Higher-Order Fourier Analysis: Applications to Algebraic Property Testing

Yuichi Yoshida

National Institute of Informatics, and
Preferred Infrastructure, Inc

October 18, 2014

Property testing

Definition

$f : \{0,1\}^n \to \{0,1\}$ is *ϵ-far* from $\mathcal{P}$ if,

$$d_{\mathcal{P}}(f) := \min_{g \in \mathcal{P}} \Pr_x[f(x) \neq g(x)] \geq \epsilon.$$

Property testing

Definition

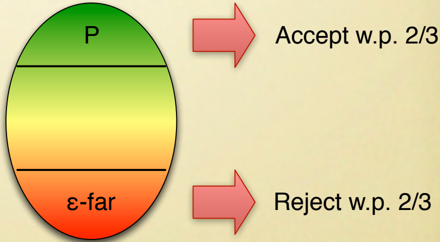
$f : \{0,1\}^n \to \{0,1\}$ is *ϵ-far* from $\mathcal{P}$ if,

$$d_{\mathcal{P}}(f) := \min_{g \in \mathcal{P}} \Pr_x[f(x) \neq g(x)] \geq \epsilon.$$

A *tester* for a property $\mathcal{P}$:

Given

- $f : \{0,1\}^n \to \{0,1\}$ as a query access.
- proximity parameter $\epsilon > 0$.



Linearity testing

Input: a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and $\epsilon > 0$.
Goal: $f(x) + f(y) = f(x + y)$ for every $x, y \in \mathbb{F}_2^n$?

Linearity testing

Input: a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and $\epsilon > 0$.
Goal: $f(x) + f(y) = f(x + y)$ for every $x, y \in \mathbb{F}_2^n$?

```
for i = 1 to O(1/ε) do
    Sample x, y ∈ F_2^n uniformly at random.
    if f(x) + f(y) ≠ f(x + y) then reject.
Accept.
```

Linearity testing

Input: a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and $\epsilon > 0$.
Goal: $f(x) + f(y) = f(x + y)$ for every $x, y \in \mathbb{F}_2^n$?

```
for i = 1 to O(1/ε) do
    Sample x, y ∈ F_2^n uniformly at random.
    if f(x) + f(y) ≠ f(x + y) then reject.
Accept.
```

Theorem ([BLR93])

- *If f is linear, always accepts. (one-sided error)*

Linearity testing

Input: a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and $\epsilon > 0$.
Goal: $f(x) + f(y) = f(x + y)$ for every $x, y \in \mathbb{F}_2^n$?

```
for i = 1 to O(1/ε) do
    Sample x, y ∈ F_2^n uniformly at random.
    if f(x) + f(y) ≠ f(x + y) then reject.
Accept.
```

Theorem ([BLR93])

- *If f is linear, always accepts. (one-sided error)*
- *If f is ϵ-far, rejects with probability at least $2/3$.*

Linearity testing

Input: a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and $\epsilon > 0$.
Goal: $f(x) + f(y) = f(x + y)$ for every $x, y \in \mathbb{F}_2^n$?

1: **for** $i = 1$ to $O(1/\epsilon)$ **do**
2: Sample $x, y \in \mathbb{F}_2^n$ uniformly at random.
3: **if** $f(x) + f(y) \neq f(x + y)$ **then** reject.
4: Accept.

Theorem ([BLR93])

- *If f is linear, always accepts. (one-sided error)*
- *If f is ϵ-far, rejects with probability at least $2/3$.*
- *Query complexity is $O(1/\epsilon) \Rightarrow$ constant!*

Backgrounds

The notion of property testing was introduced by [RS96].

Backgrounds

The notion of property testing was introduced by [RS96].

Since then, various kinds of objects have been studied.
Ex.: Functions, graphs, distributions, geometric objects, images.

Backgrounds

The notion of property testing was introduced by [RS96].

Since then, various kinds of objects have been studied.
Ex.: Functions, graphs, distributions, geometric objects, images.

Q. Why do we study property testing?

Backgrounds

The notion of property testing was introduced by [RS96].

Since then, various kinds of objects have been studied.
Ex.: Functions, graphs, distributions, geometric objects, images.

Q. Why do we study property testing?
A. Interested in

- ultra-efficient algorithms.
- relations to PCPs, locally testable codes, and learning.
- the relation between local view and global property.

Local testability of affine-Invariant properties

Definition

$\mathcal{P}$ is *affine-invariant* if a function $f : \mathbb{F}_2^n \to \{0,1\}$ satisfies $\mathcal{P}$, then $f \circ A$ satisfies $\mathcal{P}$ for any bijective affine transformation $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$.

Local testability of affine-Invariant properties

Definition

$\mathcal{P}$ is *affine-invariant* if a function $f : \mathbb{F}_2^n \to \{0, 1\}$ satisfies $\mathcal{P}$, then $f \circ A$ satisfies $\mathcal{P}$ for any bijective affine transformation $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$.

Definition

$\mathcal{P}$ is *(locally) testable* if there is a tester for $\mathcal{P}$ with $q(\epsilon)$ queries.

Ex.:

- degree-d polynomials [AKK$^+$05, BKS$^+$10]
- Fourier sparsity [GOS$^+$11]
- Odd-cycle-freeness: the Cayley graph has no odd cycle [BGRS12]

The goal

Q. Can we characterize testable affine-invariant properties? [KS08]

The goal

Q. *Can we characterize testable affine-invariant properties? [KS08]*

A. *Yes, in a satisfying sense.*

The goal

Q. Can we characterize testable affine-invariant properties? [KS08]

A. Yes, in a satisfying sense.

In this talk, we review how we have resolved this question.

- One-sided error testable $\approx$ Affine-subspace hereditary
- Testable $\Leftrightarrow$ Estimable
- Two-sided error testable $\Leftrightarrow$ Regular-reducible
- and more...

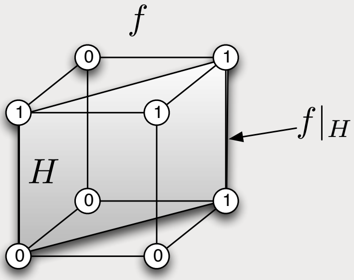
Higher order Fourier analysis has played a crucial role!

Oblivious tester

Definition

An *oblivious tester* works as follows:

- Take a restriction $f|_H$.
  - H: random affine subspace of dimension $h(\epsilon)$.
- Output based only on $f|_H$.



Motivation: avoid "unnatural" properties such as $f \in \mathcal{P} \Leftrightarrow n$ is even.
For natural properties, $\exists$ a tester $\Rightarrow \exists$ an oblivious tester.

Why is higher order Fourier analysis useful?

$\mu_{f,h}$: the distribution of $f|_H$.

Observation

A tester cannot distinguish f from g if $\mu_{f,h} \approx \mu_{g,h}$.

Why is higher order Fourier analysis useful?

$\mu_{f,h}$: the distribution of $f|_H$.

Observation

A tester cannot distinguish f from g if $\mu_{f,h} \approx \mu_{g,h}$.

Consider the *decomposition* $f = f_1 + f_2 + f_3$ for $d = d(\epsilon, h)$:

- $f_1 = \Gamma(P_1, \ldots, P_C)$ for high-rank degree-d polynomials $P_1, \ldots, P_C$.
- f_2: small L_2 norm.
- f_3: small U^{d+1} norm.

Why is higher order Fourier analysis useful?

$\mu_{f,h}$: the distribution of $f|_H$.

Observation

A tester cannot distinguish f from g if $\mu_{f,h} \approx \mu_{g,h}$.

Consider the *decomposition* $f = f_1 + f_2 + f_3$ for $d = d(\epsilon, h)$:

- $f_1 = \Gamma(P_1, \ldots, P_C)$ for high-rank degree-d polynomials $P_1, \ldots, P_C$.
- f_2: small L_2 norm.
- f_3: small U^{d+1} norm.

The *pseudorandom parts* f_2 and f_3 do not affect $\mu_{f,h}$ much.
$\Rightarrow$ we can focus on the *structured part* f_1.

One-sided error testable $\approx$ Affine-subspace hereditary

Affine-subspace hereditary

Definition

A property $\mathcal{P}$ is *affine-subspace hereditary* if
$f \in \mathcal{P} \Rightarrow f|_H \in \mathcal{P}$ for any affine subspace H.

Ex.:

- degree-d polynomials, Fourier sparsity, odd-cycle-freeness
- $f = gh$ for some polynomials g, h of degree $\leq d - 1$.
- $f = g^2$ for some polynomial g of degree $\leq d - 1$.

Characterization of one-sided error testability

Conjecture ([BGS10])

$\mathcal{P}$ is testable with one-sided error by an oblivious tester
$\Leftrightarrow$ $\mathcal{P}$ is (essentially) affine-subspace hereditary

Characterization of one-sided error testability

Conjecture ([BGS10])

$\mathcal{P}$ is testable with one-sided error by an oblivious tester
$\Leftrightarrow$ $\mathcal{P}$ is (essentially) affine-subspace hereditary

$\Rightarrow$ is true [BGS10].

Characterization of one-sided error testability

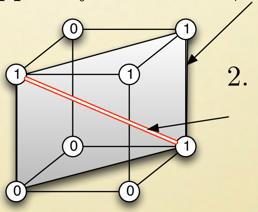
Conjecture ([BGS10])

$\mathcal{P}$ is testable with one-sided error by an oblivious tester
$\Leftrightarrow$ $\mathcal{P}$ is (essentially) affine-subspace hereditary

$\Rightarrow$ is true [BGS10].

Proof sketch:

1. Suppose $f \in \mathcal{P}$ and $f|_H \notin \mathcal{P}$



2. $\exists f|_K$, rejected by the tester

3. f is also rejected w.p.> 0, contradiction.

Alternative formulation via linear forms

Think of *affine-triangle-freeness*:
No $x, y_1, y_2 \in \mathbb{F}_2^n$ s.t. $f(x + y_1) = f(x + y_2) = f(x + y_1 + y_2) = 1$.

Alternative formulation via linear forms

Think of *affine-triangle-freeness*:
No $x, y_1, y_2 \in \mathbb{F}_2^n$ s.t. $f(x + y_1) = f(x + y_2) = f(x + y_1 + y_2) = 1$.

$\Leftrightarrow$ No $x, y_1, y_2 \in \mathbb{F}_2^n$ s.t.

$$f(L_1(x, y_1, y_2)) = \sigma_1 \text{ for } L_1(x, y_1, y_2) = x + y_1 \text{ and } \sigma_1 = 1,$$
$$f(L_2(x, y_1, y_2)) = \sigma_2 \text{ for } L_2(x, y_1, y_2) = x + y_2 \text{ and } \sigma_2 = 1,$$
$$f(L_3(x, y_1, y_2)) = \sigma_3 \text{ for } L_3(x, y_1, y_2) = x + y_1 + y_2 \text{ and } \sigma_3 = 1.$$

Alternative formulation via linear forms

Think of *affine-triangle-freeness*:
No $x, y_1, y_2 \in \mathbb{F}_2^n$ s.t. $f(x+y_1) = f(x+y_2) = f(x+y_1+y_2) = 1$.

$\Leftrightarrow$ No $x, y_1, y_2 \in \mathbb{F}_2^n$ s.t.

$f(L_1(x,y_1,y_2)) = \sigma_1$ for $L_1(x,y_1,y_2) = x+y_1$ and $\sigma_1 = 1$,
$f(L_2(x,y_1,y_2)) = \sigma_2$ for $L_2(x,y_1,y_2) = x+y_2$ and $\sigma_2 = 1$,
$f(L_3(x,y_1,y_2)) = \sigma_3$ for $L_3(x,y_1,y_2) = x+y_1+y_2$ and $\sigma_3 = 1$.

We call this *$(A = (L_1, L_2, L_3), \sigma = (\sigma_1, \sigma_2, \sigma_3))$-freeness*.

Alternative formulation via linear forms

Think of *affine-triangle-freeness*:
No $x, y_1, y_2 \in \mathbb{F}_2^n$ s.t. $f(x+y_1) = f(x+y_2) = f(x+y_1+y_2) = 1$.

$\Leftrightarrow$ No $x, y_1, y_2 \in \mathbb{F}_2^n$ s.t.

$$f(L_1(x,y_1,y_2)) = \sigma_1 \text{ for } L_1(x,y_1,y_2) = x+y_1 \text{ and } \sigma_1 = 1,$$
$$f(L_2(x,y_1,y_2)) = \sigma_2 \text{ for } L_2(x,y_1,y_2) = x+y_2 \text{ and } \sigma_2 = 1,$$
$$f(L_3(x,y_1,y_2)) = \sigma_3 \text{ for } L_3(x,y_1,y_2) = x+y_1+y_2 \text{ and } \sigma_3 = 1.$$

We call this $(A = (L_1, L_2, L_3), \sigma = (\sigma_1, \sigma_2, \sigma_3))$*-freeness*.

- A is called an *affine* system of *linear forms*.
  $\Rightarrow$ well studied in higher order Fourier analysis.

Testability of subspace hereditary properties

Observation

The following are equivalent:

- $\mathcal{P}$ *is affine-subspace hereditary.*
- *There exists a (possibly infinite) collection* $\{(A^1, \sigma^1), \ldots\}$ *s.t.* $f \in \mathcal{P} \Leftrightarrow f$ *is* (A^i, σ^i)*-free for each* i.

Testability of subspace hereditary properties

Observation

The following are equivalent:

- $\mathcal{P}$ *is affine-subspace hereditary.*
- *There exists a (possibly infinite) collection* $\{(A^1, \sigma^1), \ldots\}$ *s.t.* $f \in \mathcal{P} \Leftrightarrow f$ *is* (A^i, σ^i)*-free for each* i*.*

Theorem ([BFH+13])

If each (A^i, σ^i) *has bounded complexity, then the property is testable with one-sided error.*

Proof idea

Let's focus on the case $f = \Gamma(P_1, \ldots, P_C)$ and $\mathcal{P}$ = affine $\triangle$-freeness.

Proof idea

Let's focus on the case $f = \Gamma(P_1, \ldots, P_C)$ and $\mathcal{P}$ = affine $\triangle$-freeness.

f is ϵ-far from $\mathcal{P}$
$\Rightarrow$ There are $x^*, y_1^*, y_2^* \in \mathbb{F}_2^n$ spanning an affine triangle.

Proof idea

Let's focus on the case $f = \Gamma(P_1, \ldots, P_C)$ and $\mathcal{P}$ = affine $\triangle$-freeness.

f is ϵ-far from $\mathcal{P}$
$\Rightarrow$ There are $x^*, y_1^*, y_2^* \in \mathbb{F}_2^n$ spanning an affine triangle.

$$\Pr_{x,y_1,y_2}[f(x+y_1) = f(x+y_2) = f(x+y_1+y_2) = 1]$$
$$\geq \Pr_{x,y_1,y_2}[P_i(L_j(x,y_1,y_2)) = P_i(L_j(x^*,y_1^*,y_2^*))\ \forall i \in [C], j \in [3]],$$

which is non-negligibly high from the *equidistribution theorem*.
$\Rightarrow$ Random sampling works.

Testability $\Leftrightarrow$ Estimability

Testability $\Leftarrow$ Estimability

Definition

$\mathcal{P}$ is *estimable* if we can estimate $d_{\mathcal{P}}(\cdot)$ to within δ with $q(\delta)$ queries for any $\delta > 0$.

Testability $\Leftarrow$ Estimability

Definition

$\mathcal{P}$ is *estimable* if we can estimate $d_{\mathcal{P}}(\cdot)$ to within δ with $q(\delta)$ queries for any $\delta > 0$.

Trivial direction: $\mathcal{P}$ is estimable $\Rightarrow$ $\mathcal{P}$ is testable.

Testability $\Leftarrow$ Estimability

Definition

$\mathcal{P}$ is *estimable* if we can estimate $d_{\mathcal{P}}(\cdot)$ to within δ with $q(\delta)$ queries for any $\delta > 0$.

Trivial direction: $\mathcal{P}$ is estimable $\Rightarrow$ $\mathcal{P}$ is testable.

Theorem ([HL13])

$\mathcal{P}$ is testable $\Rightarrow$ $\mathcal{P}$ is estimable.

Testability $\Leftarrow$ Estimability

Definition

$\mathcal{P}$ is *estimable* if we can estimate $d_{\mathcal{P}}(\cdot)$ to within δ with $q(\delta)$ queries for any $\delta > 0$.

Trivial direction: $\mathcal{P}$ is estimable $\Rightarrow$ $\mathcal{P}$ is testable.

Theorem ([HL13])

$\mathcal{P}$ is testable $\Rightarrow$ $\mathcal{P}$ is estimable.

Algorithm:

1: $H \leftarrow$ a random affine subspace of a constant dimension.
2: **return** Output $d_{\mathcal{P}}(f|_H)$.

Intuition behind the proof

Why can we expect $d_{\mathcal{P}}(f) \approx d_{\mathcal{P}}(f|_H)$*?*

Intuition behind the proof

Why can we expect $d_{\mathcal{P}}(f) \approx d_{\mathcal{P}}(f|_H)$?

(Oversimplified argument)

- Since $\mathcal{P}$ is testable, $d_{\mathcal{P}}(f)$ is determined by the distribution $\mu_{f,h}$.

Intuition behind the proof

Why can we expect $d_{\mathcal{P}}(f) \approx d_{\mathcal{P}}(f|_H)$?

(Oversimplified argument)

- Since $\mathcal{P}$ is testable, $d_{\mathcal{P}}(f)$ is determined by the distribution $\mu_{f,h}$.
- If $f = \Gamma(P_1, \ldots, P_C)$, then $\mu_{f,h}$ is determined by Γ, degrees and depths of $P_1, \ldots, P_C$ (rather than *P_i's themselves*).

Intuition behind the proof

Why can we expect $d_{\mathcal{P}}(f) \approx d_{\mathcal{P}}(f|_H)$?

(Oversimplified argument)

- Since $\mathcal{P}$ is testable, $d_{\mathcal{P}}(f)$ is determined by the distribution $\mu_{f,h}$.
- If $f = \Gamma(P_1, \ldots, P_C)$, then $\mu_{f,h}$ is determined by Γ, degrees and depths of $P_1, \ldots, P_C$ (rather than *P_i's themselves*).
- $f = \Gamma(P_1, \ldots, P_C)$ and $f_H = \Gamma(P_1|_H, \ldots, P_C|_H)$ share the same Γ, degrees and depths.

  $\Rightarrow \mu_{f,h} \approx \mu_{f|_H,h}$.

  $\Rightarrow d_{\mathcal{P}}(f) \approx d_{\mathcal{P}}(f|_H)$.

Two-sided error testability $\Leftrightarrow$ Regular-reducibility

Structured part

Recall that, for $f = \Gamma(P_1, \ldots, P_C) + f_2 + f_3$,

> *$\mu_{f,h}$ is determined by Γ, and degrees and depths of P_i's.*

Let's use them as a (constant-size) sketch of f.

Regularity-instance

Definition

A *regularity-instance* I is a tuple of

- an error parameter $\gamma > 0$,
- a structure function $\Gamma : \prod_{i=1}^{C} \mathbb{U}_{h_i+1} \to [0, 1]$,
- a complexity parameter $C \in \mathbb{N}$,
- a degree-bound parameter $d \in \mathbb{N}$,
- a degree parameter $\mathbf{d} = (d_1, \ldots, d_C) \in \mathbb{N}^C$ with $d_i < d$,
- a depth parameter $\mathbf{h} = (h_1, \ldots, h_C) \in \mathbb{N}^C$ with $h_i < \frac{d_i}{p-1}$, and
- a rank parameter $r \in \mathbb{N}$.

Satisfying a regularity-instance

Definition

Let $I = (\gamma, \Gamma, C, d, \mathbf{d}, \mathbf{h}, r)$ be a regularity-instance.
f *satisfies I* if it is of the form

$$f(x) = \Gamma(P_1(x), \ldots, P_C(x)) + \Upsilon(x),$$

where

- P_i is a polynomial of degree d_i and depth h_i,
- $(P_1, \ldots, P_C)$ has rank at least r,
- $\|\Upsilon\|_{U^d} \leq \gamma$.

Testing regularity-instances

Theorem ([Yos14a])

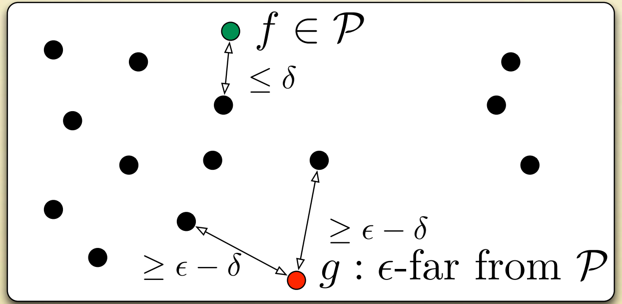
For any high-rank regularity-instance I, there is a tester for the property of satisfying I.

Algorithm:

1: $H \leftarrow$ a random affine subspace of a constant dimension.
2: **if** $f|_H$ is close to satisfying I **then** accept.
3: **else** reject.

Regular-reducibility

A property $\mathcal{P}$ is *regular-reducible* if for any $\delta > 0$, there exists a set $\mathcal{R}$ of constant number of high-rank regularity-instances such that:



Characterization of two-sided error testability

Theorem

An affine-invariant property $\mathcal{P}$ is testable

$$\Updownarrow$$

$\mathcal{P}$ is regular-reducible.

Characterization of two-sided error testability

Theorem

An affine-invariant property $\mathcal{P}$ is testable

$$\Updownarrow$$

$\mathcal{P}$ is regular-reducible.

Proof sketch:

- Regular-reducible $\Rightarrow$ testable
  Regularity-instances are testable, and testability implies estimability [HL13]. Hence, we can estimate the distance to $\mathcal{R}$.
- Testable $\Rightarrow$ regular-reducible
  The behavior of a tester depends only on $\mu_{f,h}$. Since Γ, $\mathbf{d}$, and $\mathbf{h}$ determines the distribution, we can find $\mathcal{R}$ using the tester.

Another characterization

$f, g : \mathbb{F}_2^n \to \{0,1\}$ are indistinguishable if $\mu_{f,h} \approx \mu_{g,h}$
$\Leftrightarrow \upsilon^d(f,g) := \min_A \|f - g \circ A\|_{U^d}$ is small.

Another characterization

$f, g : \mathbb{F}_2^n \to \{0,1\}$ are indistinguishable if $\mu_{f,h} \approx \mu_{g,h}$
$\Leftrightarrow \upsilon^d(f,g) := \min_A \|f - g \circ A\|_{U^d}$ is small.

Q. Can we generalize υ^d to functions over different domains?
A. Yes, with the aid of *non-standard analysis*.

Another characterization

$f, g : \mathbb{F}_2^n \to \{0,1\}$ are indistinguishable if $\mu_{f,h} \approx \mu_{g,h}$
$\Leftrightarrow \upsilon^d(f,g) := \min_A \|f - g \circ A\|_{U^d}$ is small.

Q. Can we generalize υ^d to functions over different domains?
A. Yes, with the aid of *non-standard analysis*.

We can define a counterpart of graphons [LS06] and a metric on it.

Another characterization

$f, g : \mathbb{F}_2^n \to \{0, 1\}$ are indistinguishable if $\mu_{f,h} \approx \mu_{g,h}$
$\Leftrightarrow \upsilon^d(f, g) := \min_A \|f - g \circ A\|_{U^d}$ is small.

Q. Can we generalize υ^d to functions over different domains?
A. Yes, with the aid of *non-standard analysis*.

We can define a counterpart of graphons [LS06] and a metric on it.

Theorem ([Yos14b])

A property $\mathcal{P}$ is testable
$\Leftrightarrow$ for any sequence $(f_i : \mathbb{F}_2^{n_i} \to \{0, 1\})$ that converges in the υ^d-metric for any $d \in \mathbb{N}$, the sequence $d_{\mathcal{P}}(f_i)$ converges.

Summary

Higher order Fourier analysis is useful for studying property testing as

- we care about the distribution $\mu_{f,h}$ for $h = O(1)$,
- which is determined by the structured part given by the decomposition theorem.

Summary

Higher order Fourier analysis is useful for studying property testing as

- we care about the distribution $\mu_{f,h}$ for $h = O(1)$,
- which is determined by the structured part given by the decomposition theorem.

We are almost done, *qualitatively*.

- one-sided error testability $\approx$ affine-subspace hereditary (of bounded complexity)
- two-sided error testability $\Leftrightarrow$ regular-reducibility.

Summary

Higher order Fourier analysis is useful for studying property testing as

- we care about the distribution $\mu_{f,h}$ for $h = O(1)$,
- which is determined by the structured part given by the decomposition theorem.

We are almost done, *qualitatively*.

- one-sided error testability $\approx$ affine-subspace hereditary (of bounded complexity)
- two-sided error testability $\Leftrightarrow$ regular-reducibility.

Thanks!