What's Behind the Oracle?

Discovering Fourier Structure from Queries

Arijit Ghosh

Indian Statistical Institute, Kolkata

Papers Covered in This Talk

1. **Price of Parsimony: Complexity of Fourier Sparsity Testing**
   Arijit Ghosh, Manmatha Roy
   NeurIPS 2025

2. **Testing Fourier Sparsity via Implicit Sensing**
   Arijit Ghosh, Subhamoy Maitra, Manmatha Roy
   ICLR 2026

3. **On Nonlinearity Estimation Problem**
   Arijit Ghosh, Subhamoy Maitra, Manmatha Roy — Manuscript 2026

In all papers, authors are listed in alphabetical order of surnames.

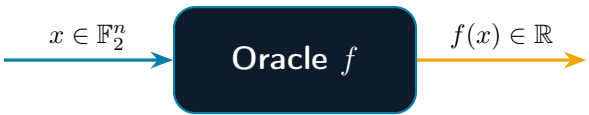
Part 1 of 6

Setup

Talk Roadmap

1. Setup
2. Fourier Sparsity Testing (ℓ_2)
3. Fourier Bias Estimation
4. Fourier Sparsity Testing (ℓ_0)
5. Lower Bounds
6. Open Problems

Model of Computation



No access to internal structure

Goal

Learn global properties of f using as **few queries as possible**, with complexity **independent of the dimension** n.

Motivation for Model of Computation

Modern systems generate data at astronomical scale

- Network traffic at routers and data centers
- High-volume logs from distributed systems
- Continuous streams of sensor and telemetry data

In such settings, we cannot afford to:

- store the entire dataset
- perform expensive exact computations

Example: Distinct Elements

A typical internet router may observe billions of packets per day. Compute the exact number of distinct IP addresses seen far.

Motivation for Model of Computation

Modern systems generate data at astronomical scale

- Network traffic at routers and data centers
- High-volume logs from distributed systems
- Continuous streams of sensor and telemetry data

In such settings, we cannot afford to:

- store the entire dataset
- perform expensive exact computations

Example: Distinct Elements

A typical internet router may observe billions of packets per day. Compute the exact number of distinct IP addresses seen far.

Property Testing: Deciding Without Reading the Entire Input

Property Testing

Given query access to an object, distinguish between:

- objects that satisfy a desired property, and
- objects that are ε-far from satisfying it

Property Testing: Deciding Without Reading the Entire Input

Property Testing

Given query access to an object, distinguish between:

- objects that satisfy a desired property, and
- objects that are ε-far from satisfying it

- Rooted in ideas from PCPs and program checking.
- Now a central algorithmic paradigm concerned with computation under limited data access.

Examples of Properties

Graph Setting:

- Is the graph connected?
- Does the graph contain a large independent set?
- Is the graph bipartite?

Examples of Properties

Graph Setting:

- Is the graph connected?
- Does the graph contain a large independent set?
- Is the graph bipartite?

Function Setting:

- Is the function close to a low-degree polynomial?
- Does the function have a sparse Fourier representation?
- Is the function monotone?

Problems Considered in This Talk

Two Spectral Questions

- **Fourier Bias** (Cryptography)
  - Measures correlation with linear functions.
  - Central to pseudorandomness and cryptanalysis.
- **Fourier Sparsity** (Learning Theory)
  - Counts significant Fourier coefficients.
  - Captures spectral complexity and learnability.

Goal: Design sublinear-time algorithms that estimate these properties using only limited access.

Parity Functions

For each $\alpha \in \mathbb{F}_2^n$, define

$$\chi_\alpha(x) = (-1)^{\langle \alpha, x \rangle}.$$

Fourier Expansion

Every Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ can be uniquely represented as

$$f(x) = \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}(\alpha) \chi_\alpha(x),$$

The coefficient $\widehat{f}(\alpha) = \mathbb{E}_x[f(x)\chi_\alpha(x)]$ measures the correlation of f with the parity χ_α.

Fourier Bias and Fourier Sparsity

$$f(x) = \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}(\alpha)\chi_\alpha(x).$$

Fourier Bias and Fourier Sparsity

$$f(x) = \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}(\alpha)\chi_\alpha(x).$$

Fourier Bias

The Fourier bias of f is

$$\mathrm{FBias}(f) = \max_{\alpha \in \mathbb{F}_2^n} |\widehat{f}(\alpha)|.$$

It measures the strongest correlation of f with any parity function.

Fourier Bias and Fourier Sparsity

$$f(x) = \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}(\alpha)\chi_\alpha(x).$$

Fourier Bias

The Fourier bias of f is

$$\mathrm{FBias}(f) = \max_{\alpha \in \mathbb{F}_2^n} |\widehat{f}(\alpha)|.$$

It measures the strongest correlation of f with any parity function.

Fourier Sparsity

The Fourier sparsity of f is

$$\mathrm{spar}(f) = \left|\{\alpha \in \mathbb{F}_2^n : \widehat{f}(\alpha) \neq 0\}\right|.$$

It counts the number of nonzero Fourier coefficients.

Questions Studied in This Work

Question 1: Estimating Fourier Bias

Given oracle access to $f : \mathbb{F}_2^n \to \{-1, +1\}$, can we estimate Fourier Bias within accuracy τ?

Questions Studied in This Work

Question 1: Estimating Fourier Bias

Given oracle access to $f : \mathbb{F}_2^n \to \{-1, +1\}$, can we estimate Fourier Bias within accuracy τ?

Question 2: Testing Fourier Sparsity (ℓ_2 Distance)

Given oracle access to $f : \mathbb{F}_2^n \to \mathbb{R}$, can we distinguish between

- f being s-Fourier sparse, and
- f being ε-far from every s-Fourier sparse function in ℓ_2 distance?

Questions Studied in This Work

Question 1: Estimating Fourier Bias

Given oracle access to $f : \mathbb{F}_2^n \to \{-1, +1\}$, can we estimate Fourier Bias within accuracy τ?

Question 2: Testing Fourier Sparsity (ℓ_2 Distance)

Given oracle access to $f : \mathbb{F}_2^n \to \mathbb{R}$, can we distinguish between

- f being s-Fourier sparse, and
- f being ε-far from every s-Fourier sparse function in ℓ_2 distance?

Question 3: Testing Fourier Sparsity (ℓ_0 Distance)

Given oracle access to $f : \mathbb{F}_2^n \to \{-1, +1\}$, can we distinguish between

- f being s-Fourier sparse, and
- f being ε-far from every s-Fourier sparse function in ℓ_0 distance?

The Naive Way

The Curse of Dimensionality

- $f : \mathbb{F}_2^n \to \{-1, +1\}$ has 2^n Fourier coefficients.
- For $n = 128$, this is approximately 3×10^{38} coefficients.
- Even the Fast Walsh–Hadamard Transform requires $O(n2^n)$ time and the entire truth table of f.

Computing the full spectrum is infeasible for large n.

Summary of Results

Problem	Prior best	This work	Lower bound
Bias estimation	$O(n/\tau^6)$	$\widetilde{O}(1/\tau^2)$	$\Omega(1/\tau^2)$
Sparsity testing (ℓ_2)	$\widetilde{O}(s/\varepsilon^4)$	$\widetilde{O}(s/\varepsilon^2)$	$\Omega(s)$
Sparsity testing (ℓ_0)	$\widetilde{O}(s^{14})$	$\widetilde{O}(s^4)$	$\Omega(s)$

Summary of Results

Problem	Prior best	This work	Lower bound
Bias estimation	$O(n/\tau^6)$	$\widetilde{O}(1/\tau^2)$	$\Omega(1/\tau^2)$
Sparsity testing (ℓ_2)	$\widetilde{O}(s/\varepsilon^4)$	$\widetilde{O}(s/\varepsilon^2)$	$\Omega(s)$
Sparsity testing (ℓ_0)	$\widetilde{O}(s^{14})$	$\widetilde{O}(s^4)$	$\Omega(s)$

Common features of all three results:

Summary of Results

Problem	Prior best	This work	Lower bound
Bias estimation	$O(n/\tau^6)$	$\widetilde{O}(1/\tau^2)$	$\Omega(1/\tau^2)$
Sparsity testing (ℓ_2)	$\widetilde{O}(s/\varepsilon^4)$	$\widetilde{O}(s/\varepsilon^2)$	$\Omega(s)$
Sparsity testing (ℓ_0)	$\widetilde{O}(s^{14})$	$\widetilde{O}(s^4)$	$\Omega(s)$

Common features of all three results:

- Complexity is measured by the number of queries issued to the function.

Summary of Results

Problem	Prior best	This work	Lower bound
Bias estimation	$O(n/\tau^6)$	$\widetilde{O}(1/\tau^2)$	$\Omega(1/\tau^2)$
Sparsity testing (ℓ_2)	$\widetilde{O}(s/\varepsilon^4)$	$\widetilde{O}(s/\varepsilon^2)$	$\Omega(s)$
Sparsity testing (ℓ_0)	$\widetilde{O}(s^{14})$	$\widetilde{O}(s^4)$	$\Omega(s)$

Common features of all three results:

- Complexity is measured by the number of queries issued to the function.
- It depends only on the testing parameters and is independent of n.

Summary of Results

Problem	Prior best	This work	Lower bound
Bias estimation	$O(n/\tau^6)$	$\widetilde{O}(1/\tau^2)$	$\Omega(1/\tau^2)$
Sparsity testing (ℓ_2)	$\widetilde{O}(s/\varepsilon^4)$	$\widetilde{O}(s/\varepsilon^2)$	$\Omega(s)$
Sparsity testing (ℓ_0)	$\widetilde{O}(s^{14})$	$\widetilde{O}(s^4)$	$\Omega(s)$

Common features of all three results:

- Complexity is measured by the number of queries issued to the function.
- It depends only on the testing parameters and is independent of n.
- All algorithms run in time polynomial in n and the underlying complexity parameters.
- Non-adaptive testers achieve the upper bounds, whereas the lower bounds apply to the more powerful class of adaptive testers.

The Common Algorithmic Idea

Fourier Analysis on Affine Subspace

Key Formula (Restriction)

$$\widehat{f_A}(\gamma) = \sum_{\beta \in \gamma + H^\perp} \widehat{f}(\beta)\, \chi_\beta(\alpha)$$

Each restricted Fourier coefficient is a *bucket* of global coefficients corresponding to a coset of $H^\perp$.

A Dimensionality Reduction Framework for Boolean function

Three Miracles of Random Affine Restriction

M1. Survival: every heavy Fourier coefficient of f survives restriction — the corresponding restricted coefficient concentrates tightly around the largest global coefficient in its bucket

M2. Separation: distinct heavy coefficients land in distinct cosets of $H^{\perp}$ with high probability
(collision probability $= 2^{-t}$, controlled by subspace dimension t)

M3. Stability: many small coefficients cannot aggregate to create a spurious large restricted coefficient
(controlled via the fourth Fourier moment)

These three miracles are the leitmotif of all three papers in this talk.

Part 2 of 6

Fourier Sparsity Testing (ℓ_2)

Talk Roadmap

1. Setup
2. Fourier Sparsity Testing (ℓ_2)
3. Fourier Bias Estimation
4. Fourier Sparsity Testing (ℓ_0)
5. Lower Bounds
6. Open Problems

Problem: How Far from Sparse? — Formulation

Setting: $f : \mathbb{F}_2^n \to \mathbb{R}$, $\|f\|_2 = 1$, real-valued.

Key structural lemma:

$$\operatorname{dist}_2^2(f, \mathcal{F}_s) = \min_{g \in \mathcal{F}_s} \mathbb{E}_x[(f(x) - g(x))^2] = 1 - \max_{\substack{T \subseteq \mathbb{F}_2^n \\ |T| \leq s}} \sum_{\beta \in T} \widehat{f}(\beta)^2$$

Problem: How Far from Sparse? — Formulation

Setting: $f : \mathbb{F}_2^n \to \mathbb{R}$, $\|f\|_2 = 1$, real-valued.

Key structural lemma:

$$\operatorname{dist}_2^2(f, \mathcal{F}_s) = \min_{g \in \mathcal{F}_s} \mathbb{E}_x[(f(x) - g(x))^2] = 1 - \max_{\substack{T \subseteq \mathbb{F}_2^n \\ |T| \leq s}} \sum_{\beta \in T} \widehat{f}(\beta)^2$$

Problem (Tolerant Testing)

Given parameters s, δ, ε, distinguish with probability $\geq 2/3$:

- **Close:** $\operatorname{dist}_2^2(f, \mathcal{F}_s) \leq \delta$ (top-s Fourier mass $\geq 1 - \delta$)
- **Far:** $\operatorname{dist}_2^2(f, \mathcal{F}_s) \geq \delta + \varepsilon$ (top-s Fourier mass $\leq 1 - \delta - \varepsilon$)

Inverse Problems: Recovering Hidden Structure

$$\text{Model Parameters} \xrightarrow{\text{Forward Operator}} \text{Observations}$$

$$\text{Observations} \xrightarrow{\text{Inverse Problem}} \text{Model Parameters}$$

Inverse Problems: Recovering Hidden Structure

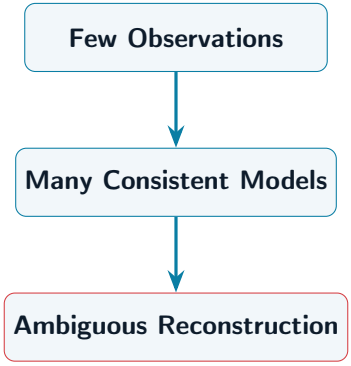
$$\text{Model Parameters} \xrightarrow{\text{Forward Operator}} \text{Observations}$$

$$\text{Observations} \xrightarrow{\text{Inverse Problem}} \text{Model Parameters}$$

Examples: MRI, signal processing, computer vision, geophysics, astronomy.

Learning is fundamentally an inverse problem.

The Fundamental Challenge: Ill-Posedness



Fundamental Question

How much information is required to uniquely recover the underlying model parameters?

Classical Answer: Nyquist–Shannon Sampling

Sampling Theorem

A band-limited signal with highest frequency N can be reconstructed exactly provided

$$f_s \geq 2N.$$

- Below the Nyquist rate, information is lost.
- Different signals may produce identical samples (aliasing).
- Reconstruction becomes impossible without additional assumptions.

Can we recover a signal from *fewer* than $2N$ measurements?

Sparsity Breaks the Barrier

- many real-world signals are sparse or approximately sparse, in the frequency domain.
- If only $k \ll N$ coefficients carry significant information, recovery is possible from far fewer than N measurements.

Sparsity Breaks the Barrier

- many real-world signals are sparse or approximately sparse, in the frequency domain.
- If only $k \ll N$ coefficients carry significant information, recovery is possible from far fewer than N measurements.

Compressed Sensing

$$\text{Sparse Structure} \implies \text{Fewer Measurements} \implies \text{Exact Recovery}$$

Applications: MRI, astronomy, geophysics, signal processing.

Fourier Sparse Signal $\Longrightarrow$ Efficient Recovery

But wait...

A Hidden Assumption

Most recovery algorithms require the sparsity level k as an input parameter.

The Missing Question

Can we determine whether a signal is sparse without first reconstructing it?

Main Result

Theorem (Upper Bound)

There is a **nonadaptive** algorithm that tolerantly tests s-Fourier sparsity with success probability $\geq 2/3$, using

$$\widetilde{O}\Big(\frac{s}{\varepsilon^2}\Big) \text{ queries.}$$

Theorem (Lower Bound)

Any randomized algorithm distinguishing s-sparse from $\frac{1}{4}$-far must use $\Omega(s)$ queries.

Both bounds are optimal up to polylogarithmic factors.

Comparison with Prior Work

Result	Query complexity	
Yaroslavtsev & Zhou 2020	$\widetilde{O}(s/\varepsilon^4)$	
This work (upper bound)	$\widetilde{O}(s/\varepsilon^2)$	$\leftarrow$ **quadratic improvement in ε**
Yaroslavtsev & Zhou (lower bd)	$\Omega(\sqrt{s})$	
This work (lower bound)	$\Omega(s)$	$\leftarrow$ **quadratic improvement**

Algorithm: FOURIER-SPARSITY-TESTER

Algorithm (3 steps)

1. **Sample:** Pick random affine subspace $A = \alpha + H \subseteq \mathbb{F}_2^n$ with $\dim(H) = \Theta\big(\log \frac{s}{\varepsilon}\big)$
2. **Estimate:** Compute estimates $\widetilde{f_A}(\gamma)$ of all Fourier coefficients of the restriction f_A using median-of-means ($\widetilde{O}(s/\varepsilon^2)$ queries total)
3. **Decide:** Output **YES** if the sum of squares of the top-s estimated coefficients is $\geq 1 - (\delta + \varepsilon/2)$

Algorithm: FOURIER-SPARSITY-TESTER — Key Lemma

Key Lemma (Spectral Concentration)

For $\gamma \in H$, $\tau > 0$, and $t \geq \log(1/\eta^4)$:

$$\Pr_{H,\alpha}\left[\left|\widehat{f_A}(\gamma) - \widehat{f}(\gamma)\chi_\gamma(\alpha)\right| > \eta + \tau\right] \leq \frac{\eta^4}{\tau^2}$$

Algorithm: FOURIER-SPARSITY-TESTER — Key Lemma

Key Lemma (Spectral Concentration)

For $\gamma \in H$, $\tau > 0$, and $t \geq \log(1/\eta^4)$:

$$\Pr_{H,\alpha}\left[\left|\widehat{f_A}(\gamma) - \widehat{f}(\gamma)\chi_\gamma(\alpha)\right| > \eta + \tau\right] \leq \frac{\eta^4}{\tau^2}$$

Proof sketch: Let $Z = \widehat{f_A}(\gamma) - \widehat{f}(\gamma)\chi_\gamma(\alpha) = \sum_{\beta \neq \gamma,\, \beta \in \gamma + H^\perp} \widehat{f}(\beta)\chi_\beta(\alpha)$.

- $\mathbb{E}[Z] \approx 0$ (character orthogonality over random α)
- $\mathrm{Var}(Z) \leq 2^{-t} = \eta^4$ (Parseval + collision probability)
- Apply Chebyshev $\Rightarrow$ the bound above

Consequence: top-s Fourier mass of f_A approximates that of f up to additive $\pm\varepsilon/4$, enabling the tester to distinguish close from far.

Part 3 of 6

Fourier Bias Estimation

Talk Roadmap

1. Setup
2. Fourier Sparsity Testing (ℓ_2)
3. Fourier Bias Estimation
4. Fourier Sparsity Testing (ℓ_0)
5. Lower Bounds
6. Open Problems

Why Do Heavy Fourier Coefficients Matter?

Heavy Fourier coefficients have played a central role in modern cryptanalysis.

Common cryptanalytic paradigm

Find Heavy Coefficient $\Longrightarrow$ Exploit Correlation

Why Do Heavy Fourier Coefficients Matter?

Why it matters in cryptography:

- **Linear cryptanalysis**: exploit $|\widehat{f}(\alpha)| = 2\varepsilon$ in block ciphers; data complexity $O(\varepsilon^{-2})$ — directly our $1/\tau^2$ bound!
- **Fast correlation attacks** on stream ciphers: LFSR + filter function; a large bias reduces the state recovery to decoding
- **Goldreich–Levin**: a heavy Fourier coefficient $\Rightarrow$ hardcore predicate is broken
- **Bit security of RSA / Discrete Log**: existence of a heavy Fourier coefficient of a bit-extraction function enables an SFT-based attack

Why Fourier Bias Estimation?

Presence of Heavy Coefficients $\Longrightarrow$ Exploit Correlation

- The cryptographic literature has extensively studied how to **exploit** heavy Fourier coefficients.
- In contrast, the complementary question of **certifying their presence** has received much less attention.
- We address this gap by developing query-efficient algorithms for estimating Fourier bias from black-box access alone.

The Problem We Consider

Our Goal

Estimate $\mathrm{FBias}(f)$ using only black-box oracle access.

$$\left|\widetilde{\mathrm{FBias}}(f) - \mathrm{FBias}(f)\right| \leq \tau.$$

This immediately yields a tester for the existence of heavy Fourier coefficients.

Prior Methods Fall Short

The Cryptanalytic Setting

f = black box (cipher or component), chosen-plaintext model $\approx$ query model.

Goal: estimate $\mathrm{FBias}(f)$ to additive error $\pm\tau$, using few queries.

Prior Methods Fall Short

The Cryptanalytic Setting

f = black box (cipher or component), chosen-plaintext model $\approx$ query model.

Goal: estimate $\mathrm{FBias}(f)$ to additive error $\pm\tau$, using few queries.

Prior methods are inadequate:

- **Full FFT:** 2^n queries — infeasible for $n = 128$
- **BLR test:** only multiplicative control ($\delta \leq p_{\mathrm{rej}} \leq 6\delta$) — useless in the low-bias regime where $\mathrm{FBias} \approx 0$
- **Goldreich–Levin:** $O(n/\tau^3)$ queries — linear in n
- **Bera et al. 2021:** $O(n/\tau^6)$ — worse in both n *and* τ

Prior Methods Fall Short

The Cryptanalytic Setting

f = black box (cipher or component), chosen-plaintext model $\approx$ query model.

Goal: estimate $\mathrm{FBias}(f)$ to additive error $\pm\tau$, using few queries.

Prior methods are inadequate:

- **Full FFT:** 2^n queries — infeasible for $n = 128$
- **BLR test:** only multiplicative control ($\delta \leq p_{\mathrm{rej}} \leq 6\delta$) — useless in the low-bias regime where $\mathrm{FBias} \approx 0$
- **Goldreich–Levin:** $O(n/\tau^3)$ queries — linear in n
- **Bera et al. 2021:** $O(n/\tau^6)$ — worse in both n *and* τ

We need a method that: (a) depends only on τ, not n, and (b) gives tight additive control.

Main Result — Fourier Bias Estimation

Theorem 1 (Upper Bound)

There exists a **nonadaptive** randomized algorithm that estimates $\mathrm{FBias}(f)$ with

$$\Pr\left[\left|\widehat{\mathrm{FBias}}(f) - \mathrm{FBias}(f)\right| \leq \tau\right] \geq \tfrac{3}{4}$$

using $\widetilde{O}(1/\tau^2)$ queries — **independent of** n.

Theorem 2 (Lower Bound)

Any randomized algorithm (even adaptive) requires $\Omega(1/\tau^2)$ queries.

Main Result — Fourier Bias Estimation

Algorithm	**Queries**	n-**independent?**
FFT	2^n	No
Goldreich–Levin	$O(n/\tau^3)$	No
Bera et al. 2021	$O(n/\tau^6)$	No
This work	$\widetilde{O}(1/\tau^2)$	**Yes**
This work	$\Omega(1/\tau^2)$	—

Problem essentially settled.

The Algorithm — Restrict and Estimate

Algorithm: Estimate-Fourier-Bias

1. Set $t = \lceil 40 + 20 \log_2(1/\tau) \rceil$
2. Sample random subspace $H \leq \mathbb{F}_2^n$ of dimension t
3. Sample $\alpha \sim \mathrm{Unif}(\mathbb{F}_2^n)$ independently; set $A = \alpha + H$
4. Estimate all $|\widehat{f_A}(\gamma)|$ for $\gamma \in W$ to within $\pm\tau/2$
5. Return $\max_\gamma |\widetilde{f_A}(\gamma)|$

Query count: estimating all 2^t coefficients of f_A to precision $\tau/2$:

$$O\left(\frac{t + \log(1/\delta)}{\tau^2}\right) = \widetilde{O}\left(\frac{1}{\tau^2}\right) \text{ queries (by Hoeffding + union bound).}$$

The Algorithm — Why It Works: Three Miracles

We need to show: $\left|\max_{\gamma \in W} |\widehat{f_A}(\gamma)| - \mathrm{FBias}(f)\right| \leq \tau$ with probability $\geq 3/4$.

This requires three things to hold simultaneously:

The Three Miracles

M1. Survival: every heavy global coefficient survives restriction — the corresponding restricted coefficient is approximately the same magnitude

M2. Separation: distinct heavy global coefficients land in distinct buckets of $H^{\perp}$ — they do not interfere with each other

M3. Stability: many small coefficients cannot aggregate into a spurious large restricted coefficient

M3 is the most delicate — it requires a fourth moment argument.

Lemma 3 (Survival)

For every $\gamma \in \mathbb{F}_2^n$, let $\beta_\gamma \in W$ be the unique representative with $\gamma \in \beta_\gamma + H^\perp$. Then:

$$\Pr_{H,\alpha}\left[\left|\widehat{f_A}(\beta_\gamma) - \widehat{f}(\gamma)\chi_\gamma(\alpha)\right| \geq 2^{-t} + \lambda\right] \leq \frac{2^{-t}}{\lambda^2}$$

Proof: Define $Z = \widehat{f_A}(\beta_\gamma) - \widehat{f}(\gamma)\chi_\gamma(\alpha) = \sum_{\substack{\beta \in \gamma + H^\perp \\ \beta \neq \gamma}} \widehat{f}(\beta)\chi_\beta(\alpha)$.

- **First moment:** $\mathbb{E}_{H,\alpha}[Z] \leq 2^{-t}$ (character orthogonality over α)
- **Second moment:** $\mathbb{E}_{H,\alpha}[Z^2] \leq 2^{-t}$ (Parseval + collision probability 2^{-t})
- **Variance:** $\mathrm{Var}(Z) \leq 2^{-t}$
- **Chebyshev** $\Rightarrow$ the lemma. ✓

Miracle 2 — Separation

Lemma 2 (Separation)

For any distinct $\alpha, \beta \in \mathbb{F}_2^n$:

$$\Pr_H[\alpha \text{ and } \beta \text{ lie in the same coset of } H^\perp] = 2^{-t}$$

Lemma 2 (Separation)

For any distinct $\alpha, \beta \in \mathbb{F}_2^n$:

$$\Pr_H[\alpha \text{ and } \beta \text{ lie in the same coset of } H^\perp] = 2^{-t}$$

Consequence for the heavy set:
Define $\mathcal{L} = \{\beta : |\widehat{f}(\beta)| \geq \tau^2/2\}$. By Parseval: $|\mathcal{L}| \leq 4/\tau^4$.

Let E_1 = event that two distinct elements of $\mathcal{L}$ land in the same coset of $H^\perp$.

$$\Pr(E_1) \leq \binom{|\mathcal{L}|}{2} \cdot 2^{-t} \leq \frac{8}{\tau^8} \cdot 2^{-t}.$$

For $t = \lceil 20 + 20 \log_2(1/\tau) \rceil$: $\quad \Pr(E_1) \leq \frac{1}{16}$.

With high probability, every heavy coefficient occupies its own bucket.

Miracle 3 — The Fourth Moment Argument

The danger: many tiny Fourier coefficients could aggregate into a single $\widehat{f_A}(\gamma)$ that *looks* heavy — creating a false positive.

Miracle 3 — The Fourth Moment Argument

The danger: many tiny Fourier coefficients could aggregate into a single $\widehat{f_A}(\gamma)$ that *looks* heavy — creating a false positive.

Tool: the *fourth Fourier moment*

$$\|\widehat{f}\|_4^4 = \sum_{\alpha} \widehat{f}(\alpha)^4 = \|f\|_{U^2}^4 \quad \text{(Gowers } U^2\text{-norm)}$$

measures spectral concentration. We show $\|\widehat{f_A}\|_4^4 \approx \|\widehat{f}\|_4^4$.

Key insight: $\max_\gamma |\widehat{f_A}(\gamma)| \leq \|\widehat{f_A}\|_4$, so small fourth moment $\Rightarrow$ no large coefficient.

Lemma 4 (Expectation)

$$\mathbb{E}_A\left[\|\widehat{f_A}\|_4^4\right] = (1 - 3\cdot 2^{-t} + 2\cdot 2^{-2t})\|\widehat{f}\|_4^4 + 3\cdot 2^{-t} - 2\cdot 2^{-2t}$$

So $\left|\mathbb{E}_A[\|\widehat{f_A}\|_4^4] - \|\widehat{f}\|_4^4\right| \le 3\cdot 2^{-t}$.

Lemma 5 (Variance)

$$\mathrm{Var}_A\left(\|\widehat{f_A}\|_4^4\right) \le 2^{1-t}$$

Lemma 6 (Concentration — via Chebyshev)

For any $\varepsilon > 0$: $\Pr_A\left[\left|\|\widehat{f_A}\|_4^4 - \|\widehat{f}\|_4^4\right| \ge \varepsilon\right] \le \dfrac{72\cdot 2^{-t}}{\varepsilon^2}$

Proof of Lemma 4 — Computing $\mathbb{E}_A[\|\widehat{f_A}\|_4^4]$

Expand, average over α: only zero-sum 4-tuples $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 0$ survive.

$$\mathbb{E}_\alpha\left[\|\widehat{f_A}\|_4^4 \,\middle|\, H\right] = \sum_{\substack{\beta_1,\ldots,\beta_4 \\ \beta_1+\cdots+\beta_4=0}} \prod_{i=1}^{4} \widehat{f}(\beta_i) \cdot \mathbf{1}_H(\beta_1,\ldots,\beta_4)$$

Classify zero-sum 4-tuples by $r = \dim \operatorname{span}\{\beta_i - \beta_1\}$. Prob. all in one coset $= 2^{-tr}$.

r	**Structure**	**Contribution**
0	all equal: $\beta_1 = \cdots = \beta_4$	$\|\widehat{f}\|_4^4$
1	two distinct pairs	$3(1 - \|\widehat{f}\|_4^4)$
2	general zero-sum	$2\|\widehat{f}\|_4^4 - 1$

Case $r = 2$ key step: $\sum_u \left|\widehat{f^2}(u)\right|^2 = \|f^2\|_2^2 = 1$

since $f^2 \equiv 1$ for Boolean f.

Combining all three cases:

$$\mathbb{E}_A\left[\|\widehat{f_A}\|_4^4\right] = (1 - 3a + 2a^2)\|\widehat{f}\|_4^4 + 3a - 2a^2$$

where $a = 2^{-t}$. So $|\mathbb{E}[\cdot] - \|\widehat{f}\|_4^4| \le 3a$. ✓

Completing the Proof — Putting It Together

Decompose: $\mathbb{F}_2^n = \mathcal{L} \cup \mathcal{S}$ where $\mathcal{L} = \{\beta : |\widehat{f}(\beta)| \geq \tau^2/2\}$ (heavy set).

On event $E_1^c \cap E_2^c \cap E_3^c$ (probability $\geq 3/4$):

1. **M2 (Separation):** heavy coefficients lie in distinct cosets of $H^\perp$
2. **M1 (Survival):** for each $\gamma \in \mathcal{L}$, $|\widehat{f_A}(\beta_\gamma)| \approx |\widehat{f}(\gamma)| \pm \tau^8/64$
3. **M3 (Stability):** total fourth moment of the light part:

$$\sum_{\gamma \notin W_{\mathcal{L}}} |\widehat{f_A}(\gamma)|^4 \;\leq\; \frac{\tau^4}{4} + \frac{\tau^4}{2} + \frac{\tau^4}{4} = \tau^4$$

Completing the Proof — The Final Bound

From M3:

$$\max_{\gamma \notin W_{\mathcal{L}}} |\widehat{f_A}(\gamma)| \;\le\; \left(\sum_{\gamma \notin W_{\mathcal{L}}} |\widehat{f_A}(\gamma)|^4 \right)^{1/4} \;\le\; \tau$$

No spurious heavy coefficients can emerge from the light tail.

From M1 + M2: the maximum of $|\widehat{f_A}(\gamma)|$ over heavy cosets equals $\mathrm{FBias}(f)$ up to $\pm\,\tau^8/64 \ll \tau$.

Combining:

$$\left| \max_{\gamma \in W} |\widehat{f_A}(\gamma)| - \mathrm{FBias}(f) \right| \le \tau$$

with probability $\ge 3/4$, using $\widetilde{O}(1/\tau^2)$ queries.

Theorem 1 Proved

Survival + Separation + Stability $\Rightarrow$ restriction preserves FBias.

Part 4 of 6

Fourier Sparsity Testing (ℓ_0)

Talk Roadmap

1. Setup
2. Fourier Sparsity Testing (ℓ_2)
3. Fourier Bias Estimation
4. Fourier Sparsity Testing (ℓ_0)
5. Lower Bounds
6. Open Problems

Why Hamming Distance Is Harder — The Key Distinction

Two distance measures:

$$\text{Hamming: } \delta(f,g) = \Pr_x[f(x) \neq g(x)] \qquad \ell_2\text{: } \operatorname{dist}_2(f,g)^2 = \mathbb{E}_x[(f(x)-g(x))^2]$$

Key Distinction

- ℓ_2 **tester:** certify that *enough large* Fourier coefficients exist (top-s mass is large)
- **Hamming tester:** must also certify the *absence of a Fourier tail* (no many small nonzero coefficients)

Hamming is **strictly harder** — certifying absence is harder than certifying presence.

Why Hamming Distance Is Harder — The Gap Illustrated

Two functions with the same large coefficients but different tails:

Function A — sparse, accepted by both tests:

Function B — ℓ_2 accepts, Hamming rejects:

Main Result — ICLR 2026

Theorem 1.1 (Upper Bound)

There is a **non-adaptive** tester with query complexity $\widetilde{O}(s^4)$ that distinguishes s-Fourier sparse from ε-far in Hamming distance.

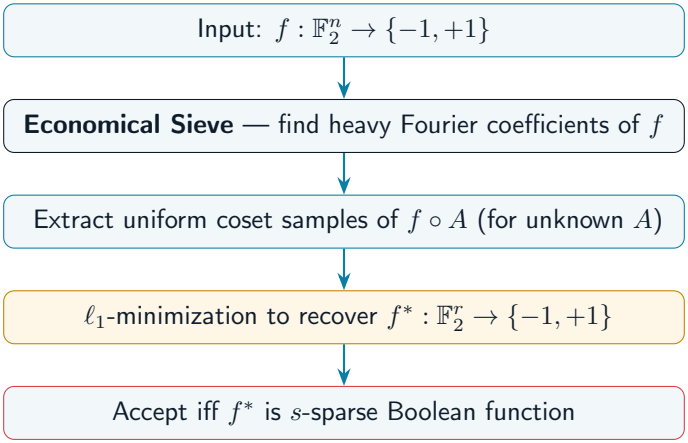
Theorem 1.2 (Lower Bound)

Any tester (adaptive or not) requires $\Omega(s)$ queries.

	Upper bound	**Lower bound**
Gopalan et al. 2011	$\widetilde{O}(s^{14})$	$\Omega(\sqrt{s})$
This work	$\widetilde{O}(s^4) \leftarrow$ exponent $14 \to 4$	$\Omega(s) \leftarrow \sqrt{s} \to s$

Algorithm: Testing via Implicit Sensing — Pipeline

High-level algorithm pipeline:



Algorithm: Testing via Implicit Sensing — Key Ingredient

Dimension Reduction (Sanyal 2019)

For any s-sparse $f : \mathbb{F}_2^n \to \{-1, +1\}$, the linear span of $\mathrm{supp}(\widehat{f})$ has dimension $r = O(\sqrt{s})$.

Therefore $f = f^* \circ L$ where $f^* : \mathbb{F}_2^r \to \{-1, +1\}$ and $L : \mathbb{F}_2^n \to \mathbb{F}_2^r$.

Why $r = O(\sqrt{s})$ is the key:

- The reduced function f^* lives on $\mathbb{F}_2^r$ with $r = O(\sqrt{s})$
- RIP (Haviv & Regev 2017): exact recovery of f^* from $O(s^{3/2} \operatorname{polylog} s)$ random samples — *independent of* n
- We generate these samples via coset sampling of $f \circ A$ for an *unknown* invertible A — without ever knowing A explicitly

This is "implicit" sensing: we sense without knowing the sensing matrix.

The Economical Sieve — What It Does

Lemma (Datta et al., STACS 2026)

Economical Sieve uses $\widetilde{O}\big(\max\{\frac{1}{\theta^4}, \frac{\lambda}{\theta^2}\}\big)$ queries to f and outputs simultaneously:

- All $\alpha \in \mathbb{F}_2^n$ with $|\widehat{f}(\alpha)| \geq \theta$ (heavy coefficients)
- λ uniformly random labeled samples $(x, f(x))$
- **Linear relations** among the heavy Fourier coefficients

The linear relations are the key: they let us discover the *unknown* linear transformation A implicitly, without knowing A explicitly.

We use these relations to extract **coset samples** of $f \circ A$ — uniform random samples from the projected function $f^* = f \circ T \circ A$ — without ever knowing A.

The Economical Sieve — Query Complexity

Plugging into the algorithm:

- Set threshold $\theta = \frac{1}{4s}$ (smallest possible nonzero coefficient of an s-sparse Boolean function)
- Set sample count $\lambda = \widetilde{O}(s^2)$ (sufficient for RIP-based recovery of f^*)

Total query complexity of the Economical Sieve:

$$\widetilde{O}\left(\max\left\{\frac{1}{\theta^4},\, \frac{\lambda}{\theta^2}\right\}\right) = \widetilde{O}\big(\max\{s^4,\, s^2 \cdot s^2\}\big) = \widetilde{O}(s^4)$$

Final Query Complexity

The full tester uses $\widetilde{O}(s^4)$ queries — matching Theorem 1.1.

Part 5 of 6

Lower Bounds

Talk Roadmap

1. Setup
2. Fourier Sparsity Testing (ℓ_2)
3. Fourier Bias Estimation
4. Fourier Sparsity Testing (ℓ_0)
5. Lower Bounds
6. Open Problems

Maiorana–McFarland Functions

Definition: for $r \leq n$, $n = r + \log r$, and $\phi : \mathbb{F}_2^{n-r} \to \mathbb{F}_2^r$ with r linearly independent outputs,

$$g_L(x, y) = (-1)^{\langle Lx, \phi(y) \rangle}, \quad (x, y) \in \mathbb{F}_2^r \times \mathbb{F}_2^{n-r}, \quad L \in \mathbb{F}_2^{r \times r}.$$

Spectral structure:

$$\widehat{g_L}(u, v) = \frac{1}{2^{n-r}} \sum_{y \in (L^T \phi)^{-1}(u)} (-1)^{v \cdot y}$$

$\Rightarrow |\mathrm{supp}(\widehat{g_L})| = \mathrm{rank}(L) \cdot r$, all nonzero coefficients equal in magnitude $= 1/\sqrt{\mathrm{rank}(L) \cdot r}$.

Why useful:

- $g_C = g_A \cdot g_B$ when $C = A + B$ — product factorizes Alice/Bob
- Sparsity/bias of g_C determined by $\mathrm{rank}(C)$ — links to comm. complexity
- Cryptographically hard: used in stream cipher design for confusion/diffusion

Approximate Matrix Rank

Communication Hardness (Sherstov & Storozhenko 2024)

Alice has $A \in \mathbb{F}_2^{r \times r}$, Bob has $B \in \mathbb{F}_2^{r \times r}$. Goal: is $\operatorname{rank}(A + B) = r$ or $r/4$?

$$R^*_{1/3}\big(\mathrm{RANK}_{r,r/4}\big) = \Omega(r^2)$$

Lower Bound for ℓ_2 Fourier Sparsity

Reduction from Approximate Matrix Rank

1. Alice and Bob construct g_A and g_B locally.
2. Their product satisfies

$$g_C = g_A \cdot g_B, \qquad C = A + B.$$

3. Each query to g_C costs **2 bits** of communication.
4. If $\mathrm{rank}(C) = r$, then g_C is $1/4$-far from every $r^2/4$-sparse function.
5. A q-query tester for sparsity $s = r^2/4$ yields a $2q$-bit communication protocol.

$$2q(s, \tfrac{1}{4}) \geq \Omega(r^2) = \Omega(s) \qquad \Longrightarrow \qquad q(s, \tfrac{1}{4}) = \Omega(s).$$

Any tester for ℓ_2 Fourier sparsity requires $\Omega(s)$ queries.

Lower Bound for Bias Estimation

Reduction from Approximate Matrix Rank

1. Alice and Bob construct g_A and g_B locally.
2. Their product satisfies

$$g_C = g_A \cdot g_B, \qquad C = A + B.$$

3. Each query to g_C costs **2 bits** of communication.
4. Setting $\tau = 1/(2r)$, a q_τ-query estimator yields a $2q_\tau$-bit communication protocol.
5. This gives a protocol for approximate matrix rank.

$$2q_{1/(2r)} \geq \Omega(r^2) \qquad \Longrightarrow \qquad q_\tau = \Omega\left(\frac{1}{\tau^2}\right).$$

Any estimator achieving additive error τ requires $\Omega(1/\tau^2)$ queries.

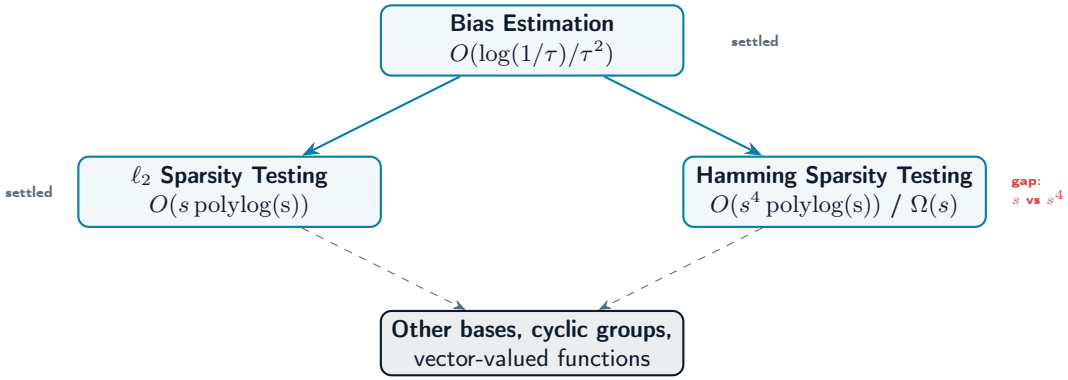
Part 6 of 6

Open Problems & Conclusion

Talk Roadmap

1. Setup
2. Fourier Sparsity Testing (ℓ_2)
3. Fourier Bias Estimation
4. Fourier Sparsity Testing (ℓ_0)
5. Lower Bounds
6. Open Problems

Research Landscape — What We Know



Open Problems

Open Problems

Problem 1 — Close the Hamming Gap (most urgent)

Best upper bound: $\widetilde{O}(s^4)$. Best lower bound: $\Omega(s)$.

Is the true complexity $\Theta(s)$, $\Theta(s^2)$, or $\Theta(s^4)$?

Open Problems

Problem 1 — Close the Hamming Gap (most urgent)

Best upper bound: $\widetilde{O}(s^4)$. Best lower bound: $\Omega(s)$.

Is the true complexity $\Theta(s)$, $\Theta(s^2)$, or $\Theta(s^4)$?

Problem 2 — Extend Lower Bounds to Arbitrary ε

All current lower bounds hold only for constant ε (e.g., $1/4$).

Extend to *arbitrary* $\varepsilon > 0$.

Open Problems

Open Problems

Problem 3 — Other Bases and Domains

Dimension-independent sparsity testing for wavelets, Fourier over $\mathbb{Z}_N$, or general dictionaries.

Open Problems

Problem 3 — Other Bases and Domains

Dimension-independent sparsity testing for wavelets, Fourier over $\mathbb{Z}_N$, or general dictionaries.

Problem 4 — Locate the Heavy Frequency

We can estimate $\mathrm{FBias}(f)$ in $\widetilde{O}(1/\tau^2)$ queries. Can we also *find* the α achieving it in $\widetilde{O}(n/\tau^2)$ queries, replacing the $O(n/\tau^3)$ cost of Goldreich–Levin?

Problem 3 — Other Bases and Domains

Dimension-independent sparsity testing for wavelets, Fourier over $\mathbb{Z}_N$, or general dictionaries.

Problem 4 — Locate the Heavy Frequency

We can estimate $\mathrm{FBias}(f)$ in $\widetilde{O}(1/\tau^2)$ queries. Can we also *find* the α achieving it in $\widetilde{O}(n/\tau^2)$ queries, replacing the $O(n/\tau^3)$ cost of Goldreich–Levin?

Thank You