Testing Linear-Invariant Non-Linear Properties

Arnab Bhattacharyya, Victor Chen, Madhu Sudan, Ning Xie
CSAIL, Massachusetts Institute of Technology

IIT Kanpur Theory Seminar
January 27, 2009

Testing Properties of Functions

$f : D \to R$ function from huge finite domain D to (small?) finite range R.

A function property is a family of functions $\mathcal{F} \subseteq \{D \to R\}$.

1-sided Property Tester

A q-local tester for function property $\mathcal{F}$ is an algorithm that, when given a parameter $\epsilon \in (0, 1)$ and oracle access to the function f,

- makes at most q oracle queries to f
- accepts with probability 1 if $f \in \mathcal{F}$
- rejects with high probability if $\min_{g \in \mathcal{F}} \Pr_{x \in D}[f(x) \neq g(x)] > \epsilon$ (a.k.a. f is ϵ-far from $\mathcal{F}$)

$\mathcal{F}$ is called q-locally testable if there is a local tester with query complexity q independent of $|D|$.

Testing Properties of Functions

$f : D \to R$ function from huge finite domain D to (small?) finite range R.

A function property is a family of functions $\mathcal{F} \subseteq \{D \to R\}$.

1-sided Property Tester

A q-local tester for function property $\mathcal{F}$ is an algorithm that, when given a parameter $\epsilon \in (0, 1)$ and oracle access to the function f,

- makes at most q oracle queries to f
- accepts with probability 1 if $f \in \mathcal{F}$
- rejects with high probability if $\min_{g \in \mathcal{F}} \Pr_{x \in D}[f(x) \neq g(x)] > \epsilon$ (a.k.a. f is ϵ-far from $\mathcal{F}$)

$\mathcal{F}$ is called q-locally testable if there is a local tester with query complexity q independent of $|D|$.

Locally Testable Properties: Abbreviated History

- Prehistoric: Statistical Sampling
  - Is input function identically 0?
- Linearity Testing [BLR '90], Multilinearity Testing [BFL '91]
- Graph/Combinatorial Property Testing [GGR '98]
  - Is a graph "close" to being 3-colorable?
- Algebraic Testing [GLRSW, RS, FS, AKKLR, KR, JPRZ]
  - Is multivariate function a polynomial of bounded degree?

Locally Testable Properties: Abbreviated History

- Prehistoric: Statistical Sampling
  - Is input function identically 0?
- Linearity Testing [BLR '90], Multilinearity Testing [BFL '91]
- Graph/Combinatorial Property Testing [GGR '98]
  - Is a graph "close" to being 3-colorable?
- Algebraic Testing [GLRSW, RS, FS, AKKLR, KR, JPRZ]
  - Is multivariate function a polynomial of bounded degree?

Our Quest: **What makes a property testable?**

Necessary Conditions for Testability

- One-sided error and testability:
  - Suppose f is rejected by a k-query tester. Suppose queried points are $x_1, \dots, x_k \in D$ and $f(x_i) = \alpha_i$.
  - For every function $g \in \mathcal{F}$, $\langle g(x_1), \dots, g(x_k) \rangle \neq \langle \alpha_1, \dots, \alpha_k \rangle$.

Necessary Conditions for Testability

- One-sided error and testability:
  - Suppose f is rejected by a k-query tester. Suppose queried points are $x_1, \ldots, x_k \in D$ and $f(x_i) = \alpha_i$.
  - For every function $g \in \mathcal{F}$, $\langle g(x_1), \ldots, g(x_k) \rangle \neq \langle \alpha_1, \ldots, \alpha_k \rangle$.

- k-Local Constraint: $C = \langle x_1, \ldots, x_k \rangle; S \subsetneq R^k$. Function g satisfies C if $\langle g(x_1), \ldots, g(x_k) \rangle \in S$.
  - k-Local Testability implies k-Local Constraints.

Necessary Conditions for Testability

- One-sided error and testability:
  - Suppose f is rejected by a k-query tester. Suppose queried points are $x_1, \ldots, x_k \in D$ and $f(x_i) = \alpha_i$.
  - For every function $g \in \mathcal{F}$, $\langle g(x_1), \ldots, g(x_k) \rangle \neq \langle \alpha_1, \ldots, \alpha_k \rangle$.

- k-Local Constraint: $C = \langle x_1, \ldots, x_k \rangle; S \subsetneq R^k$. Function g satisfies C if $\langle g(x_1), \ldots, g(x_k) \rangle \in S$.
  - k-Local Testability implies k-Local Constraints.

- Every $f \notin \mathcal{F}$ rejected by some k-local constraint. Set of k-local constraints characterize $\mathcal{F}$:

$$\exists C_1, \ldots, C_m \text{ s.t. } f \in \mathcal{F} \Leftrightarrow f \text{ satisfies } C_j \ \forall j$$

  - Local Testability implies Local Characterization.

Local Characterizations Sufficient?

- NO! [Ben-Sasson, Harsha, Raskhodnikova '03]
  - $D = [n]$; $R = \{0, 1\}$; $\mathcal{F} =$ set of functions that satisfy some random 3-ary constraints
  - $\mathcal{F}$ not $o(n)$-locally testable
- Criticism: Random constraints too "asymmetric"
- Perhaps should consider more "symmetric" properties

Towards the Holy Grail

What are sufficient conditions for testability?

Towards the Holy Grail

What are sufficient conditions for testability?

The best understood class of nontrivial properties is graph properties

- Invariant under renaming of vertices
- [AFNS, Borgs et al] show that for graph properties in the "dense graph model", testability is *all about regularity*. A graph property is testable if and only if the property is "regular".

Towards the Holy Grail

What are sufficient conditions for testability?

The best understood class of nontrivial properties is graph properties

- Invariant under renaming of vertices
- [AFNS, Borgs et al] show that for graph properties in the "dense graph model", testability is *all about regularity*. A graph property is testable if and only if the property is "regular".

What about testability of algebraic properties?

Linear-Invariant Algebraic Properties

- Domain D is a vector space
  - In this talk, domain $D = \mathbb{F}_2^n$ and range $R = \{0, 1\}$
- Property $\mathcal{F}$ is invariant under linear transformations of the domain.
  - If $f \in \mathcal{F}$, then $f \circ L \in \mathcal{F}$ for any linear transformation $L \in \mathbb{F}_2^{n \times n}$

Example

- Linear functions, and more generally, polynomials in $\mathbb{F}_2[x_1, \ldots, x_n]$ of degree $\leqslant d$
- Homogeneous polynomials of a given degree
- Dual-BCH codes
- $\mathcal{F}_1 \cap \mathcal{F}_2$ and $\mathcal{F}_1 + \mathcal{F}_2$, where $\mathcal{F}_1, \mathcal{F}_2$ are linear-invariant

Linear-Invariant Algebraic Properties

- Domain D is a vector space
  - In this talk, domain $D = \mathbb{F}_2^n$ and range $R = \{0, 1\}$
- Property $\mathcal{F}$ is invariant under linear transformations of the domain.
  - If $f \in \mathcal{F}$, then $f \circ L \in \mathcal{F}$ for any linear transformation $L \in \mathbb{F}_2^{n \times n}$

Example

- Linear functions, and more generally, polynomials in $\mathbb{F}_2[x_1, \ldots, x_n]$ of degree $\leqslant d$
- Homogeneous polynomials of a given degree
- Dual-BCH codes
- $\mathcal{F}_1 \cap \mathcal{F}_2$ and $\mathcal{F}_1 + \mathcal{F}_2$, where $\mathcal{F}_1, \mathcal{F}_2$ are linear-invariant

All the above properties known to be locally testable. Does testability arise from linear-invariance?

Linear-Invariant + Linear Properties

- Suppose that the range R has algebraic structure as a field $\mathbb{F}$.
  - Property $\mathcal{F}$ is *linear* if $f, g \in \mathcal{F}$ and $\alpha, \beta \in \mathbb{F}$ implies $\alpha f + \beta g \in \mathcal{F}$.
- All the properties on previous slide were linear-invariant and linear.

Linear-Invariant + Linear Properties

- Suppose that the range R has algebraic structure as a field $\mathbb{F}$.
  - Property $\mathcal{F}$ is *linear* if $f, g \in \mathcal{F}$ and $\alpha, \beta \in \mathbb{F}$ implies $\alpha f + \beta g \in \mathcal{F}$.
- All the properties on previous slide were linear-invariant and linear.

Theorem ([Kaufman, Sudan '07])

If $\mathcal{F}$ is a linear-invariant, linear property that is locally characterized, then $\mathcal{F}$ is locally testable.

Theorem explains the testability of diverse properties proved previously using diverse techniques.

Linear-Invariant Non-Linear Properties: An Example

- The property of triangle-freeness
  - A function $f : \mathbb{F}_2^n \rightarrow \{0, 1\}$ is triangle-free if for every $x, y \in \mathbb{F}_2^n$, $\langle f(x), f(y), f(x + y)\rangle \neq \langle 1, 1, 1\rangle$
- Linear-invariant, yet non-linear

Linear-Invariant Non-Linear Properties: An Example

- The property of triangle-freeness
  - A function $f : \mathbb{F}_2^n \to \{0, 1\}$ is triangle-free if for every $x, y \in \mathbb{F}_2^n$, $\langle f(x), f(y), f(x + y)\rangle \neq \langle 1, 1, 1\rangle$
- Linear-invariant, yet non-linear
- Green in [Green '05] showed that triangle-freeness is testable!
- Analysis is very different from that of typical algebraic tests and is more reminiscent of graph property testing.
  - An algebraic Regularity Lemma

Matroid Freeness Properties

A more systematically defined class of linear-invariant, non-linear properties:

Definition (Matroid Freeness)

Given a binary matroid $\mathcal{M}$ represented by $v_1, \ldots, v_k \in \mathbb{F}_2^k$, the property of being $\mathcal{M}$-*free* is given by the family:

$$\{f : \mathbb{F}_2^n \to \{0,1\} \,|\, \forall \text{ linear } L : \mathbb{F}_2^k \to \mathbb{F}_2^n, \langle f(L(v_1)), \ldots, f(L(v_k))\rangle \neq 1^k\}$$

Example: Triangle freeness is described by the matroid $\mathcal{M}_\Delta = \{e_1, e_2, e_1 + e_2\}$.

- Under different linear transforms, $e_1, e_2, e_1 + e_2$ get mapped to all the triples of the form $x, y, x + y$.

Matroid Freeness Properties

A more systematically defined class of linear-invariant, non-linear properties:

Definition (Matroid Freeness)

Given a binary matroid $\mathcal{M}$ represented by $v_1, \ldots, v_k \in \mathbb{F}_2^k$, the property of being $\mathcal{M}$-*free* is given by the family:

$$\{f : \mathbb{F}_2^n \to \{0,1\} \mid \forall \text{ linear } L : \mathbb{F}_2^k \to \mathbb{F}_2^n, \langle f(L(v_1)), \ldots, f(L(v_k))\rangle \neq 1^k\}$$

Example: Triangle freeness is described by the matroid $\mathcal{M}_\Delta = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2\}$.

- Under different linear transforms, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2$ get mapped to all the triples of the form $x, y, x + y$.

For our purposes, a matroid is a collection of vectors in $\mathbb{F}_2^k$.

Matroid Freeness Properties

A more systematically defined class of linear-invariant, non-linear properties:

Definition (Matroid Freeness)

Given a binary matroid $\mathcal{M}$ represented by $v_1, \ldots, v_k \in \mathbb{F}_2^k$, the property of being $\mathcal{M}$-*free* is given by the family:

$$\{f : \mathbb{F}_2^n \to \{0,1\} \mid \forall \text{ linear } L : \mathbb{F}_2^k \to \mathbb{F}_2^n, \langle f(L(v_1)), \ldots, f(L(v_k))\rangle \neq 1^k\}$$

Example: Triangle freeness is described by the matroid $\mathcal{M}_\Delta = \{e_1, e_2, e_1 + e_2\}$.

- Under different linear transforms, $e_1, e_2, e_1 + e_2$ get mapped to all the triples of the form $x, y, x + y$.

Matroid Freeness Properties

A more systematically defined class of linear-invariant, non-linear properties:

Definition (Matroid Freeness)

Given a binary matroid $\mathcal{M}$ represented by $v_1, \ldots, v_k \in \mathbb{F}_2^k$, the property of being $\mathcal{M}$-*free* is given by the family:

$$\{f : \mathbb{F}_2^n \to \{0,1\} \mid \forall \text{ linear } L : \mathbb{F}_2^k \to \mathbb{F}_2^n, \langle f(L(v_1)), \ldots, f(L(v_k))\rangle \neq 1^k\}$$

Example: Triangle freeness is described by the matroid $\mathcal{M}_\Delta = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2\}$.

- Under different linear transforms, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2$ get mapped to all the triples of the form $x, y, x + y$.

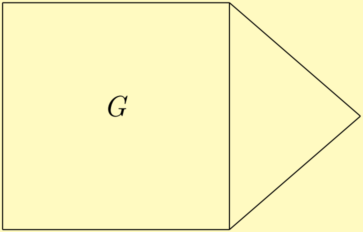
We say f *contains* $\mathcal{M}$ at L if $\langle f(L(v_1)), \ldots, f(L(v_k))\rangle = 1^k$

Graphic Matroids

A matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ forms a graphic matroid if there exists a simple graph G on k edges with each edge associated with an element of $\mathcal{M}$ such that a set $S \subset \{v_1, \ldots, v_k\}$ has a linear dependency iff the associated set of edges contains a cycle in G.

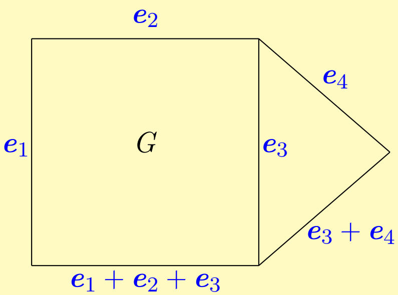
Graphic Matroids

A matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ forms a graphic matroid if there exists a simple graph G on k edges with each edge associated with an element of $\mathcal{M}$ such that a set $S \subset \{v_1, \ldots, v_k\}$ has a linear dependency iff the associated set of edges contains a cycle in G.



Graphic Matroids

A matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ forms a graphic matroid if there exists a simple graph G on k edges with each edge associated with an element of $\mathcal{M}$ such that a set $S \subset \{v_1, \ldots, v_k\}$ has a linear dependency iff the associated set of edges contains a cycle in G.



Testability of G-freeness

For any matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$, there is a natural k-local $\mathcal{M}$-freeness tester: pick a linear map $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ uniformly at random and check that f does not contain $\mathcal{M}$ at L.

[Green '05] showed that C_k-freeness is testable with natural test.

Testability of G-freeness

For any matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$, there is a natural k-local $\mathcal{M}$-freeness tester: pick a linear map $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ uniformly at random and check that f does not contain $\mathcal{M}$ at L.

Theorem (Main Result)

Given any graphic matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$, there is a function $\tau : \mathbb{R}^+ \to \mathbb{R}^+$ such that the natural k-local tester for $\mathcal{M}$-freeness accepts $\mathcal{M}$-free functions with probability 1 and rejects functions ϵ-far from being $\mathcal{M}$-free with probability $\tau(\epsilon)$.

Testability of G-freeness

For any matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$, there is a natural k-local $\mathcal{M}$-freeness tester: pick a linear map $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ uniformly at random and check that f does not contain $\mathcal{M}$ at L.

Theorem (Main Result)

Given any graphic matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$, there is a function $\tau : \mathbb{R}^+ \to \mathbb{R}^+$ such that the natural k-local tester for $\mathcal{M}$-freeness accepts $\mathcal{M}$-free functions with probability 1 and rejects functions ϵ-far from being $\mathcal{M}$-free with probability $\tau(\epsilon)$.

$\tau(\epsilon) = 2^{-O(W(\mathrm{poly}(1/\epsilon)))}$ where $W(t)$ is a tower of two's of height $\lceil t \rceil$.

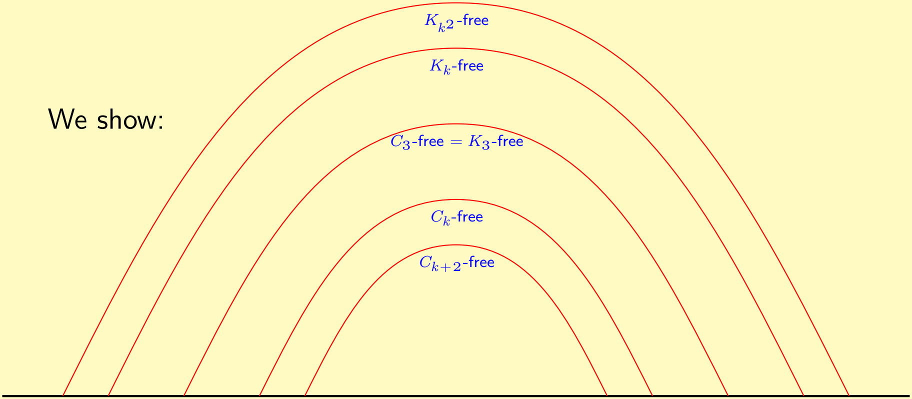
The Family of G-free Properties

When are G_1-freeness and G_2-freeness distinct properties?

The Family of G-free Properties

When are G_1-freeness and G_2-freeness distinct properties?

We show:



Summary of Contributions

- We uniformly describe a new infinite family of testable functions.

Summary of Contributions

- We uniformly describe a new infinite family of testable functions.
- In subsequent work, [Shapira '08] and [Král, Serra, Vena '08] showed testability for non-graphic matroid patterns using regularity lemma for hypergraphs.

Summary of Contributions

- We uniformly describe a new infinite family of testable functions.
- In subsequent work, [Shapira '08] and [Král, Serra, Vena '08] showed testability for non-graphic matroid patterns using regularity lemma for hypergraphs.
- Future Work
  - Testability of matroid-freeness where the forbidden pattern is not 1^k but an arbitrary binary string
  - Allow for conjunctions of constraints
  - Show separation between more classes of matroid-freeness properties
  - Lower bounds for query complexity of matroid-freeness properties

Outline of Proof for Testability

Based on notion of "complexity" of matroids, introduced by [Green, Tao '06]. Complexity is an integer and the simplest matroids have "complexity one". We prove:

Lemma (Complexity of Graphic Matroids)

For all graphs G, the graphic matroid of G has complexity 1.

Lemma (Testability of 1-complex Matroid Freeness)

If $\mathcal{M} = \{v_1, \ldots, v_k\}$ is a matroid of complexity 1, then the natural k-local tester for $\mathcal{M}$-freeness accepts $\mathcal{M}$-free functions with probability 1 and rejects functions ϵ-far from $\mathcal{M}$-free with probability at least $\tau(\epsilon)$.

The Complexity of Matroids

Definition (Matroid complexity [Green, Tao '06])

Given a binary matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$, $\mathcal{M}$ has complexity c if for every $i \in [k]$, the set $\{v_j\}_{j \in [k] \setminus \{i\}}$ can be partitioned into $c + 1$ classes such that v_i is not in the linear span of any of the classes.

The Complexity of Matroids

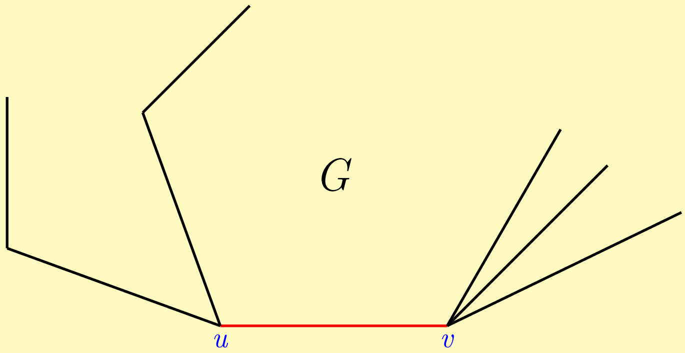
Definition (Matroid complexity [Green, Tao '06])

Given a binary matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$, $\mathcal{M}$ has complexity c if for every $i \in [k]$, the set $\{v_j\}_{j \in [k] \setminus \{i\}}$ can be partitioned into $c+1$ classes such that v_i is not in the linear span of any of the classes.

Motivating Example:

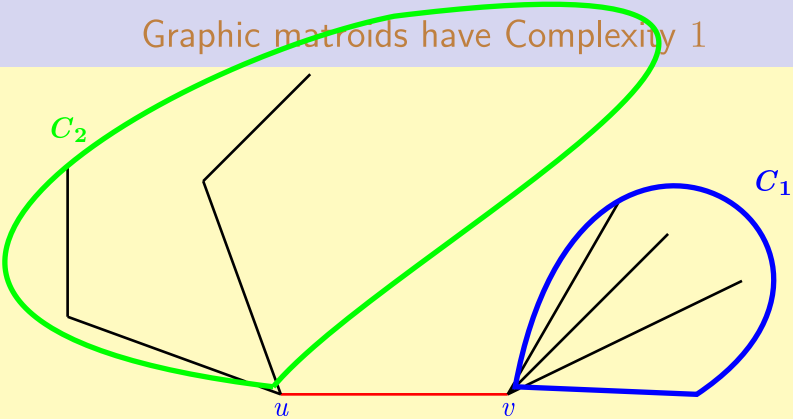
- Let $\mathcal{M} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_{k-1}, \boldsymbol{e}_1 + \cdots + \boldsymbol{e}_{k-1}\}$ be the graphic matroid of C_k. Then, $\mathcal{M}$ has complexity 1.

Graphic matroids have Complexity 1



For each edge (u, v), we need to partition rest of the edges into two classes such that neither contains (u, v) as part of a cycle.

Graphic matroids have Complexity 1



If $C_1 = \{\text{edges incident to } v\}$ and $C_2 = E(G) \setminus (C_1 \cup \{(u, v)\})$, edge (u, v) is not in the span of either of these sets because G is simple.

Hence, graphic matroids have complexity 1.

Key Property of Complexity 1 Matroids

[Green, Tao '06] showed that if a matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ has complexity 1, and if $A \subseteq \mathbb{F}_2^n$, then the number of linear maps $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ such that $L(v_i) \in A$ for all $i \in [k]$ is controlled by the Fourier spectrum of A.

Key Property of Complexity 1 Matroids

[Green, Tao '06] showed that if a matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ has complexity 1, and if $A \subseteq \mathbb{F}_2^n$, then the number of linear maps $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ such that $L(v_i) \in A$ for all $i \in [k]$ is controlled by the Fourier spectrum of A.

Overview of Boolean Harmonic Analysis

For $f : \mathbb{F}_2^n \to \mathbb{R}$, define $\widehat{f} : \mathbb{F}_2^n \to \mathbb{R}$ as

$$\widehat{f}(\alpha) = \mathop{\mathbb{E}}_{x \in \{0,1\}^n} \left[f(x)(-1)^{\sum \alpha_i x_i} \right]$$

Properties:

- Fourier inversion: $f(x) = \sum_\alpha \widehat{f}(\alpha)(-1)^{\sum \alpha_i x_i}$
- Parseval: $\sum_\alpha \widehat{f}^2(\alpha) = \mathbb{E}_x[f^2(x)]$
- Zero'th coefficient: $\widehat{f}(0) = \mathbb{E}_x[f(x)]$

Key Property of Complexity 1 Matroids

[Green, Tao '06] showed that if a matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ has complexity 1, and if $A \subseteq \mathbb{F}_2^n$, then the number of linear maps $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ such that $L(v_i) \in A$ for all $i \in [k]$ is controlled by the Fourier spectrum of A.

Lemma (Generalized von Neumann)

Suppose the matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ *has complexity* 1 *and let* $f_1, \ldots, f_k : \mathbb{F}_2^n \to \{0,1\}$. *Then:*

$$\mathop{\mathbb{E}}_{L:\mathbb{F}_2^k \to \mathbb{F}_2^n}\left[\prod_{i=1}^{k} f_i(L(v_i))\right] \leqslant \min_{i \in [k]} \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f_i}(\alpha)^4$$

Key Property of Complexity 1 Matroids

[Green, Tao '06] showed that if a matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ has complexity 1, and if $A \subseteq \mathbb{F}_2^n$, then the number of linear maps $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ such that $L(v_i) \in A$ for all $i \in [k]$ is controlled by the Fourier spectrum of A.

Lemma (Generalized von Neumann)

Suppose the matroid $\mathcal{M} = \{v_1, \ldots, v_k\}$ has complexity 1 and let $f_1, \ldots, f_k : \mathbb{F}_2^n \to \{0, 1\}$. Then:

$$\mathop{\mathbb{E}}_{L:\mathbb{F}_2^k \to \mathbb{F}_2^n} \left[\prod_{i=1}^{k} f_i(L(v_i)) \right] \leqslant \min_{i \in [k]} \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f_i}(\alpha)^4$$

So, we need to be able to control the Fourier spectrum of f, or some functions related to it.

Green Regularity Lemma

Proved by Green in [Green '05], the Green Regularity Lemma states that for any $f : \mathbb{F}_2^n \to \{0, 1\}$, it is possible to find a subgroup $H \leqslant \mathbb{F}_2^n$ such that f restricted to most of the cosets of H is pseudo-random with respect to linear tests.

Green Regularity Lemma

Proved by Green in [Green '05], the Green Regularity Lemma states that for any $f : \mathbb{F}_2^n \to \{0,1\}$, it is possible to find a subgroup $H \leqslant \mathbb{F}_2^n$ such that f restricted to most of the cosets of H is pseudo-random with respect to linear tests.

Definition (Uniformity)

Given $\epsilon \in (0,1)$, a function $f : \mathbb{F}_2^n \to \{0,1\}$ is ϵ-*uniform* if for every nonzero $\alpha \in \mathbb{F}_2^n$, $|\widehat{f}(\alpha)| \leqslant \epsilon$.

Lemma (Regularity Lemma over $\mathbb{F}_2^n$)

Given $f : \mathbb{F}_2^n \to \{0,1\}$ and $\epsilon \in (0,1)$, there exists a subspace H of $\mathbb{F}_2^n$ of co-dimension at most $W(\epsilon^{-3})$ such that $\Pr_{g\in\mathbb{F}_2^n}[f_{g+H}$ is ϵ-uniform$] \geqslant 1-\epsilon$.

Testability of Complexity 1 Matroids

$\mathcal{M}$ matroid of complexity 1. To analyze soundness of the natural test, suppose f ϵ-far free from $\mathcal{M}$-free.

Testability of Complexity 1 Matroids

$\mathcal{M}$ matroid of complexity 1. To analyze soundness of the natural test, suppose f ϵ-far free from $\mathcal{M}$-free.

- Use Green Regularity Lemma to obtain subspace $H \leqslant \mathbb{F}_2^n$ of co-dimension at most $W(a(\epsilon)^{-3})$, where $a(\epsilon)$ specified later.

Testability of Complexity 1 Matroids

$\mathcal{M}$ matroid of complexity 1. To analyze soundness of the natural test, suppose f ϵ-far free from $\mathcal{M}$-free.

- Use Green Regularity Lemma to obtain subspace $H \leqslant \mathbb{F}_2^n$ of co-dimension at most $W(a(\epsilon)^{-3})$, where $a(\epsilon)$ specified later.
- Define reduced function $f^R : \mathbb{F}_2^n \to \{0,1\}$. For each $g \in \mathbb{F}_2^n$, if $f_{g+H} : H \to \{0,1\}$ is $a(\epsilon)$-uniform, then:

$$f^R_{g+H}(x) = \begin{cases} 0 & \text{if } \mu(f_{g+H}) \leqslant b(\epsilon) \\ f_{g+H} & \text{otherwise.} \end{cases}$$

where $b(\epsilon)$ specified later. If f_{g+H} is not $a(\epsilon)$-uniform, then define $f^R_{g+H} = 0$.

Testability of Complexity 1 Matroids Contd.

- If $a(\epsilon) + b(\epsilon) < \epsilon$, f^R contains $\mathcal{M}$ at some linear $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$.

Testability of Complexity 1 Matroids Contd.

- If $a(\epsilon) + b(\epsilon) < \epsilon$, f^R contains $\mathcal{M}$ at some linear $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$.
- Complexity 1 implies f contains $\mathcal{M}$ at many linear maps.

Testability of Complexity 1 Matroids Contd.

- If $a(\epsilon) + b(\epsilon) < \epsilon$, f^R contains $\mathcal{M}$ at some linear $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$.
- Complexity 1 implies f contains $\mathcal{M}$ at many linear maps.
  - Because of rounding, $f_i \stackrel{\text{def}}{=} f_{L(v_i)+H}$ must be $a(\epsilon)$-uniform and $b(\epsilon)$-dense for all $i \in [k]$

Testability of Complexity 1 Matroids Contd.

- If $a(\epsilon) + b(\epsilon) < \epsilon$, f^R contains $\mathcal{M}$ at some linear $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$.
- Complexity 1 implies f contains $\mathcal{M}$ at many linear maps.
  - Because of rounding, $f_i \stackrel{\text{def}}{=} f_{L(v_i)+H}$ must be $a(\epsilon)$-uniform and $b(\epsilon)$-dense for all $i \in [k]$
  - Lower-bound number of linear maps $\phi : \mathbb{F}_2^k \to H$ such that f contains $\mathcal{M}$ at $L + \phi$:

$$\mathop{\mathbb{E}}_{\phi:\mathbb{F}_2^k \to H}\left[\prod_{i\in[k]} f_i(\phi(v_i))\right] = \mathop{\mathbb{E}}_{\phi:\mathbb{F}_2^k \to H}\left[\prod_{i\in[k]} \left(\widehat{f_i}(\emptyset) + (f_i(\phi(v_i)) - \widehat{f_i}(\emptyset))\right)\right]$$

$$\geqslant \underbrace{b(\epsilon)^k}_{\text{(density)}} - \underbrace{(2^k - 1)a(\epsilon)^2}_{\text{(uniformity)}}$$

Testability of Complexity 1 Matroids Contd.

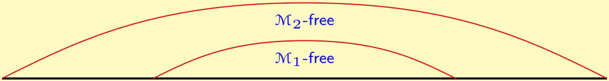
- If $a(\epsilon) + b(\epsilon) < \epsilon$, f^R contains $\mathcal{M}$ at some linear $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$.
- Complexity 1 implies f contains $\mathcal{M}$ at many linear maps.
  - Because of rounding, $f_i \stackrel{\text{def}}{=} f_{L(v_i)+H}$ must be $a(\epsilon)$-uniform and $b(\epsilon)$-dense for all $i \in [k]$
  - Lower-bound number of linear maps $\phi : \mathbb{F}_2^k \to H$ such that f contains $\mathcal{M}$ at $L + \phi$:

$$\mathop{\mathbb{E}}_{\phi:\mathbb{F}_2^k \to H}\left[\prod_{i\in[k]} f_i(\phi(v_i))\right] = \mathop{\mathbb{E}}_{\phi:\mathbb{F}_2^k \to H}\left[\prod_{i\in[k]}\left(\widehat{f}_i(\emptyset) + (f_i(\phi(v_i)) - \widehat{f}_i(\emptyset))\right)\right]$$

$$\geqslant \underbrace{b(\epsilon)^k}_{\text{(density)}} - \underbrace{(2^k-1)a(\epsilon)^2}_{\text{(uniformity)}}$$

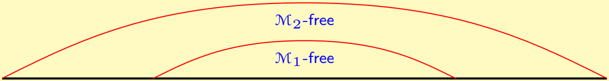
- Set $a(\epsilon) = \left(\frac{\epsilon}{2}\right)^k$, $b(\epsilon) = \frac{\epsilon}{2}$. Rejection probability is $2^{-O(W(\text{poly}(1/\epsilon)))}$. □

Separation between Matroid Freeness Properties



Given two matroids $\mathcal{M}_1$ and $\mathcal{M}_2$, suppose we want to show that there exist functions that are far from $\mathcal{M}_1$-free yet contain $\mathcal{M}_2$.

Separation between Matroid Freeness Properties



Given two matroids $\mathcal{M}_1$ and $\mathcal{M}_2$, suppose we want to show that there exist functions that are far from $\mathcal{M}_1$-free yet contain $\mathcal{M}_2$.

General strategy:

- Based on matroid $\mathcal{M}_1$, we construct a "canonical" function $f_{\mathcal{M}_1} : \mathbb{F}_2^n \to \{0, 1\}$ that is far from $\mathcal{M}_1$-free
- If the canonical function $f_{\mathcal{M}_1}$ contains $\mathcal{M}_2$ at some linear map, then there is a "matroid homomorphism" from $\mathcal{M}_2$ to $\mathcal{M}_1$
- Based upon properties of $\mathcal{M}_1$ and $\mathcal{M}_2$, we show impossibility of such matroid homomorphisms

Canonical Function that is far from $\mathcal{M}_1$-free

Given $\mathcal{M}_1$ represented by $v_1, \ldots, v_k \in \mathbb{F}_2^k$ and integer $n \geqslant k$, define the *canonical function* $f_{\mathcal{M}_1} : \mathbb{F}_2^n \to \{0, 1\}$ as $f_{\mathcal{M}_1}(x, y) = 1$ iff $x \in \{v_1, \ldots, v_k\}$, where $x \in \mathbb{F}_2^k$ and $y \in \mathbb{F}_2^{n-k}$.

Canonical Function that is far from $\mathcal{M}_1$-free

Given $\mathcal{M}_1$ represented by $v_1, \ldots, v_k \in \mathbb{F}_2^k$ and integer $n \geqslant k$, define the *canonical function* $f_{\mathcal{M}_1} : \mathbb{F}_2^n \to \{0, 1\}$ as $f_{\mathcal{M}_1}(x, y) = 1$ iff $x \in \{v_1, \ldots, v_k\}$, where $x \in \mathbb{F}_2^k$ and $y \in \mathbb{F}_2^{n-k}$.

Lemma

The canonical function $f_{\mathcal{M}_1}$ is 2^{-k}-far from $\mathcal{M}_1$-free.

Canonical Function that is far from $\mathcal{M}_1$-free

Given $\mathcal{M}_1$ represented by $v_1, \ldots, v_k \in \mathbb{F}_2^k$ and integer $n \geqslant k$, define the *canonical function* $f_{\mathcal{M}_1} : \mathbb{F}_2^n \to \{0, 1\}$ as $f_{\mathcal{M}_1}(x, y) = 1$ iff $x \in \{v_1, \ldots, v_k\}$, where $x \in \mathbb{F}_2^k$ and $y \in \mathbb{F}_2^{n-k}$.

Lemma

The canonical function $f_{\mathcal{M}_1}$ is 2^{-k}-far from $\mathcal{M}_1$-free.

Proof Sketch:

$f_{\mathcal{M}_1}$ contains $\mathcal{M}_1$ at the linear map $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ that sends x to $\langle x, 0 \rangle$. So, $f_{\mathcal{M}_1}$ is clearly not $\mathcal{M}_1$-free. Let g be a function 2^{-k}-close to f. Choose a random linear map $L_1 : \mathbb{F}_2^k \to \mathbb{F}_2^{n-k}$, and let $L : \mathbb{F}_2^k \to \mathbb{F}_2^n$ send x to $\langle x, L_1(x) \rangle$. With probability > 0, g contains $\mathcal{M}_1$ at L. $\square$

Matroid Homomorphisms

Definition

Suppose $\mathcal{M}_1$ and $\mathcal{M}_2$ matroids given by $v_1, \dots, v_k \in \mathbb{F}_2^k$ and $w_1, \dots, w_\ell \in \mathbb{F}_2^\ell$. A map $\phi : \{w_1, \dots, w_\ell\} \to \{v_1, \dots, v_\ell\}$ is a *matroid homomorphism* from $\mathcal{M}_2$ to $\mathcal{M}_1$ if for every set $T \subseteq [\ell]$ such that $\sum_{i \in T} w_i = 0$, then $\sum_{i \in T} \phi(w_i) = 0$.

Matroid Homomorphisms

Definition

Suppose $\mathcal{M}_1$ and $\mathcal{M}_2$ matroids given by $v_1, \ldots, v_k \in \mathbb{F}_2^k$ and $w_1, \ldots, w_\ell \in \mathbb{F}_2^\ell$. A map $\phi : \{w_1, \ldots, w_\ell\} \to \{v_1, \ldots, v_\ell\}$ is a *matroid homomorphism* from $\mathcal{M}_2$ to $\mathcal{M}_1$ if for every set $T \subseteq [\ell]$ such that $\sum_{i \in T} w_i = 0$, then $\sum_{i \in T} \phi(w_i) = 0$.

Lemma

If $f_{\mathcal{M}_1}$ contains $\mathcal{M}_2$, then $\mathcal{M}_2$ has a homomorphism to $\mathcal{M}_1$.

Matroid Homomorphisms

Definition

Suppose $\mathcal{M}_1$ and $\mathcal{M}_2$ matroids given by $v_1, \ldots, v_k \in \mathbb{F}_2^k$ and $w_1, \ldots, w_\ell \in \mathbb{F}_2^\ell$. A map $\phi : \{w_1, \ldots, w_\ell\} \to \{v_1, \ldots, v_\ell\}$ is a *matroid homomorphism* from $\mathcal{M}_2$ to $\mathcal{M}_1$ if for every set $T \subseteq [\ell]$ such that $\sum_{i \in T} w_i = 0$, then $\sum_{i \in T} \phi(w_i) = 0$.

Lemma

If $f_{\mathcal{M}_1}$ contains $\mathcal{M}_2$, then $\mathcal{M}_2$ has a homomorphism to $\mathcal{M}_1$.

Proof Sketch:

Suppose $f_{\mathcal{M}_1}$ contains $\mathcal{M}_2$ at L. If $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^k$ is the map sending $\langle x, y \rangle$ to x, then $\pi \circ L$ is the desired homomorphism. □

Homomorphism Impossibility Proof

Lemma

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the graphic matroids of C_a and C_b. If a, b are odd and $b < a$, then there is no homomorphism from $\mathcal{M}_2$ to $\mathcal{M}_1$.

Homomorphism Impossibility Proof

Lemma

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the graphic matroids of C_a and C_b. If a, b are odd and $b < a$, then there is no homomorphism from $\mathcal{M}_2$ to $\mathcal{M}_1$.

Proof Sketch:

Suppose ϕ is such a homomorphism. Then, if $v_1, \ldots, v_b$ represent $\mathcal{M}_2$, $v_1 + \cdots + v_b = 0$; so $\phi(v_1) + \cdots + \phi(v_b) = 0$ in $\mathcal{M}_1$. But then $b \geqslant a$ because there is no odd-sized set of dependent elements in $\mathcal{M}_1$ of size smaller than a.

Homomorphism Impossibility Proof

Lemma

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the graphic matroids of C_a and C_b. If a, b are odd and $b < a$, then there is no homomorphism from $\mathcal{M}_2$ to $\mathcal{M}_1$.

Theorem (Separation Theorem 1)

For every odd k, there exist functions g that are C_k-free but far from C_{k+2}-free.

Homomorphism Impossibility Proof

Lemma

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the graphic matroids of K_a and K_b. If $a \geqslant 3$ and $b \geqslant \binom{a}{2} + 2$, then there is no homomorphism from $\mathcal{M}_2$ to $\mathcal{M}_1$.

Homomorphism Impossibility Proof

Lemma

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the graphic matroids of K_a and K_b. If $a \geqslant 3$ and $b \geqslant \binom{a}{2} + 2$, then there is no homomorphism from $\mathcal{M}_2$ to $\mathcal{M}_1$.

Proof Sketch:

Suppose ϕ is such a homomorphism. By pigeonhole principle, there must exist incident edges e_i, e_j in K_b such that $\phi(e_i) = \phi(e_j)$. If f forms triangle with e_i and e_j in K_b, then because $e_i + e_j + f = 0$ in $\mathcal{M}_2$, $\phi(f) = \phi(e_i) + \phi(e_j) = 0 \notin \mathcal{M}_1$, an impossibility. □

Homomorphism Impossibility Proof

Lemma

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the graphic matroids of K_a and K_b. If $a \geqslant 3$ and $b \geqslant \binom{a}{2} + 2$, then there is no homomorphism from $\mathcal{M}_2$ to $\mathcal{M}_1$.

Theorem (Separation Theorem 2)

If $k \geqslant 3$ and $\ell \geqslant \binom{k}{2} + 2$, then there is a function that is K_ℓ-free but far from K_k-free.

Conclusions

- Goal is to find the "smallest" set of invariances and other conditions that explain testability.
- We show linearity is not a necessary condition for testability and suggest that linear invariance solely might be sufficient for properties with small description.
- We show that the class of testable matroid-freeness properties is provably infinite.