y_i - y_j\|_2\sqrt{1 - \sin^2\big(y_i - y_j,\, x_i - x_j\big)} \geq \|y_i - y_j\|_2 \cdot \sqrt{1 - 4\gamma_t^2}.$$

Whenever $\gamma_t \leq \frac{1}{4}$, we have

$$\frac{1}{\sqrt{1 - 4\gamma_t^2}} \leq 1 + 4\gamma_t^2,$$

and hence

$$\|y_i - y_j\| \leq 1 + 4\gamma_t^2.$$

Since y_i, y_j were arbitrary, it follows that

$$\operatorname{diam}(\mathrm{Witness}_{t+1}) \leq 1 + 4\gamma_t^2.$$

Finally, we apply the Robust Jung Theorem (Theorem 5) to $\mathrm{Witness}_{t+1}$. It follows that there exists a regular simplex Δ_{t+1} such that

$$\mathrm{Witness}_{t+1} \subseteq \Delta_{t+1} + B(4C_d\,\gamma_t^2).$$

Setting $\gamma_{t+1} := 4C_d\,\gamma_t^2$ completes the induction step. ■

B.3.2. LOWER BOUND IN MAIN THEOREM 2

We begin with a lemma describing how the adversary can answer the reconstructor's queries so as to maintain $\Delta + B(\alpha)$ inside the feasible region. Here Δ denotes the vertex set of a regular simplex in $\mathbb{R}^d$, and $\Delta + B(\alpha)$ is the Minkowski sum

$$\Delta + B(\alpha) \;=\; \{\, x + y : x \in \Delta,\ \|y\|_2 \le \alpha \,\}.$$

We then use this lemma to derive lower bounds on the excess error.

Lemma 17 (Criteria for the good neighborhood of regular simplex) *Fix $0 < \alpha < \frac{1}{4}$, and let $\Delta = x_0x_1\ldots x_d$ be a regular simplex with edge length 1. For a given direction $v \in \mathbb{R}^d$, without loss of generality choose the edge x_0x_1 of Δ so that*

$$\langle v, x_0 - x_1\rangle \;=\; \max_{i,j}\,\langle v, x_i - x_j\rangle.$$

If $\langle v, x_0 - x_1\rangle \le 1 - 2\alpha$, then, with $r := \langle x_0 + x_1, v\rangle/2$, one has

$$\Delta + B(\alpha) \subset \Phi(\{v, r\}).$$

Proof

First, observe that for any vertex x_i one has

$$\langle x_1 - x_0, v\rangle \;\le\; \langle x_0 + x_1 - 2x_i, v\rangle \;\le\; \langle x_0 - x_1, v\rangle.$$

Indeed, suppose for contradiction that $\langle x_0 + x_1 - 2x_i, v\rangle > \langle x_0 - x_1, v\rangle$. Then $\langle x_1 - x_i, v\rangle > 0$, and hence $\langle x_0 - x_i, v\rangle > \langle x_0 - x_1, v\rangle$, contradicting the maximality of the latter. The other inequality is proved analogously.

Now assume $x \in B(x_i, \alpha)$. We claim that $x \in \Phi(v, r)$. Write $x = x_i + e$ with $\|e\|_2 \le \alpha$. Then

$$\Big\langle \tfrac{x_0+x_1}{2} - x,\, v \Big\rangle \;=\; \Big\langle \tfrac{x_0+x_1}{2} - x_i,\, v \Big\rangle - \langle e, v\rangle \;\le\; \tfrac{1}{2}\langle x_0 - x_1, v\rangle + \alpha \;\le\; \tfrac{1}{2},$$

where the second inequality uses the bound $\langle x_0 + x_1 - 2x_i, v\rangle \le \langle x_0 - x_1, v\rangle$ shown above (and $|\langle e, v\rangle| \le \alpha$), and the third uses the hypothesis $\langle x_0 - x_1, v\rangle \le 1 - 2\alpha$. The opposite inequality, $\langle (x_0 + x_1)/2 - x, v\rangle \ge -\frac{1}{2}$, is proved in the same way. ■

Now we are ready to prove the lower bound. Recall the statement of the theorem:

Main Theorem 2 restatement. *(Lower bound) There exist constant $A_d > 0$ depending only on the dimension d, such that for all $T \ge d$*

$$\mathsf{ExcessErr}_d(T, \tfrac{1}{2}) \ge 2^{-2^{A_d \cdot T}}.$$

Proof It suffices to show that there exists a universal constant $M > 0$ such that, for all sufficiently small $\alpha > 0$, for every regular simplex Δ of edge length 1 and every direction v, there exist a response r and a simplex Δ' with

$$\Delta' + B(\alpha') \subset \Delta + B(\alpha) \qquad \text{and} \qquad \Delta' + B(\alpha') \subset \Phi(v, r),$$

where $\alpha' = M\alpha^2$. We claim that the choice $M = \frac{1}{17}$ is sufficient.

Without loss of generality, assume that the simplex is

$$\Delta = \{0, x_1, \ldots, x_d\}, \qquad \langle v, x_1 \rangle = \max_{i,j} \langle v, x_i - x_j \rangle.$$

If $\langle v, x_1 \rangle \leq 1 - 2\alpha^2/17$, then Lemma 17 directly produces a response r with $\Delta + B(\alpha^2/17) \subset \Phi(v, r)$, and we are done with $\Delta' = \Delta$. Suppose instead that

$$\langle v, x_1 \rangle > 1 - 2\alpha^2/17 \geq 1 - \alpha^2/8 + O(\alpha^4) = \cos(\alpha/2).$$

In this case we rotate Δ slightly so that Lemma 17 becomes applicable to the rotated simplex. The rotation is provided by Lemma 34 (Appendix G): given any direction v and any angle $\theta \leq \pi/18$, it constructs an isometry R_θ such that the rotated simplex $\Delta' := R_\theta \Delta$ satisfies $\langle v, x_1' \rangle \leq \cos\theta$ while still attaining the maximum at the edge $0x_1'$.

Applying Lemma 34 with $\theta = \alpha/2$ produces a rotated simplex Δ' with

$$\langle v, x_1' \rangle \leq \cos(\alpha/2) \leq 1 - 2\alpha^2/17, \qquad \langle v, x_1' \rangle = \max_{i,j} \langle v, x_i' - x_j' \rangle.$$

Lemma 17 then yields a response r such that $\Delta' + B(\alpha^2/17) \subset \Phi(v, r)$.

It remains to verify that $\Delta' + B(\alpha^2/17) \subset \Delta + B(\alpha)$. For this we use Lemma 35 (Appendix G), which reduces this inclusion to a bound on the single quantity $\|x_1 - x_1'\|_2$: namely, it suffices to check that $\|x_1 - x_1'\|_2 \leq \alpha - \alpha^2/17$. Indeed,

$$\|x_1 - x_1'\|_2^2 = 2(1 - \cos(\alpha/2)) = \alpha^2/8 + O(\alpha^4) \leq (\alpha - \tfrac{1}{17}\alpha^2)^2$$

for all sufficiently small α.

Iterating this construction over T rounds yields $\Delta_T + B(\alpha_T) \subset \Phi_T$ with $\alpha_T \geq 2^{-2^{A_d T}}$, and hence $\mathsf{ExcessErr}_d(T, \frac{1}{2}) \geq \mathtt{rad}(\Phi_T) - \mathrm{Jung}_d \geq \alpha_T$. This concludes the proof. ■

B.4. Proof of Main Theorem 3

The goal of this section is to quantify how the query budget $T(d)$ required to control the excess error depends on the dimension d.

Recall that the excess error is

$$\mathsf{ExcessErr}_d(T, \delta) := \mathrm{OPT}_d(T, \delta) - \mathrm{OPT}_d(\infty, \delta),$$

where $\mathrm{OPT}_d(\infty, \delta)$ denotes the limiting optimal error from Main Theorem 1.

We use standard asymptotic notation as $d \to \infty$. For $f, g : \mathbb{N} \to \mathbb{R}_+$, we write $f(d) = o(g(d))$ if $\frac{f(d)}{g(d)} \to 0$, and $f(d) = \omega(g(d))$ if $\frac{f(d)}{g(d)} \to +\infty$.

Main Theorem 3 restatement. *Let $T : \mathbb{N} \to \mathbb{N}$ be a query budget as a function of the dimension d.*

1. *If $\ln T(d) = o(d)$ for all sufficiently large d, then*

$$\mathsf{ExcessErr}_d(T(d), \delta) \xrightarrow{d \to \infty} +\infty.$$

2. *If $\ln T(d) = \omega(d)$ for all sufficiently large d, then*

$$\mathsf{ExcessErr}_d(T(d), \delta) \xrightarrow{d \to \infty} 0.$$

Proof We treat the two regimes separately. The subexponential regime relies on a Gaussian construction reminiscent of Dinur and Nissim (2003); the superexponential regime reduces to the covering-number bound of Lemma 14.

Proof of Item 1 (subexponential budgets). By the scaling property (Lemma 9), we have $\mathrm{OPT}_d(T, \delta) = \delta \mathrm{OPT}_d(T, 1)$. Thus, it suffices to show that

$$\mathrm{OPT}_d(T_d, 1) - \mathrm{OPT}_d(\infty, 1) \xrightarrow{d \to \infty} +\infty.$$

From Main Theorem 1, the asymptotic optimal error is bounded by $\mathrm{OPT}_d(\infty, 1) = \sqrt{\frac{2d}{d+1}} \le \sqrt{2}$. Therefore, we need only show that

$$\mathrm{OPT}_d(T_d, 1) \xrightarrow{d \to \infty} +\infty.$$

To this end, let $\delta_d := 2\sqrt{\ln T_d}$. We will prove that for all sufficiently large d,

$$\mathrm{OPT}_d(T_d, \delta_d) \ge \mathrm{Jung}_d \cdot \sqrt{d}. \tag{11}$$

Assuming (11), the scaling lemma implies

$$\mathrm{OPT}_d(T_d, 1) = \frac{\mathrm{OPT}_d(T_d, \delta_d)}{\delta_d} \ge \frac{\mathrm{Jung}_d \cdot \sqrt{d}}{\delta_d} = \mathrm{Jung}_d \cdot \sqrt{\frac{d}{4 \ln T_d}}.$$

Since $\ln T_d = o(d)$, the right-hand side tends to $+\infty$ as $d \to \infty$.

It remains to establish (11). Consider the a posteriori adversary that answers 0 to every query. For a transcript with query directions $v_1, \ldots, v_{T_d} \in S^{d-1}$, the feasible region is

$$\Phi_{T_d} := \bigcap_{t=1}^{T_d} \Big\{ x \in \mathbb{R}^d : \ |\langle v_t, x \rangle| \le \delta_d \Big\}.$$

It suffices to show that $\mathrm{diam}(\Phi_{T_d}) \ge \sqrt{d}$ for all large enough d.

Let $X, Y \sim \mathcal{N}(0, I_d)$ be independent. Using the standard Gaussian tail bound

$$\Pr(|g| > t) \le e^{-t^2/2} \qquad (g \sim \mathcal{N}(0,1),\ t \ge 0),$$

we have, for any query direction $v \in S^{d-1}$,

$$\Pr\big(|\langle v, X \rangle| > \delta_d\big) \le e^{-\delta_d^2/2},$$

and the same bound holds with X replaced by Y. A union bound over all T_d queries therefore yields

$$\Pr(X \notin \Phi_{T_d}) + \Pr(Y \notin \Phi_{T_d}) \le 2T_d e^{-\delta_d^2/2} = \frac{2}{T_d}.$$

On the other hand, since $X - Y \sim \mathcal{N}(0, 2I_d)$, we may write

$$\|X - Y\|^2 \sim 2\,\chi_d^2,$$

where χ_d^2 is chi-square with d degrees of freedom. Thus

$$\Pr(\|X - Y\|_2 < \sqrt{d}) = \Pr(\chi_d^2 < d/2).$$

By a standard concentration inequality for the Chi-square distribution

$$\Pr(\chi_d^2 < d/2) \le \left(\frac{e^{1/2}}{2}\right)^{d/2},$$

which decays exponentially in d.

Combining the above bounds, we obtain

$$\Pr\Big[X \in \Phi_{T_d},\ Y \in \Phi_{T_d},\ \|X - Y\| \ge \sqrt{d}\Big] \ \ge\ 1 - \frac{2}{T_d} - \left(\frac{e^{1/2}}{2}\right)^{d/2}.$$

We may assume $T_d \ge 3$ for all sufficiently large d, and hence the right-hand side is strictly positive for all large d. Hence there exist $x, y \in \Phi_{T_d}$ with $\|x - y\| \ge \sqrt{d}$, and therefore $\mathrm{diam}(\Phi_{T_d}) \ge \sqrt{d}$. Apply Jung's Theorem (Theorem 4). This proves (11) and completes the proof.

Proof of Item 2 (superexponential budgets). The argument parallels the upper-bound proof of Main Theorem 1; the only difference is that here we use the explicit dimension-dependent covering bounds from Lemma 14.

Fix $\alpha \in (0, \pi/2)$ and let $T_\alpha := \mathcal{N}(S^{d-1}, \alpha)$. By Lemma 15, after querying an α-cover of S^{d-1} the feasible region satisfies

$$\mathrm{diam}(\Phi_{T_\alpha}) \le \frac{2\delta}{\cos\alpha}.$$

By Lemma 12, the optimal worst-case error at time T_α is equal to the Chebyshev radius of the feasible region Φ_{T_α}. By Jung's theorem (Theorem 4),

$$\mathtt{rad}(\Phi_{T_\alpha}) \le \mathrm{Jung}_d \cdot \mathrm{diam}(\Phi_{T_\alpha}) \le \mathrm{Jung}_d \cdot \frac{2\delta}{\cos\alpha},$$

and hence

$$\mathrm{OPT}_d(T_\alpha, \delta) \le \mathrm{Jung}_d \cdot \frac{2\delta}{\cos\alpha}.$$

Using $\mathrm{OPT}_d(\infty, \delta) = 2\delta\mathrm{Jung}_d$ (Main Theorem 1), we obtain

$$\mathsf{ExcessErr}_d(T_\alpha, \delta) = \mathrm{OPT}_d(T_\alpha, \delta) - \mathrm{OPT}_d(\infty, \delta) \le 2\delta\mathrm{Jung}_d\left(\frac{1}{\cos\alpha} - 1\right).$$

Now set

$$\beta_d := \frac{1}{\sqrt[d-1]{T(d)}}.$$

Since $\ln T(d) = \omega(d)$, we have $\frac{\ln T(d)}{d-1} \to +\infty$, hence $\sqrt[d-1]{T(d)} = \exp\left(\frac{\ln T(d)}{d-1}\right) \to \infty$, and thus $\beta_d \to 0$. By Lemma 14 (applied with $\alpha = \beta_d$), we have $T(d) \geq \mathcal{N}(S^{d-1}, \beta_d) = T_{\beta_d}$ for all sufficiently large d. Using monotonicity in T (Lemma 8), and therefore also of $\mathsf{ExcessErr}_d(\cdot, \delta)$, we get

$$\mathsf{ExcessErr}_d(T(d), \delta) \leq \mathsf{ExcessErr}_d(T_{\beta_d}, \delta) \leq 2\delta \mathrm{Jung}_d \left(\frac{1}{\cos \beta_d} - 1\right).$$

Finally, $2\delta \cdot \mathrm{Jung}_d = \sqrt{\frac{2d}{d+1}}\, \delta < \sqrt{2}\, \delta$, and since $\beta_d \to 0$ and cos is continuous with $\cos 0 = 1$, the right-hand side tends to 0. This proves the claim. ■

Appendix C. Improper reconstruction

In this section, we analyze an *improper* variant of the reconstruction game. In the improper reconstruction model, the reconstructor does not output a single estimate $\hat{x}_T \in \mathbb{R}^d$. Instead, after observing the interaction transcript, it produces a prediction rule $G : S^{d-1} \to \mathbb{R}$, whose goal is to predict future answers $\langle x^*, v\rangle$ for every direction $v \in S^{d-1}$.

To define the optimal reconstruction error for the improper version, we must specify a loss for a prediction rule relative to the secret point x^*. In this work, we use a worst-case loss: after T rounds, imagine an additional "evaluation round" in which the adversary chooses the worst direction $v \in S^{d-1}$. The loss of a prediction rule $\hat{G}_T : S^{d-1} \to \mathbb{R}$ is then

$$\sup_{v \in S^{d-1}} \left|\hat{G}_T(v) - \langle x^*, v\rangle\right|.$$

Definition 18 (Improper optimal reconstruction error) *Fix a dimension $d \geq 1$ and work in $\mathbb{R}^d$. Let $\mathsf{RC}_{\text{improper}}$ be the set of improper reconstruction strategies for T rounds, i.e., strategies that after T rounds output a prediction rule $\hat{G}_T : S^{d-1} \to \mathbb{R}$. We define the* improper optimal reconstruction error *by*

$$\mathrm{OPT}_d^{\text{improper}}(T, \delta) := \inf_{\mathcal{R} \in \mathsf{RC}_{\text{improper}}} \sup_{x^* \in \mathbb{R}^d} \sup_{\mathcal{A} \in \mathsf{ADV}(x^*)} \sup_{v \in S^{d-1}} \left|\hat{G}_T(v) - \langle x^*, v\rangle\right|.$$

Here $\mathsf{ADV}(x^*)$ denotes the set of adversary response strategies consistent with the secret point x^*. All strategies are assumed deterministic.

The performance of an improper reconstructor is likewise governed by the geometry of the feasible region, as in the usual reconstruction setting. Recall that given a transcript $(v_1, r_1), \ldots, (v_T, r_T)$ at noise level $\delta > 0$, the feasible region after T rounds is

$$\Phi_T := \left\{x \in \mathbb{R}^d : |\langle x, v_t\rangle - r_t| \leq \delta \text{ for all } t \in [T]\right\}.$$

As in the usual reconstruction setting, we may take the adversary to be a posteriori. Once the transcript is fixed, the adversary may choose any secret point $x^* \in \Phi_T$. The relevant geometric

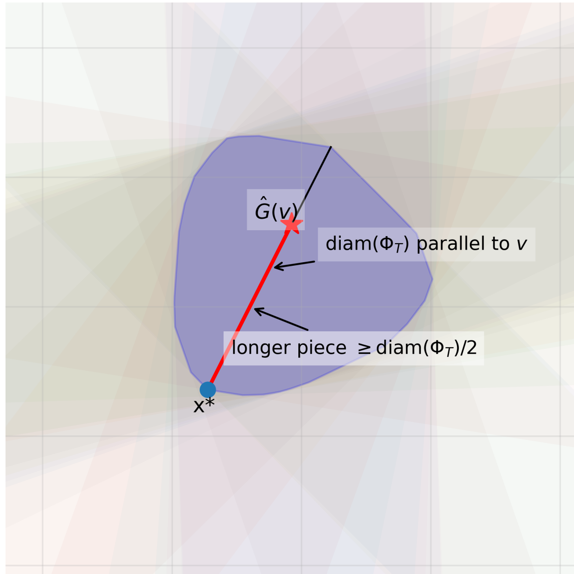

Figure 4: For any prediction rule $\hat{G}$ there exists a secret point $x^* \in \Phi_T$ and a direction v such that $\left|\hat{G}(v) - \langle x^*, v\rangle\right| \geq \mathrm{diam}(\Phi_T)/2$.

invariant is now the diameter: given a feasible region Φ_T, the worst-case improper error is exactly $\mathrm{diam}(\Phi_T)/2$.

Indeed, for any direction v, the set of possible values of $\langle x^*, v\rangle$ as x^* ranges over Φ_T is the projection of Φ_T onto the line spanned by v, hence it forms an interval. If the diameter of Φ_T is attained along a line parallel to v, then regardless of the prediction made by the reconstructor, the adversary can choose a secret point $x^* \in \Phi_T$ such that the prediction error in direction v is at least $\mathrm{diam}(\Phi_T)/2$.

Conversely, if the reconstructor predicts, for each direction v, the midpoint of this projection interval, then the prediction error in direction v is at most $\mathrm{diam}(\Phi_T)/2$. Therefore, in the improper model, the worst-case error induced by a feasible region Φ_T is exactly

$$\frac{1}{2}\,\mathrm{diam}(\Phi_T).$$

See Figure 4 for a geometric illustration.

Formally, this observation is captured by the following lemma. We omit the proof, as it is a direct formalization of the argument above.

Lemma 19 (Improper OPT via the feasible region) *Fix a transcript with a nonempty feasible region Φ_T. For any reconstructor output $\hat{G} : S^{d-1} \to \mathbb{R}$,*

$$\sup_{x^* \in \Phi_T} \sup_{v \in S^{d-1}} \left|\hat{G}(v) - \langle x^*, v\rangle\right| \;\geq\; \frac{1}{2}\,\mathrm{diam}(\Phi_T).$$

Moreover, equality holds for a prediction rule outputting the midpoint of the projection of Φ_T onto the line spanned by queried direction v. Consequently,

$$\mathrm{OPT}_d^{\mathrm{improper}}(T,\delta) = \inf_{\mathcal{R}} \sup_{\substack{\textit{feasible transcripts} \\ \textit{generated against } \mathcal{R}}} \frac{1}{2}\,\mathrm{diam}(\Phi_T).$$

Consequently, any upper (respectively, lower) bound on $\mathrm{diam}(\Phi_T)$ that holds for all transcripts after T rounds immediately yields the corresponding upper (respectively, lower) bound on the improper minimax error, up to the factor $1/2$.

C.1. Analogue of the Main Theorems for improper reconstruction

In the improper model surprisingly the asymptotic optimal reconstruction error is strictly smaller than usual asymptotic reconstruction error:

$$\mathrm{OPT}_d^{\mathrm{improper}}(\infty,\delta) := \lim_{T\to\infty} \mathrm{OPT}_d^{\mathrm{improper}}(T,\delta) = \delta.$$

Indeed, by Lemma 19, the loss is exactly $\frac{1}{2}\mathrm{diam}(\Phi_T)$. Querying a sufficiently fine angular covering of the sphere ensures that the diameter of the feasible region is at most $(2+o(1))\delta$, and hence the improper reconstruction error is at most $(1+o(1))\delta$. Conversely, an adversary that answers 0 to every query ensures that the diameter of the feasible region remains at least 2δ, forcing improper error at least δ.

It follows that the improper excess error is captured by exactly the same strategies that already yield the asymptotic improper error. This stands in contrast to the (usual) reconstruction model: there, controlling the diameter alone is not sufficient, since the loss is governed by the Chebyshev radius, and one must additionally relate radius to diameter (via Jung-type arguments).

To control the diameter of the feasible region, the reconstructor must query a sufficiently dense net of directions on the sphere. Quantitatively, this leads to a convergence rate dictated by the angular covering number of S^{d-1}.

Theorem 20 (Improper excess error) *Fix $d \ge 1$. There exist constants $a_d, A_d > 0$, depending only on d, and an integer $T'_d \ge 1$ such that for all $T \ge T'_d$,*

$$\delta\big(1 + a_d\,T^{-2/(d-1)}\big) \;\le\; \mathrm{OPT}_d^{\mathrm{improper}}(T,\delta) \;\le\; \delta\big(1 + A_d\,T^{-2/(d-1)}\big).$$

Proof

Upper bound. Let $T \ge 16^{d-1}$. Then

$$\alpha \;:=\; \frac{8}{T^{1/(d-1)}} \;\le\; \frac{8}{16} \;=\; \frac{1}{2} \;<\; \frac{\pi}{4}.$$

By Lemma 14, the sphere admits an angular α-covering of size at most T. Fix such a covering $\mathcal{V} = \{v_1,\dots,v_T\} \subseteq S^{d-1}$. Consider the reconstructor that, in rounds $i = 1,\dots,T$, queries v_i. Let r_i be the corresponding answers, satisfying $|r_i - \langle x^*, v_i\rangle| \le 1$, and let Φ_T be the feasible region induced by the transcript:

$$\Phi_T \;=\; \Big\{x \in \mathbb{R}^d :\; |\langle x, v_i\rangle - r_i| \le 1 \;\text{ for all } i \in [T]\Big\}.$$

By Lemma 15, for any answers consistent with noise level 1 we have

$$\mathrm{diam}(\Phi_T) \;\le\; \frac{2}{\cos\alpha}.$$

Hence,

$$\sup_{\text{adversaries}} \mathrm{diam}(\Phi_T) \;\le\; \frac{2}{\cos\alpha}.$$

Since $\alpha \in [0, \pi/4]$, we may use the elementary inequality $\frac{1}{\cos\alpha} \le 1 + 2\alpha^2$ to obtain

$$\sup_{\text{adversaries}} \mathrm{diam}(\Phi_T) \;\le\; 2\big(1 + 2\alpha^2\big) \;=\; 2 + O(\alpha^2).$$

Finally

$$\mathrm{OPT}_d^{\text{improper}}(T) \;\le\; \frac{1}{2} \cdot \sup_{\text{adversaries}} \mathrm{diam}(\Phi_T) \;\le\; \frac{1}{2} \cdot \frac{2}{\cos\alpha} \;=\; \frac{1}{\cos\alpha} \;\le\; 1 + O(\alpha^2).$$

Substituting $\alpha = 8\,T^{-1/(d-1)}$ yields

$$\mathrm{OPT}_d^{\text{improper}}(T) \;\le\; 1 + O\big(T^{-2/(d-1)}\big),$$

as claimed.

Lower bound. Consider the adversary that answers $y_t = 0$ for every query v_t. Fix any (possibly adaptive) sequence of queries $v_1, \dots, v_T$.

Let $\alpha > 0$ be

$$\alpha := \frac{1}{\sqrt[d-1]{T}}.$$

By Lemma 14 one has $\mathcal{N}_\angle(S^{d-1}, \alpha) > 2T$ and hence, there exists $v \in S^{d-1}$ such that for every $i \in [T]$, $\rho(v, v_i) \ge \alpha$ and $\rho(v, -v_i) \ge \alpha$, where ρ denotes the angular metric (see Equation (7)). Equivalently, for all $i \in [T]$,

$$|\langle v, v_i\rangle| \;\le\; \cos\alpha.$$

Define two antipodal points

$$x := \frac{v}{\cos\alpha}, \qquad y := -\frac{v}{\cos\alpha}.$$

Then for every query v_i one has $|\langle x, v_i\rangle| \le 1$, and similarly for y. Hence $x, y \in \Phi_T$, and therefore

$$\mathrm{diam}(\Phi_T) \;\ge\; \|x - y\|_2 \;=\; \frac{2}{\cos\alpha}.$$

This implies

$$\mathrm{OPT}_d^{\text{improper}}(T) \;\ge\; \frac{1}{\cos\alpha}.$$

Expanding $1/\cos\alpha = 1 + \Theta(\alpha^2)$ for small α, we obtain

$$\mathrm{OPT}_d^{\text{improper}}(T) - 1 \;\ge\; \Theta(\alpha^2) \;=\; \Theta\big(T^{-2/(d-1)}\big),$$

which gives the desired lower bound. ■

Analogue of Main Theorem 3 Main Theorem 3 applies equally to the improper excess error $\mathrm{OPT}_d^{\mathrm{improper}}(T,\delta)-\delta$, since its proof proceeds by controlling the diameter of the feasible region and then translating this control into a bound on the loss. In the usual reconstruction setting, this translation uses Jung's theorem for the upper bound and the triangle inequality for the lower bound:

$$\frac{1}{2}\operatorname{diam}(\Phi_T) \;\le\; \mathtt{rad}(\Phi_T) \;\le\; \sqrt{\frac{d}{2(d+1)}}\operatorname{diam}(\Phi_T).$$

In the improper setting, the loss is exactly $\operatorname{diam}(\Phi_T)/2$, so no passage through the Chebyshev radius is needed. Hence

1. If $\ln T(d) = o(d)$ and $T(d) \ge 3$ for all sufficiently large d, then

$$\mathrm{OPT}_d^{\mathrm{improper}}(T(d),\delta) \xrightarrow{d\to\infty} +\infty.$$

2. If $\ln T(d) = \omega(d)$ for all sufficiently large d, then

$$\mathrm{OPT}_d^{\mathrm{improper}}(T(d),\delta) \xrightarrow{d\to\infty} \delta.$$

Appendix D. Robust Jung Theorem

In this appendix we formalize and prove the main technical tool used in the proof of Main Theorem 2, concerning sets whose diameter and Chebyshev radius are close to extremal in Jung's inequality.

We work here with a slightly different formulation than in the Technical Overview; it will turn out that the two formulations are equivalent. The main notion we use is the Hausdorff distance.

D.1. Supporting definitions and lemmas

Definition 21 (Hausdorff distance) *For nonempty compact sets $A, B \subset \mathbb{R}^d$, the* Hausdorff distance *is defined by*

$$\operatorname{dist}_H(A,B) := \max\Big\{\sup_{a\in A}\inf_{b\in B}\|a-b\|_2, \sup_{b\in B}\inf_{a\in A}\|a-b\|_2\Big\}.$$

The fact that dist_H is indeed a metric on the space of nonempty compact subsets of $\mathbb{R}^d$ can be found in, e.g., Illanes and Nadler (1999).

For a set $S \subset \mathbb{R}^d$ and $r > 0$, we denote by

$$S + B(r) := \{x + y : x \in S,\ \|y\|_2 \le r\}$$

the Minkowski sum of S with the Euclidean ball of radius r.

It follows directly from the definitions that $\operatorname{dist}_H(S_1,S_2) \le r$ if and only if

$$S_1 \subseteq S_2 + B(r) \qquad \text{and} \qquad S_2 \subseteq S_1 + B(r),$$

that is, S_1 and S_2 are r-close in the sense used in the Technical Overview. When $r = 0$, this condition reduces to $S_1 = S_2$.

Chebyshev radius. For a bounded set $S \subset \mathbb{R}^d$, its *Chebyshev radius* is

$$\mathtt{rad}(S) := \inf_{x \in \mathbb{R}^d} \sup_{y \in S} \|x - y\|_2.$$

Any point attaining the infimum is called a *Chebyshev center* of S.

Recall Jung's theorem, for a proof see (Gruber, 2007, Theorem 3.3).

Theorem *[Jung's theorem] Let $S \subseteq \mathbb{R}^d$ be a bounded set with diameter $D := \mathrm{diam}(S)$. Then its Chebyshev radius satisfies*

$$\mathtt{rad}(S) \leq \sqrt{\frac{d}{2(d+1)}}\, D.$$

The factor $\sqrt{\frac{d}{2(d+1)}}$ is called Jung's constant*; we denote it by* Jung_d.

Recall that a *witness* of S is a subset $W \subseteq S$ consisting of $d+1$ points such that

$$\mathtt{rad}(W) \geq \mathrm{Jung}_d.$$

The following lemma is a standard consequence of Helly's theorem for convex sets in $\mathbb{R}^d$; see, for example, Chapter 3 of Gruber (2007).

Lemma 22 *Let $S \subset \mathbb{R}^d$ be a compact set. There exists a subset $W \subseteq S$ with $|W| \leq d+1$ such that $\mathtt{rad}(W) = \mathtt{rad}(S)$.*

Proof By definition of the Chebyshev radius, for every $r < \mathtt{rad}(S)$ we have

$$\bigcap_{x \in S} B(x, r) = \emptyset.$$

By Helly's theorem, there exist points $x_0, \ldots, x_d \in S$ such that

$$\bigcap_{i=0}^{d} B(x_i, r) = \emptyset.$$

Equivalently, $\mathtt{rad}(\{x_0, \ldots, x_d\}) > r$. Now set $r_n := \mathtt{rad}(S) - \frac{1}{n}$. Applying the above with $r = r_n$, we can choose a $(d{+}1)$-tuple

$$W_n := (x_0^{(n)}, \ldots, x_d^{(n)}) \in S^{d+1}$$

such that $\mathtt{rad}(W_n) > r_n$, and hence $\mathtt{rad}(W_n) \geq \mathtt{rad}(S) - \frac{1}{n}$. Since S is compact, S^{d+1} is compact as well, so (W_n) has a convergent subsequence $W_{n_k} \to W \in S^{d+1}$. By continuity of the Chebyshev radius for finite point sets,

$$\mathtt{rad}(W) \;\geq\; \limsup_{k \to \infty} \mathtt{rad}(W_{n_k}) \;\geq\; \mathtt{rad}(S).$$

On the other hand, $W \subseteq S$ implies $\mathtt{rad}(W) \leq \mathtt{rad}(S)$.

Therefore $\mathtt{rad}(W) = \mathtt{rad}(S)$, as claimed. ■

In particular, if $\mathrm{rad}(S) \geq \mathrm{Jung}_d$, then W is a witness. If $|W| < d+1$, a witness of cardinality exactly $d+1$ can be constructed by duplicating elements of W.

D.2. Proof of Robust Jung Theorem

Theorem 5 restatement. *Let $d \geq 1$ and let $S \subset \mathbb{R}^d$ be a compact set satisfying*

$$\operatorname{diam}(S) \leq 1 + \beta \qquad \textit{and} \qquad \texttt{rad}(S) \geq \mathrm{Jung}_d,$$

for some $0 \leq \beta \leq \beta_d$, where $\beta_d > 0$ depends only on d. Then, there exists a set of points

$$\Delta = \{x_0, x_1, \ldots, x_d\} \subset \mathbb{R}^d$$

forming the vertex set of a regular simplex of edge length 1 such that every witness $W \subseteq S$ is $C_d\beta$-close to Δ, for a constant $C_d > 0$ depending only on d, i.e. $\operatorname{dist}_H(\Delta, W) \leq C_d\beta$.

To prove this theorem, we first show that every witness is close (in Hausdorff distance) to the vertex set of a regular simplex. We then combine this with a stability statement for regular simplices (proved in Appendix E), which implies that any two such vertex sets must be close to each other. Putting these two ingredients together yields the claim.

Lemma 23 (The witnesses is nearly regular) *Suppose closed set $S \subseteq \mathbb{R}^d$ has diameter 1 and Chebyshev radius $r \geq (1-\epsilon)\sqrt{\frac{d}{2(d+1)}}$. Then there exist $x_0, \ldots, x_d \in S$ such that $\|x_i - x_j\|_2 \geq 1 - O(d^2\epsilon)$ for all $i \neq j$.*

Proof Lemma 22 implies that there exist $x_0, \ldots, x_d \in S$ whose Chebyshev radius is r.

Suppose without loss of generality that the minimal enclosing ball is centered at the origin, and let $i_0, \ldots, i_{n-1}$ be the indices of the x_i such that $\|x_i\|_2 = r$. The origin is contained in the convex hull of the x_{i_s}'s, say $\sum_s \lambda_s x_{i_s} = 0$, where $\lambda_s \geq 0$ and $\sum_s \lambda_s = 1$.

Write $\|x_{i_s} - x_{i_t}\|^2 = 1 - \delta_{st}$. For all s,

$$1 - \lambda_s = \sum_{t \neq s} \lambda_t = \sum_{t \neq s} \lambda_t(\|x_{i_s} - x_{i_t}\|^2 + \delta_{st}).$$

We have

$$\sum_{t \neq s} \lambda_t \|x_{i_s} - x_{i_t}\|^2 = \sum_t \lambda_t \|x_{i_s} - x_{i_t}\|^2 = \sum_t \lambda_t(2r^2 - 2\langle x_{i_s}, x_{i_t}\rangle) = 2r^2,$$

since $\sum_t \lambda_t x_{i_t} = 0$. Therefore

$$1 - \lambda_s = 2r^2 + \sum_{t \neq s} \lambda_t \delta_{st}. \tag{12}$$

Summing over all s,

$$n - 1 = 2nr^2 + \sum_{s \neq t} \lambda_t \delta_{st},$$

and so

$$r^2 = \frac{n-1}{2n} - \frac{1}{2n} \sum_s \sum_{t \neq s} \lambda_t \delta_{st}.$$

On the other hand, $r^2 \geq (1-2\epsilon)\frac{d}{2(d+1)}$. Since $r^2 \leq \frac{n-1}{2n}$, if $(1-2\epsilon)\frac{d}{2(d+1)} > \frac{d-1}{2d}$ (which happens when $\epsilon < \frac{1}{2d^2}$) then $n = d+1$, and we can take $i_t = t$. Moreover,

$$\sum_s \sum_{t \neq s} \lambda_t \delta_{st} = d - 2(d+1)r^2 \leq 2d\epsilon. \tag{13}$$

Substituting (13) in (12) yields

$$1 - \lambda_s \leq 2r^2 + 2d\epsilon \leq \frac{d}{d+1} + 2d\epsilon,$$

hence

$$\lambda_s \geq \frac{1}{d+1} - 2d\epsilon.$$

If $\epsilon \leq \frac{1}{2d(d+1)}$ then $\lambda_s \geq \frac{1}{2(d+1)}$. Since

$$\sum_s \sum_{t \neq s} \lambda_t \delta_{st} \geq \lambda_s \max_{t \neq s} \delta_{st},$$

in view of (13) this means that $\delta_{st} \leq 2d(d+1)\epsilon$ for all $s \neq t$.

Finally, $\|x_s - x_t\|_2 \geq \sqrt{1 - 2d(d+1)\epsilon} \geq 1 - 2d(d+1)\epsilon$. This bound holds under the assumption $\epsilon \leq \frac{1}{2d(d+1)}$. Otherwise, the bound in the theorem trivially holds. ■

Lemma 24 (From nearly regular to regular) *There exist constants $c_d > 0$ and $\epsilon_d > 0$, depending only on d, such that the following holds. Let $S = \{x_0, \ldots, x_d\} \subset \mathbb{R}^d$ satisfy*

$$\operatorname{diam}(S) \leq 1 \qquad \textit{and} \qquad \texttt{rad}(S) \geq (1-\epsilon)\sqrt{\frac{d}{2(d+1)}}$$

for some $0 < \epsilon \leq \epsilon_d$. Then there exists a regular simplex Δ of diameter 1 such that

$$\operatorname{dist}_H(S, \Delta) \leq c_d\, \epsilon.$$

Proof We proceed by induction on the dimension d. The base case $d = 1$ is trivial. By the induction hypothesis and Theorem 23, it suffices to prove the following: given points $y_1, \ldots, y_d$ forming a regular simplex in a subspace of $\mathbb{R}^d$ (with edge length 1), assume there is a point y_0 such that

$$\|y_0 - y_k\|^2 = 1 + f_k(\epsilon), \qquad f_k(\epsilon) = O(\epsilon).$$

Then there exists a point y_* such that $y_*, y_1, \ldots, y_d$ form a regular simplex and

$$\|y_* - y_0\|_2 \leq O(\epsilon).$$

Let v be an affine functional with $v\big(\operatorname{aff}\{y_1, \ldots, y_d\}\big) = 0$ and $v(y_0) > 0$, and let y_* be a point in the half-space $\{x \in \mathbb{R}^d : v(x) > 0\}$ that, together with $y_1, \ldots, y_d$, forms a regular simplex. Define

$$F \colon \mathbb{R}^d \to \mathbb{R}^d, \qquad F(x) = \big(\|x - y_1\|^2 - 1, \ldots, \|x - y_d\|^2 - 1\big).$$

The ith row of $DF(x)$ is $2(x - y_i)^\top$. A direct computation at y_* yields

$$DF(y_*)\, DF(y_*)^\top \;=\; 4\big(\langle y_* - y_i,\; y_* - y_j\rangle\big)_{i,j} \;=\; 2\,(I + J),$$

where J is the all-ones matrix. Thus the singular values of $DF(y_*)$ are $\sqrt{2(d+1)}$ (once) and $\sqrt{2}$ (multiplicity $d-1$), so $DF(y_*)$ is invertible and

$$\big\|DF(0)^{-1}\big\| \;=\; \frac{1}{\sigma_{\min}(DF(y_*))} \;=\; \frac{1}{\sqrt{2}}.$$

Applying the inverse function theorem at y_* gives neighborhoods $U \ni y_*$ and $V \ni 0$ together with a local C^1-inverse G. Hence, for all ϵ sufficiently small, we have $a_0 := F(y_0) \in V$. Note that there are at most two preimages of a_0, since d spheres centered at $y_1, y_2, \ldots, y_d$ can intersect in at most two points by an induction-on-dimension argument. As one of these preimages lies in the half-space where $v < 0$, while $v(y_0) > 0$, it follows that $y_0 \in U$.

By the multivariate version of Taylor's theorem applied to G, for every $a \in V$ we can write

$$G(a) = DF^{-1}(0)\, a + H(a)\, a, \qquad H\colon V \to \mathcal{M}(d \times d),$$

where H is a continuous matrix-valued function with $\lim_{a\to 0} H(a) = 0$. Therefore,

$$\|y_0 - y_*\|_2 = \|G(a_0) - G(0)\|_2 \;\le\; \big(\|DF(0)^{-1}\|_2 + \|H(a_0)\|\big) \cdot \sqrt{\sum_{k=1}^{d} f_k^2(\epsilon)} = O(\epsilon).$$

■

Finally we prove Robust Jung Theorem. Recall it's formulation

Theorem 5 restatement. *Let $d \ge 1$ and let $S \subset \mathbb{R}^d$ be a compact set satisfying*

$$\operatorname{diam}(S) \le 1 + \beta \qquad \textit{and} \qquad \mathtt{rad}(S) \ge \mathrm{Jung}_d,$$

for some $0 \le \beta \le \beta_d$, where $\beta_d > 0$ depends only on d. Then, there exists a set of points

$$\Delta = \{x_0, x_1, \ldots, x_d\} \subset \mathbb{R}^d$$

forming the vertex set of a regular simplex of edge length 1 such that every witness $W \subseteq S$ is $C_d\beta$-close to Δ, for a constant $C_d > 0$ depending only on d, i.e. $\operatorname{dist}_H(\Delta, W) \le C_d\beta$.

Proof Fix two witnesses $W, W' \subseteq S$. By Lemma 24, there exist regular simplices Δ, Δ' of edge length 1 such that

$$\operatorname{dist}_H(W, \Delta) \le c_d\,\beta \qquad \text{and} \qquad \operatorname{dist}_H(W', \Delta') \le c_d\,\beta.$$

Since $W, W' \subseteq S$ and $\operatorname{diam}(S) \le 1 + \beta$, we have

$$\operatorname{diam}(\Delta \cup \Delta') \;\le\; \operatorname{diam}(W \cup W') + 2c_d\beta \;\le\; 1 + \beta + 2c_d\beta.$$

Shrinking $\beta_d > 0$ if necessary, we may assume $1 + \beta + 2c_d\beta \le 1 + \beta_d$. Therefore, by Theorem 25,

$$\operatorname{dist}_H(\Delta, \Delta') \le (d+1)d^2\,\beta.$$

Applying the triangle inequality for the Hausdorff distance,

$$\operatorname{dist}_H(W, W') \;\le\; \operatorname{dist}_H(W, \Delta) + \operatorname{dist}_H(\Delta, \Delta') + \operatorname{dist}_H(\Delta', W') \;\le\; \big(2c_d + (d+1)d^2\big)\beta.$$

In particular, all witnesses are contained in a common $C_d\beta$-neighborhood of Δ, for a constant $C_d > 0$ depending only on d. This completes the proof. ■

Appendix E. Regular simplex rotation

The aim of this section is to prove the following theorem.

Theorem 25 *Let $\Delta, \Delta' \subset \mathbb{R}^d$ be the vertex sets of regular simplices of edge length 1. Suppose that*

$$\operatorname{diam}(\Delta \cup \Delta') \leq 1 + \beta$$

for some $0 \leq \beta < \beta_d$, where $\beta_d > 0$ depends only on d. Then

$$\operatorname{dist}_H(\Delta, \Delta') \leq (d+1)d^2\,\beta.$$

The proof uses Lie–algebra techniques together with several linear–algebraic properties of regular simplices. First, in Theorem 27 we show that $\operatorname{diam}(\Delta\cup\Delta') = 1$ only if $\Delta' = \Delta$. Next, in Theorem 29 we establish a weaker continuity statement: if $\operatorname{diam}(\Delta\cup\Delta')$ is small, then $\operatorname{dist}_H(\Delta, \Delta')$ is also small. After that, in Theorem 30 we relate $\operatorname{dist}_H(\Delta, \Delta')$ to the size of a Euclidean transformation (τ, R) (translation + rotation) carrying Δ to Δ'; specifically, we show that for $\operatorname{dist}_H(\Delta, \Delta') < \frac{1}{2}$ there exists (τ, R) with

$$\operatorname{dist}_H(\Delta, \Delta') \;>\; B_d\big(\|R - I_d\|_2 + \|\tau\|\big),$$

where B_d is a universal constant. Finally, in Theorem 32 assuming $\|R - I_d\|_2 + \|\tau\|$ is sufficiently small, we pass to the Lie–algebra level and express $\operatorname{dist}_H(\Delta, \Delta')$ and $\operatorname{diam}(\Delta\cup\Delta')$ in terms of the principal logarithm $(\tau_0, B) \in \mathfrak{e}(d)$; taking $\operatorname{diam}(\Delta \cup \Delta') - 1$ small enough then yields the claim.

Remark 26 *Let Δ_0 be a regular unit simplex with vertices $x_0, x_1, \ldots, x_d$. Placing the centroid of Δ_0 at the origin, the vertices form a* tight frame *with frame constant $\frac{1}{2}$. In other words, if $X \in \mathbb{R}^{d\times(d+1)}$ denotes the matrix whose columns are the vectors x_i, then*

$$XX^\top = \tfrac{1}{2}\, I_d.$$

In particular, for any point $y \in \mathbb{R}^d$ one has

$$\|y\|^2 = \frac{1}{2}\sum \langle x_i, y\rangle^2$$

Lemma 27 *Let $\Delta = \{x_0, x_1, \ldots, x_d\}$ and $\Delta' = \{y_0, y_1, \ldots, y_d\}$ be two regular simplices of edge length 1. If*

$$\operatorname{diam}(\Delta \cup \Delta') = 1,$$

then $\Delta = \Delta'$.

Proof For convenience, scale by $\sqrt{2}$. Thus we may assume both simplices have edge length $\sqrt{2}$, and

$$\operatorname{diam}(\Delta \cup \Delta') = \sqrt{2}.$$

Let $R := \sqrt{\frac{d}{d+1}}$. A regular simplex of edge length $\sqrt{2}$ has circumradius R, and its minimum enclosing ball (its circumball) is unique.

By Jung's theorem, $\texttt{rad}(\Delta \cup \Delta') \leq R$, hence $\Delta \cup \Delta'$ is contained in some ball $B(c, R)$. This ball contains Δ, so by uniqueness of the circumball of Δ, its center c must be the circumcenter of Δ.

The same argument applied to Δ' shows that c is also the circumcenter of Δ'. Denote this common circumcenter by O, and translate so that $O = 0$.

Then $\|x_i\| = \|y_j\| = R$ for all i, j, and $\sum_{i=0}^{d} x_i = 0$. In particular,

$$\langle x_i, x_j \rangle = \begin{cases} R^2 = \frac{d}{d+1}, & i = j, \\ -\frac{1}{d+1}, & i \neq j, \end{cases}$$

and the simplex frame identity holds (see Remark 26):

$$\|y\|^2 \;=\; \sum_{i=0}^{d} \langle y, x_i \rangle^2 \qquad \text{for all } y \in \mathbb{R}^d.$$

Fix j and set $x := y_j$. Since $\operatorname{diam}(\Delta \cup \Delta') = \sqrt{2}$, we have $\|x - x_i\| \leq \sqrt{2}$ for every i. Expanding,

$$\|x - x_i\|^2 = \|x\|^2 + \|x_i\|^2 - 2\langle x, x_i \rangle = 2R^2 - 2\langle x, x_i \rangle \leq 2,$$

so $\langle x, x_i \rangle \geq R^2 - 1 = -\frac{1}{d+1}$ for all i.

Define $b_i := \langle x, x_i \rangle + \frac{1}{d+1} \geq 0$. Using $\sum_i x_i = 0$ and $\|x\|^2 = R^2$,

$$\sum_{i=0}^{d} b_i = \sum_{i=0}^{d} \langle x, x_i \rangle + 1 = 1,$$

and using the inner products of the simplex,

$$\sum_{i=0}^{d} b_i^2 = \sum_{i=0}^{d} \langle x, x_i \rangle^2 + \tfrac{2}{d+1} \sum_{i=0}^{d} \langle x, x_i \rangle + \tfrac{1}{d+1} = \|x\|^2 + \tfrac{1}{d+1} = 1.$$

(Here we used the frame identity 26 and $\sum_i \langle x, x_i \rangle = 0$.)

Since $b_i \geq 0$ and $\sum_i b_i = 1$, we have $\sum_i b_i^2 \leq 1$, with equality iff exactly one $b_i = 1$ and the others are 0. Thus $b_{i^*} = 1$ for some i^*, i.e. $\langle x, x_{i^*} \rangle = R^2$. Because $\|x\| = \|x_{i^*}\| = R$, equality in Cauchy–Schwarz gives $x = x_{i^*}$. Hence every y_j equals some x_i, so $\Delta' \subseteq \Delta$. By symmetry $\Delta \subseteq \Delta'$, and therefore $\Delta = \Delta'$. ■

Definition 28 *Consider the subspace of regular unit simplices with ordered vertices:*

$$\mathcal{Y} = \Big\{ (x_0, x_1, \ldots, x_d) \in (\mathbb{R}^d)^{d+1} \;\Big|\; \|x_i - x_j\| = 1 \textit{ for all } i \neq j \Big\}.$$

This is a smooth algebraic submanifold of $(\mathbb{R}^d)^{d+1}$. The symmetric group S_{d+1} acts freely on $\mathcal{Y}$ by permuting vertices; hence the quotient $\mathcal{X} := \mathcal{Y}/S_{d+1}$ is a smooth manifold. Notice that on $\mathcal{X}$ the Hausdorff distance is well-defined, and the resulting metric topology agrees with the manifold topology.

Lemma 29 *Assume Δ_0 and Δ' are regular simplices with edge length 1. For every $\varepsilon > 0$ there exists $\beta_\varepsilon > 0$ such that*

$$\operatorname{diam}(\Delta_0 \cup \Delta') < 1 + \beta_\varepsilon$$

implies

$$\operatorname{dist}_H(\Delta_0, \Delta') \leq \varepsilon.$$

Proof

Recall that $\mathcal{X}$ is the space of regular simplices with edge length 1, represented by their vertices (see Theorem 28).

Define the function $D_{\Delta_0}\colon \mathcal{X} \to \mathbb{R}$ by

$$D_{\Delta_0}(\Delta') \;=\; \big(\operatorname{diam}(\Delta_0 \cup \Delta')\big) - 1.$$

This function is continuous since

$$\operatorname{diam}(\Delta_0 \cup \Delta') = \max\Big(1, \max_{i,j} \|x_i - x_j'\|\Big),$$

which is the maximum of finitely many continuous functions. Moreover, the equation

$$D_{\Delta_0}(\Delta) = 0$$

has the unique solution $\Delta = \Delta_0$ by Theorem 27. Now fix any $\beta_0 > 0$. The set of simplices

$$\{\Delta \in \mathcal{X} : D_{\Delta_0}(\Delta) \le \beta_0\}$$

is compact. Hence, we can argue by compactness: for every $\varepsilon > 0$ there exists $0 < \beta \le \beta_0$ such that whenever $D_{\Delta_0}(\Delta) \le \beta$ one has

$$\operatorname{dist}_H(\Delta, \Delta_0) \le \varepsilon.$$

Indeed, suppose the contrary. Then there exists $\varepsilon_0 > 0$ and a sequence of simplices Δ_n such that

$$\operatorname{diam}(\Delta_n \cup \Delta_0) \le 1 + \tfrac{1}{n} \qquad \text{and} \qquad \operatorname{dist}_H(\Delta_n, \Delta_0) \ge \varepsilon_0.$$

By compactness, a subsequence Δ_{n_k} converges to some $\Delta^\star$. Passing to the limit we obtain

$$\operatorname{diam}(\Delta^\star \cup \Delta_0) = 1, \qquad \operatorname{dist}_H(\Delta^\star, \Delta_0) \ge \varepsilon_0,$$

which contradicts the uniqueness of the solution to $D_{\Delta_0}(\Delta) = 0$. ■

Lemma 30 *Let Δ_0 and Δ' be two regular unit simplices. If $\operatorname{dist}_H(\Delta_0, \Delta') < \frac{1}{2}$, then there exists a Euclidean transformation $(\tau, R) \in E(d)$, with translation part $\tau \in \mathbb{R}^d$ and rotation part $R \in O(d)$, such that*

$$\operatorname{dist}_H(\Delta_0, \Delta') \;\ge\; \frac{1}{2\sqrt{d+1}}\,\big(\|R - I_d\|_2 + \|\tau\|\big).$$

Proof Consider the Euclidean group of isometries of $\mathbb{R}^d$, denoted $E(d)$. It admits the semidirect product decomposition

$$E(d) = \mathbb{R}^d \rtimes O(d), \qquad (\tau_1, A_1) \cdot (\tau_0, A_0) = \big(\tau_1 + A_1\tau_0,\; A_1 A_0\big).$$

Recall that $\mathcal{Y}$ denotes the set of ordered regular n-simplices of unit edge length (see Theorem 28):

$$\mathcal{Y} = \Big\{(x_0, x_1, \ldots, x_d) \in (\mathbb{R}^d)^{d+1} \;\Big|\; \|x_i - x_j\| = 1 \text{ for all } i \ne j\Big\}.$$

Notice that $E(d)$ acts on $\mathcal{Y}$ by

$$(\tau, A) \cdot (x_0, \dots, x_d) = (Ax_0 + \tau, \dots, Ax_d + \tau).$$

This is a free and transitive action; fixing a regular simplex Δ_0 (with ordered vertices) identifies $E(d)$ diffeomorphically with $\mathcal{Y}$. Denote by ε the Hausdorff distance between the simplices Δ_0 and Δ'. Since the factorization map

$$\mathcal{Y} \longrightarrow \mathcal{Y}/S_{d+1} = \mathcal{X}$$

is a covering and $\varepsilon < \frac{1}{2}$, there exists a permutation $\sigma \in S_{d+1}$ such that the vertices of $\Delta = \{x_0, \dots, x_d\}$ and $\Delta' = \{x'_0, \dots, x'_d\}$ can be paired by

$$x_i \mapsto x'_{\sigma(i)}, \qquad \|x_i - x'_{\sigma(i)}\|_2 \leq \varepsilon \quad \text{for all } i.$$

Consider a Euclidean transformation $(\tau, R) \in E(d)$ such that $R(x_i) + \tau = x'_{\sigma(i)}$. We will prove that

$$(d+1)\varepsilon^2 \geq \tfrac{1}{2}\|R - I_d\|_2^2 + \|\tau\|^2.$$

By Remark 26, the matrix $X \in \mathbb{R}^{d\times(d+1)}$ formed by the columns x_i satisfies

$$XX^\top = \tfrac{1}{2} I_d.$$

Note also that for any matrix M, one has the following connections between Frobenius and operator norms:

$$\|M\|_2 \leq \|M\|_{\mathrm{Fr}} \leq \sqrt{d}\|M\|_2.$$

By assumption, each column of

$$Y := (R - I_d)X + \tau \cdot (1, \dots, 1)$$

has norm at most ε. Hence

$$\|Y\|_{\mathrm{Fr}}^2 \leq (d+1)\varepsilon^2.$$

Since $\sum_i x_i = 0$ and $\sum_i Rx_i = 0$, the cross terms vanish, and therefore

$$\|Y\|_{\mathrm{Fr}}^2 = \|(R - I_d)X\|_{\mathrm{Fr}}^2 + \|\tau\|^2 \leq (d+1)\varepsilon^2.$$

Using the tight-frame identity $XX^\top = \frac{1}{2}I_d$, we obtain

$$\|(R - I_d)X\|_2^2 \geq \|(R - I_d)X \cdot ((R - I_d)X)^\top\|_2 = \tfrac{1}{2}\,\|(R - I_d)(R - I_d)^\top\|_2.$$

Let $M := R - I_d$, and let $\sigma_1, \dots, \sigma_d$ be its singular values. Then $MM^\top$ has $\sigma_1^2, \dots, \sigma_d^2$ singular values; since the operator norm of a matrix is the maximal singular value, one has

$$\|M\|_2^2 = \|MM^\top\|_2.$$

Combining the estimates, we obtain

$$\begin{aligned}
(d+1)\,\varepsilon^2 &\geq \|(R - I_d)X\|_{\mathrm{Fr}}^2 + \|\tau\|^2 \\
&\geq \tfrac{1}{2}\,\|(R - I_d)(R - I_d)^\top\|_2 + \|\tau\|^2 \\
&\geq \tfrac{1}{2}\,\|R - I_d\|_2^2 + \|\tau\|^2.
\end{aligned}$$

Set $a := \|R - I_d\|_2$ and $b := \|\tau\|$. Then $(d+1)\varepsilon^2 \geq \frac{1}{2}a^2 + b^2$, and using $\sqrt{x^2+y^2} \geq (x+y)/\sqrt{2}$ (valid for $x, y \geq 0$),

$$\sqrt{\tfrac{1}{2}a^2 + b^2} \;\geq\; \frac{1}{\sqrt{2}}\left(\frac{a}{\sqrt{2}} + b\right) \;\geq\; \frac{a+b}{2}.$$

Hence

$$\varepsilon \;\geq\; \frac{1}{\sqrt{d+1}}\sqrt{\tfrac{1}{2}a^2 + b^2} \;\geq\; \frac{\|R - I_d\|_2 + \|\tau\|}{2\sqrt{d+1}}.$$

■

We work in the homogeneous (matrix) model of the Euclidean group:

$$(\tau, R) \in E(d) \quad \longleftrightarrow \quad \begin{pmatrix} R & \tau \\ 0 & 1 \end{pmatrix}, \qquad (\tau_1, R_1) \circ (\tau_2, R_2) = (\tau_1 + R_1\tau_2,\ R_1R_2).$$

The Lie algebra is

$$(u, \Omega) \in \mathfrak{e}(d) \quad \longleftrightarrow \quad \begin{pmatrix} \Omega & u \\ 0 & 0 \end{pmatrix}, \qquad \Omega \in \mathfrak{so}(d),\ u \in \mathbb{R}^d.$$

The matrix exponential satisfies

$$\exp\begin{pmatrix} \Omega & u \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^{\Omega} & J(\Omega)\,u \\ 0 & 1 \end{pmatrix}, \qquad J(\Omega) := \sum_{k\geq 0} \frac{\Omega^k}{(k+1)!}.$$

Equivalently,

$$J(\Omega) = \int_0^1 e^{t\Omega}\,dt = \Omega^{-1}\big(e^{\Omega} - I\big), \quad \text{with the last identity understood by power series if } \Omega \text{ is singular.}$$

Define $j(z) := (e^z - 1)/z$ with the convention $j(0) := 1$. Then, by the holomorphic functional calculus,

$$\sigma\big(J(\Omega)\big) = j\big(\sigma(\Omega)\big),$$

where σ denotes the spectrum of the matrix, that is, the multiset of its eigenvalues.

Hence

$$J(\Omega) \text{ invertible} \iff \sigma(\Omega) \cap \big(2\pi i\,\mathbb{Z} \setminus \{0\}\big) = \varnothing.$$

In particular, if $\Omega \in \mathfrak{so}(d)$ has eigen-angles $\{\theta_j\}$ (so $\sigma(\Omega) = \{\pm i\theta_j\}$), then $J(\Omega)$ is invertible iff $\theta_j \notin 2\pi\mathbb{Z}$ for all j. A convenient sufficient condition is $\|\Omega\|_2 < \pi$ (equivalently, $|\theta_j| < \pi$ for all j).

If R has no eigenvalue -1, the principal matrix logarithm $B = \log(R) \in \mathfrak{so}(d)$ is well-defined and real-analytic, and we set

$$\log(\tau, R) := (\tau_0', B), \qquad \tau_0' := J(B)^{-1}\tau.$$

Lemma 31 (Technical lemma for principal logarithm) *Define a neighborhood of* $(0, I_d)$ *in* $E(d)$ *by*

$$U \;=\; \Big\{ (\tau, R) \in E(d) \;\Big|\; \|R - I_d\|_2 < \sqrt{2} \Big\}.$$

Then the principal logarithm

$$\log : U \longrightarrow \mathfrak{e}(d), \qquad (\tau, R) \longmapsto (\tau_0', B)$$

is well-defined with $\|B\|_2 \le \frac{\pi}{2}$.

Moreover,

$$\|R - I_d\|_2 + \|\tau\| \;\le\; \|B\|_2 + \|\tau_0'\| \;\le\; \tfrac{\pi}{2}\big(\|R - I_d\|_2 + \|\tau\|\big). \tag{14}$$

Proof We again use the technique of eigen-angles of orthogonal matrices. Over $\mathbb{C}$, the eigenvalues of R are $e^{\pm i\theta_j}$ with $0 \le \theta_j < \pi$. Equivalently, in an orthonormal basis R is block-diagonal with 2×2 rotation blocks

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

(and acts as the identity on the orthogonal complement). We define B in the same orthonormal basis by 2×2 skew blocks

$$\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}$$

(and is zero on the orthogonal complement). Since $\log : U \to V := \mathrm{Im}(\log)$ is bijective and $D\exp_{(\tau_0', B)}$ is invertible for all $(\tau_0', B) \in V$, the inverse function theorem makes $\exp : V \to U$ a bijective local diffeomorphism, hence its inverse $\log$ is smooth and $\log : U \to V$ is a diffeomorphism.

Since $R - I_d$ is normal,

$$\|R - I_d\|_2 = \max_j |e^{i\theta_j} - 1| = 2\max_j \sin\!\big(\tfrac{\theta_j}{2}\big) = 2\sin\!\big(\tfrac{\|B\|_2}{2}\big),$$

which yields

$$\|R - I_d\|_2 \;\le\; \|B\|_2 \;\le\; \frac{\pi}{2}\,\|R - I_d\|_2$$

from $\sin t \le t$ and $\sin t \ge \frac{2}{\pi} t$ on $[0, \frac{\pi}{2}]$ (hence also on $[0, \frac{\pi}{4}]$).

On each such 2-dimensional invariant subspace,

$$J(B) := (e^B - I)B^{-1} \quad \text{acts by} \quad \frac{e^{i\theta} - 1}{i\theta} \;=\; e^{i\theta/2}\,\frac{2\sin(\theta/2)}{\theta},$$

so $\|J(B)\|_2 \le 1$. Moreover,

$$\|J(B)^{-1}\|_2 \le \sup_{\theta \in [0, \|B\|_2]} \frac{\theta}{2\sin(\theta/2)} \le \frac{\pi}{2\sqrt{2}}$$

giving

$$\|\tau\|_2 \;\le\; \|\tau_0'\|_2 \;\le\; \frac{\pi}{2\sqrt{2}}\,\|\tau\|$$

Consequently,

$$\|R - I_d\|_2 + \|\tau\|_2 \;\le\; \|B\|_2 + \|\tau_0'\| \;\le\; \frac{\pi}{2}\big(\|R - I_d\|_2 + \|\tau\|\big).$$

■

Lemma 32 *There exists* $\varepsilon_0 > 0$ *such that for any Euclidean transformation*

$$R_\tau := (\tau, R) \in E(d), \qquad \|\tau\|_2 + \|R - I_d\|_2 \le \varepsilon_0,$$

one has

$$\mathrm{dist}_H\big(\Delta_0, R_\tau[\Delta_0]\big) \;<\; d^2(d+1)\,\big(\mathrm{diam}(\Delta_0 \cup R_\tau[\Delta_0]) - 1\big).$$

Proof Since the proof is technical, we split it into several steps. The guiding idea is this: any Euclidean motion near the identity can be written as $\exp\big(t(\tau_0, A)\big)$ with a *normalized* Lie–algebra direction $\|A\|_2 + \|\tau_0\| = 1$ and a small scalar $t > 0$. We then show that, regardless of the chosen normalized direction (τ_0, A), both the Hausdorff distance $\mathrm{dist}_H(\Delta_0, \Delta')$ and the diameter $\mathrm{diam}(\Delta_0 \cup \Delta') - 1$ are linear in t to first order, with uniform $O(t^2)$ remainders; moreover, the corresponding linear coefficients are *uniformly bounded away from zero*. This uniform control allows us to absorb the quadratic errors and conclude the desired inequality.

Step 1: Reduction to the Lie–algebra level and normalization.
By Lemma 31 we may work directly at the Lie–algebra level. Let $(\tau_0', B) \in \mathfrak{e}(d)$ with

$$\|B\|_2 + \|\tau_0'\|_2 \;\le\; \varepsilon_0,$$

and set $R_\tau := \exp\big((\tau_0', B)\big)$. It then suffices to prove that

$$\mathrm{dist}_H\big(\Delta_0, R_\tau[\Delta_0]\big) \;<\; d^2(d+1)\,\big(\mathrm{diam}(\Delta_0 \cup R_\tau[\Delta_0]) - 1\big).$$

Writing $t := \|\tau_0'\| + \|B\|_2 > 0$ and $(\tau_0, A) := (\tau_0', B)/t$, we obtain the normalized form

$$(\tau, R) = \exp\big(t(\tau_0, A)\big), \qquad \|A\| + \|\tau_0\| = 1.$$

Denote $R_\tau(\Delta_0)$ by $\Delta(t)$.

Step 2: First–order expansions for vertices and the translation.
From now on, we assume $t \le \frac{1}{2}$. Expanding the integral, one obtains

$$\tau = t\,\tau_0 + \tfrac{t^2}{2}\,A\tau_0 + t^3 \cdot \tau_t, \qquad \|\tau_t\|_2 < \frac{1}{6}.$$

For each vertex x_i set

$$y_i(t) \;:=\; e^{tA}x_i + \xi(t).$$

Introduce the notation

$$u_{ij} := x_i - x_j, \qquad v_i := Ax_i + \tau_0.$$

Since we normalized (τ_0, A) so that $\|A\|_2 + \|\tau_0\|_2 = 1$, and because $\|x_i\|_2 = \sqrt{\frac{d}{2(d+1)}}$, we have the uniform bound

$$\|v_i\|_2 \leq 1.$$

Using the Taylor expansion

$$e^{tA}x_j \;=\; x_j + t\,Ax_j + \tfrac{t^2}{2}\,A^2x_j + t^3x_j(t), \qquad \|x_j(t)\| \leq \frac{1}{6},$$

together with the expansion of $\xi(t)$, we obtain

$$y_j(t) - x_i = -u_{ij} \;+\; t\,v_j \;+\; \tfrac{t^2}{2}\,Av_j \;+\; t^3v_j(t), \qquad \|v_j(t)\|_2 \leq \frac{1}{3}.$$

Step 3: Expansions for dist_H **and** diam **with uniform remainders.** Observe that $\|y_i(t) - x_i\|^2 = \langle y_i(t) - x_i, y_i(t) - x_i\rangle$. A direct calculation yields for $i = j$ (since $u_{ii} = 0$),

$$\|y_i(t) - x_i\|^2 = \|v_i\|^2\,t^2 + r_i(t)\cdot t^4, \qquad |r_i(t)| \leq 1$$

and for $i \neq j$,

$$\|y_j(t) - x_i\|^2 = 1 + 2\langle -u_{ij},\, v_j\rangle\,t + f_{ji}(t)\cdot t^2, \qquad |f_{ji}(t)| \leq 3.$$

Hence, for all sufficiently small t (e.g., $|t| \leq 1$ under the normalization $\|A\| + \|\tau_0\| = 1$),

$$\mathrm{dist}_H\big(\Delta_0, \Delta(t)\big) = \Big(\max_i \|v_i\|\Big)\,t \;+\; r_{\mathrm{dist}}(t)\,t^2, \qquad |r_{\mathrm{dist}}(t)| \leq 2,$$

and

$$\mathrm{diam}\big(\Delta_0 \cup \Delta(t)\big) = 1 \;+\; \Big(\max_{i,j}\langle -u_{ij},\, v_j\rangle\Big)\,t \;+\; r_{\mathrm{diam}}(t)\,t^2, \qquad |r_{\mathrm{diam}}(t)| \leq 2.$$

Step 4: Coercivity linking $\max_j \|v_j\|$ **and** $\max_{i,j}\langle -u_{ij}, v_j\rangle$**.** Next, we show that

$$\max_j \|v_j\|_2 \;<\; (d+1)d^2 \cdot \max_{i,j}\,\langle -u_{ij},\, v_j\rangle.$$

Note that for any regular simplex $\{x_1, \ldots, x_d\} = \Delta'' \subset \mathbb{R}^{d-1}$ with edge length 1 and centroid at the origin, for any $s \in \mathbb{R}^{d-1}$,

$$\sum_{i=1}^{d}\langle s, x_i\rangle^2 = \tfrac{1}{2}\|s\|^2.$$

This follows once more from Remark 26: the vertices of a regular simplex form a tight frame with frame constant $\frac{1}{2}$.

Let us fix the vertex x_j and denote $\frac{x_j}{\|x_j\|}$ by e_j. Now let us decompose any vector into parts lying at $\mathbb{R}\cdot e_j$ and $e_j^{\perp}$:

$$v_j = Ax_j + \tau_0, \qquad v_j = a\cdot e_j + v_j^{\perp}, \quad v_j^{\perp} \in e_j^{\perp}$$

and

$$u_{ij} = \gamma e_j + t_i, \qquad t_i \in e_j^{\perp}, \quad \gamma = -\sqrt{\tfrac{d+1}{2d}}$$

Note that the vectors t_i form a regular simplex with edge 1 in the $d-1$ dimensional subspace of $\mathbb{R}^d$. Hence

$$\|v_j^\perp\|^2 = 2 \cdot \sum_{i \neq j} \langle v_j^\perp, t_i \rangle^2 = 2 \cdot \sum_{i \neq j} \langle v_j, u_{ij} \rangle^2 - a^2 \cdot (d+1),$$

and hence

$$\|v_j\|^2 = 2 \cdot \sum_{i \neq j} \langle v_j, -u_{ij} \rangle^2 - d \cdot a^2.$$

Observe that

$$\sum_j \sum_{i \neq j} \langle v_j,\, -u_{ij} \rangle = \sum_j \sum_{i \neq j} \langle v_j,\, x_j - x_i \rangle = (d+1) \sum_j \langle v_j,\, x_j \rangle = (d+1) \sum_j \langle \tau_0,\, x_j \rangle = 0,$$

since $\langle Ax_j, x_j \rangle = 0$[2] and $\sum_j x_j = 0$.

Partition the multiset $\{\langle v_j, -u_{ij} \rangle \;:\; j,\; i \neq j\}$ into positives $a_1, \ldots, a_k > 0$ and negatives $b_1, \ldots, b_m < 0$, with $k + m = (d+1)d$. Because

$$\sum_{\ell=1}^{k} a_\ell \;=\; \sum_{r=1}^{m} |b_r|,$$

we have

$$k \,\max_\ell a_\ell \;\geq\; \sum_{\ell=1}^{k} a_\ell \;=\; \sum_{r=1}^{m} |b_r| \;\geq\; \max_r |b_r| \quad\Longrightarrow\quad k^2 \big(\max_\ell a_\ell\big)^2 \;\geq\; \big(\max_r |b_r|\big)^2.$$

Hence, with $N := (d+1)d$,

$$N^2 \left(\max_{i,j} \langle v_j,\, -u_{ij} \rangle\right)^2 \;\geq\; \max_{i,j} \langle v_j,\, -u_{ij} \rangle^2.$$

Consequently,

$$\|v_j\|^2 \;\leq\; 2d \,\max_{i,j} \langle v_j,\, -u_{ij} \rangle^2 \;\leq\; 2d\, N^2 \left(\max_{i,j} \langle v_i,\, -u_{ij} \rangle\right)^2,$$

and therefore

$$\max_j \|v_j\| \;<\; (d+1)d^2 \;\max_{i,j} \langle v_i,\, -u_{ij} \rangle.$$

Step 5: Lower control on $\max_j \|v_j\|$.

We estimate $\max_j \|v_j\|$ using the Frobenius/operator–norm relation and $XX^\top = \frac{1}{2} I_d$.
As before,

$$(d+1)\, \max_j \|v_j\|^2 \;\geq\; \sum_{j=0}^{d} \|v_j\|^2 \;=\; \|AX + \tau_0 \mathbf{1}^\top\|_{\mathrm{Fr}}^2 \;=\; \|AX\|_{\mathrm{Fr}}^2 \;+\; (d+1)\, \|\tau_0\|^2,$$

where the cross term vanishes since $\sum_j x_j = 0$.

2. This is because $A \in \mathfrak{so}(d)$ which means it is skew-symmetric.

On the other hand,

$$\|AX\|_{\mathrm{Fr}}^2 = \mathrm{tr}\big(X^\top A^\top AX\big) = \mathrm{tr}\big(A\,XX^\top A^\top\big) = \tfrac{1}{2}\ \mathrm{tr}(AA^\top) = \tfrac{1}{2}\,\|A\|_{\mathrm{Fr}}^2 \ \geq\ \tfrac{1}{2}\,\|A\|_2^2.$$

Hence

$$\big(\max_j \|v_j\|\big)^2 \ \geq\ \|\tau_0\|^2 \ +\ \frac{1}{2(d+1)}\,\|A\|_2^2.$$

With the normalization $\|A\|_2 + \|\tau_0\| = 1$, write $a := \|A\|_2$, $b := \|\tau_0\|$ with $a, b \geq 0$ and $a + b = 1$. Then

$$\big(\max_j \|v_j\|\big)^2 \ \geq\ b^2 + \frac{1}{2(d+1)}\,a^2 \ \geq\ \min_{a+b=1}\Big(b^2 + \frac{1}{2(d+1)}\,a^2\Big) = \frac{1}{2d+3},$$

so

$$\max_j \|v_j\|_2 \ \geq\ \frac{1}{\sqrt{2d+3}}.$$

Uniform choice of t_0. Set $\alpha := \max_{i,j}\langle -u_{ij}, v_j\rangle$. By Steps 4–5 we have the uniform lower bound

$$\alpha \ \geq\ \frac{1}{(d+1)d^2\sqrt{2d+3}}.$$

Fix

$$t_0 := \frac{1}{4\,(d+1)d^2\sqrt{2d+3}}.$$

Then for every $t \in (0, t_0]$ and every normalized pair (A, τ_0), the quadratic remainders are uniformly bounded by the linear terms:

$$2t^2 \ \leq\ \tfrac{1}{2}\,\alpha t \quad \text{and} \quad 2t^2 \ \leq\ \tfrac{1}{2}\,(d+1)d^2\,\alpha t.$$

Hence

$$\mathrm{diam}(\Delta_0 \cup \Delta(t)) - 1 \ \geq\ \alpha t - 2t^2 \ \geq\ \tfrac{1}{2}\,\alpha t,$$

and

$$\mathrm{dist}_H(\Delta_0, \Delta(t)) \ \leq\ (d+1)d^2\,\alpha t + 2t^2 \ \leq\ (d+1)d^2\,\alpha t + \tfrac{1}{2}\,(d+1)d^2\,\alpha t.$$

Combining two inequalities yeilds

$$\mathrm{dist}_H\big(\Delta_0, \Delta(t)\big) \ \leq\ (d+1)d^2\,\big(\mathrm{diam}(\Delta_0 \cup \Delta(t)) - 1\big),$$

for all $t \in (0, t_0]$. ■

Appendix F. Covering number of the sphere

We work on the unit sphere $S^{d-1} \subset \mathbb{R}^d$ with the angular metric

$$\rho(x, y) \ :=\ \angle(x, y) \ :=\ \arccos(\langle x, y\rangle) \in [0, \pi].$$

For $\alpha > 0$, let $\mathcal{N}_\angle(S^{d-1}, \alpha)$ be the smallest N such that S^{d-1} can be covered by N angular balls of radius α (equivalently, by spherical caps of angular radius α). We write $|S^k|$ for the surface area of S^k.

Lemma 33 (Bounds on the covering number) *For every $d \geq 2$ and every $\alpha \in (0, \pi/2]$,*

$$\sqrt{d}\,\alpha^{-(d-1)} \;<\; \mathcal{N}_{\angle}(S^{d-1}, \alpha) \;<\; 2^{2d}\,\alpha^{-(d-1)}. \tag{15}$$

Proof Fix $\alpha \in (0, \pi/2]$. For a center $u \in S^{d-1}$ define the spherical cap

$$C_\alpha(u) \;:=\; \{x \in S^{d-1} : \rho(x, u) \leq \alpha\}.$$

By symmetry, the surface area of $C_\alpha(u)$ does not depend on u; we denote it by $|C_\alpha|$.

Its surface area can be computed using spherical coordinates in a standard way (see e.g. Li (2011)):

$$|C_\alpha| = \frac{2\pi^{(d-1)/2}}{\Gamma(\frac{d-1}{2})} \int_0^\alpha \sin^{d-2}\theta\, d\theta.$$

The surface area of the sphere is

$$|S^{d-1}| = \frac{2\pi^{d/2}}{\Gamma(d/2)},$$

where Γ denotes the Gamma function.

We claim that for every $\alpha \in (0, \pi/2]$,

$$\frac{|S^{d-1}|}{|C_\alpha|} \;\leq\; \mathcal{N}_{\angle}(S^{d-1}, \alpha) \;\leq\; \frac{|S^{d-1}|}{|C_{\alpha/2}|}. \tag{16}$$

Indeed, if $S^{d-1} \subseteq \bigcup_{i=1}^N C_\alpha(u_i)$, then by subadditivity of area,

$$|S^{d-1}| \leq \sum_{i=1}^N |C_\alpha(u_i)| = N|C_\alpha|,$$

hence $N \geq |S^{d-1}|/|C_\alpha|$.

For the upper bound, take a maximal α-separated set $\{x_1, \ldots, x_m\} \subset S^{d-1}$, so that $\rho(x_i, x_j) > \alpha$ for $i \neq j$. Maximality implies that $\{C_\alpha(x_i)\}_{i=1}^m$ covers S^{d-1}, hence $\mathcal{N}_{\angle}(S^{d-1}, \alpha) \leq m$. Moreover, the caps $C_{\alpha/2}(x_i)$ are pairwise disjoint, so

$$m\,|C_{\alpha/2}| \leq |S^{d-1}|,$$

and therefore $m \leq |S^{d-1}|/|C_{\alpha/2}|$. This proves Equation (16).

We now derive explicit bounds. Writing

$$\frac{|S^{d-1}|}{|C_\alpha|} = \frac{\sqrt{\pi}\,\Gamma(\frac{d-1}{2})/\Gamma(\frac{d}{2})}{\int_0^\alpha \sin^{d-2}\theta\, d\theta},$$

and using that for $0 \leq \theta \leq \pi/2$,

$$\frac{\theta}{2} \leq \sin\theta \leq \theta,$$

we obtain

$$\frac{\alpha^{d-1}}{(d-1)2^{d-2}} \;<\; \int_0^\alpha \sin^{d-2}\theta\, d\theta \;<\; \frac{\alpha^{d-1}}{d-1}.$$

Hence

$$\frac{|S^{d-1}|}{|C_\alpha|} > \frac{\sqrt{\pi}(d-1)\,\Gamma(\frac{d-1}{2})/\Gamma(\frac{d}{2})}{\alpha^{d-1}},$$

and

$$\frac{|S^{d-1}|}{|C_{\alpha/2}|} < \frac{\sqrt{\pi}(d-1)2^{2d-3}\,\Gamma(\frac{d-1}{2})/\Gamma(\frac{d}{2})}{\alpha^{d-1}}.$$

Let $x = \frac{d-1}{2}$. Gautschi's inequality implies that for all $x > \frac{1}{2}$,

$$\frac{1}{\sqrt{x+\frac{1}{2}}} \;\leq\; \frac{\Gamma(x)}{\Gamma(x+\frac{1}{2})} \;\leq\; \frac{1}{\sqrt{x-\frac{1}{2}}}.$$

Substituting $x = \frac{d-1}{2}$ yields

$$\sqrt{\frac{2}{d}} \;\leq\; \frac{\Gamma(\frac{d-1}{2})}{\Gamma(\frac{d}{2})} \;\leq\; \sqrt{\frac{2}{d-2}}.$$

Combining these inequalities gives

$$\sqrt{d}\,\alpha^{-(d-1)} \;<\; \mathcal{N}_\angle(S^{d-1},\alpha) \;<\; 2^{2d}\,\alpha^{-(d-1)},$$

as claimed. ■

Appendix G. Simplex rotation lemmas

This appendix collects two technical lemmas used in the lower-bound proof of Main Theorem 2. Together they realize the adversary's rotation step: given a direction v along which the current simplex Δ is nearly aligned, Lemma 34 constructs an explicit isometry R_θ that tilts Δ away from v by a controlled angle, and Lemma 35 controls how far the resulting rotated simplex moves in the ambient space.

Lemma 34 (Simplex rotation) *Consider a regular simplex $\Delta = \{0, x_1, \ldots, x_d\}$ of edge length 1, an angle $\theta \leq \pi/18$, and a direction v. Without loss of generality, assume the edge $0x_1$ satisfies*

$$\langle v, x_1 - 0\rangle = \max_{i,j}\langle v, x_i - x_j\rangle, \qquad \langle v, x_1\rangle > \cos\theta.$$

Then there exists an isometry $R_\theta : \mathbb{R}^d \to \mathbb{R}^d$ such that the rotated simplex $\Delta' := R_\theta\Delta$ satisfies

$$\langle v, x_1' - 0\rangle = \max_{i,j}\langle v, x_i' - x_j'\rangle, \qquad x_i' := R_\theta x_i.$$

Moreover,

$$\langle v, x_1' - 0\rangle \leq \cos\theta.$$

Proof Consider the rotation $R_\theta : \mathbb{R}^d \to \mathbb{R}^d$ that acts in the plane $\Pi = \mathrm{span}\{v, x_1\}$ as a rotation by angle θ from v toward x_1 along the smaller angle between them, and acts as the identity on the orthogonal complement $\Pi^\perp$. If $v = x_1$, take any plane $\Pi \ni x_1$ and either of the two possible rotations.

In an orthonormal basis $d_1, \dots, d_d$ with $\operatorname{span}\{d_1, d_2\} = \Pi$, $d_1 = v$, and

$$x_1 = k_1 d_1 + k_2 d_2 \quad \text{with } k_1, k_2 > 0,$$

the map R_θ is represented by

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & \cdots & 0 \\ \sin\theta & \cos\theta & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & 1 \end{pmatrix}. \tag{17}$$

First, note that

$$1 \geq k_1 = \langle x_1, v\rangle \geq \cos\theta, \qquad 0 \leq k_2 = \sqrt{1 - \langle x_1, v\rangle^2} \leq \sin\theta,$$

since $\langle d_1, x_1\rangle = k_1$ and $k_1^2 + k_2^2 = 1$. Also, $\langle x_1, x_i - x_j\rangle \in \{0, \pm\frac{1}{2}\}$.

A direct calculation gives

$$\cos 2\theta \;=\; \cos^2\theta - \sin^2\theta \;\leq\; \langle x_1', v\rangle \;\leq\; \cos\theta.$$

Set $w' := v - x_1'$. Then $\|w'\|_2 \leq \sqrt{2 - 2\cos\theta}$, and hence

$$\langle v, x_i' - x_j'\rangle \;=\; \langle x_1' + w', x_i' - x_j'\rangle \;\leq\; \tfrac{1}{2} + \|w'\|_2 \;\leq\; \cos 2\theta \;\leq\; \langle x_1', v\rangle$$

for $\theta \leq \pi/18$. The case $x_1 = v$ is identical. This proves the claim. ■

Lemma 35 (Rotation and neighborhood of the simplex) *Fix two positive numbers α, α'. Assume the isometry R_θ from Equation (17) satisfies*

$$B(R_\theta x_1, \alpha') \subset B(x_1, \alpha).$$

Denote $\Delta' := R_\theta(\Delta)$. Then

$$\Delta' + B(\alpha') \subset \Delta + B(\alpha).$$

Proof To ensure the inclusion $\Delta' + B(\alpha') \subset \Delta + B(\alpha)$, it suffices to check that no vertex of the simplex moves by more than $\alpha - \alpha'$ under the rotation R_θ.

Fix any vertex $x_i \neq x_1$, and decompose it as $x_i = w + p$, where $w \in \Pi$ and $p \in \Pi^\perp$. Note that $\|w\|_2 \leq 1$, since $\|x_i\|_2^2 = \|w\|_2^2 + \|p\|_2^2 = 1$. Moreover,

$$x_i' = R_\theta[w] + p,$$

so

$$\|x_i - x_i'\|_2^2 = \|w - R_\theta[w]\|_2^2 = 2\|w\|_2^2 - 2\langle w, R_\theta[w]\rangle = 2\|w\|_2^2\,(1 - \cos\theta).$$

Since

$$\|x_1 - x_1'\|_2^2 = 2(1 - \cos\theta)$$

and $\|w\|_2 \leq 1$, we obtain

$$\|x_i - x_i'\|_2 \leq \|x_1 - x_1'\|_2.$$

By the assumption of the lemma, vertex x_1 moves by at most $\alpha - \alpha'$, and hence every other vertex does as well. This proves the claim. ■