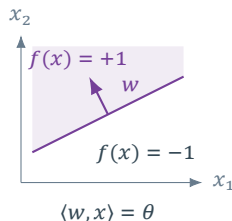


Problem setup

Boolean halfspaces

$$f(x) = \text{sgn}(\langle w, x \rangle - \theta).$$

$(X, Y) \sim \mathcal{D}$ on $\{\pm 1\}^n \times \{\pm 1\}$, with uniform X -marginal.



Smoothed benchmark [KM25]: flip each bit independently with probability σ .

$$\text{opt}_\sigma = \inf_{f \in \text{Halfspaces}} \mathbb{E}_{Z \sim B(\sigma)^n} \Pr[f(X \odot Z) \neq Y].$$

Goal: find $h : \{\pm 1\}^n \rightarrow \{\pm 1\}$ such that $\Pr[h(X) \neq Y] \leq \text{opt}_\sigma + \varepsilon$.

For $\rho = 1 - 2\sigma$, the noise operator averages over these bit flips:

$$(T_\rho f)(x) = \mathbb{E}_Z[f(x \odot Z)], \quad \widehat{T_\rho f}(S) = \rho^{|S|} \widehat{f}(S).$$

Statistical Query (SQ) learning

Access the distribution through approximate expectations of bounded queries:

$$q : \mathcal{X} \times \{\pm 1\} \rightarrow [-1, 1], \quad \text{STAT}(\tau) : \quad |v - \mathbb{E}_{(X,Y) \sim \mathcal{D}}[q(X,Y)]| \leq \tau.$$

Queries may be adaptive. One query can be simulated with $\tilde{O}(1/\tau^2)$ labeled samples.

Most standard learning techniques fit the SQ model.

Examples include gradient-based methods, spectral methods, EM, and many convex optimization algorithms.

Why SQ lower bounds matter

They give unconditional barriers for this broad class of algorithms: either many queries or very high precision is necessary.

Reyzin (2020), *Statistical Queries and Statistical Algorithms: Foundations and Applications*.

arXiv:2004.00557. SQ implementations depend on the algorithm and the precision model.

Upper and lower bounds for smoothed agnostic learning of halfspaces

Upper bounds

Work	Marginal	Complexity
[KW25]	subexp. $\mathbb{R}^n$	$n^{\tilde{O}(\log(1/\epsilon)/\sigma^2)}$
[KM25]	subexp. $\{\pm 1\}^n$	$\tilde{O}(n^{\text{poly}(1/(\sigma\epsilon))})$
This work	uniform $\{\pm 1\}^n$	$\tilde{O}(n^{O(\log(1/\epsilon)/\sigma)})$

SQ complexity lower bounds

Work	Setting	Complexity
[DK26]	continuous	$n^{\Omega(1/\sigma^2 + \log(1/\epsilon))}$
This work	uniform cube	$n^{\Omega(\log(1+\sigma/\epsilon^2)/\sigma)}$

[KW25] Koehler–Wu, STOC 2026. [KM25] Kou–Meka, NeurIPS 2025. [DK26] Diakonikolas–Kane, 2026.

Our key ingredient: a bounded dual witness

For $m = o(\sqrt{n})$, construct $\psi : \{\pm 1\}^n \rightarrow [-1, 1]$ with

$$\psi \perp \mathcal{P}_{\leq m}, \quad \langle T_\rho \text{Maj}_n, \psi \rangle \gtrsim \frac{\rho^{2m+1}}{\sqrt{m}}.$$

Weak duality turns correlation into an approximation lower bound:

$$\inf_{\deg p \leq m} \|T_\rho \text{Maj}_n - p\|_1 \gtrsim \frac{\rho^{2m+1}}{\sqrt{m}}.$$

Obstacle: Fourier modes on the cube are not uniformly bounded.

Our construction: a sine kernel with Laguerre weights.

Approximate-resilience framework: [DFTWW15]. Gaussian smoothed-learning comparison: [DK26].

The sine-Laguerre construction

1. Align Fourier signs

$$\phi_{y,n}(x) = \sin\left(\arctan \sqrt{y/n} \sum_i x_i\right).$$

This is odd, symmetric and bounded. Fourier signs align with majority.

2. Cancel low degrees

$$\Omega_{\ell,n} = (-1)^\ell \int_0^\infty L_\ell^{(1/2)}(y) (1 + y/n)^{n/2} e^{-y} \phi_{y,n} dy.$$

Set $k = \lfloor (m+1)/2 \rfloor$ and $\tilde{\Omega} = \frac{1}{2}(\Omega_{k,n} + \Omega_{k+1,n})$. Laguerre orthogonality gives $\tilde{\Omega} \perp \mathcal{P}_{\leq m}$.

3. Control the norm

With $y = s^2$, a Hermite-function bound yields $\|\tilde{\Omega}\|_\infty = O(1)$. Normalize to $\psi = \tilde{\Omega}/\|\tilde{\Omega}\|_\infty$.

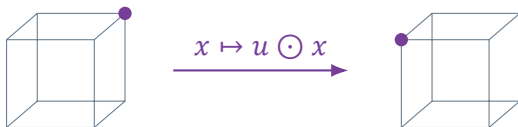
Hide the witness in many directions

$$\psi^{(u)}(x) = \psi(u \odot x), \quad \mathcal{D}_u(x, y) = 2^{-(n+1)}(1 + y\psi^{(u)}(x)).$$

Inputs remain uniform. Labels hide u through $\mathbb{E}[Y \mid X = x] = \psi(u \odot x)$.

$$x = (+, +, +)$$

$$u \odot x = (-, +, -)$$



Choose $|\mathcal{U}| = 2^{\Omega(\sqrt{n})}$ near-orthogonal directions. Then

$$\chi_{\mathcal{D}_0}(\mathcal{D}_u, \mathcal{D}_v) = |\langle \psi^{(u)}, \psi^{(v)} \rangle| \leq n^{-\Omega(m)} \quad (u \neq v).$$

Hidden-direction paradigm: [DKS17, DK26]. Hypercube sign-flip embedding: this work.

The SQ lower bound

A learner solves the hidden-direction search problem when $\kappa_m \geq 4\varepsilon$. Choose

$$m = \Theta\left(\frac{\log(1 + \sigma/\varepsilon^2)}{\sigma}\right).$$

The many-vs-one principle gives (via a standard SDA bound):

$$2^{\Omega(\sqrt{n})} \text{ queries} \quad \text{or} \quad \tau \leq n^{-\Omega(\log(1+\sigma/\varepsilon^2)/\sigma)}.$$

For $\sigma \in (0, 0.499]$ and $1/\text{poly}(n) \leq \varepsilon \leq \varepsilon_0$.

L^1 regression is essentially optimal.

It gives the upper bound $\tilde{O}(n^{O(\log(1/\varepsilon)/\sigma)})$. The remaining exponent gap is discussed in the paper.

[DFTWW15] Approximate resilience, SODA 2015. [DKS17] SQ lower bounds for robust estimation, FOCS 2017.

[DK26] SQ lower bounds for smoothed agnostic learning, 2026. [KM25] Smoothed agnostic learning on the hypercube, NeurIPS 2025.

[FGRVX17] Statistical algorithms and planted cliques, J. ACM 2017.