Invariance in Property Testing

Madhu Sudan
Microsoft/MIT

Happy Birthday, Laci!

- Thanks for …
  - All the wonderful results !
    - (PCP, Alg. Group Theory)
  - The inspiring talks !
  - The entertaining + educational writings !

Modern challenge to Algorithm Design

- Data = Massive; Computers = Tiny
  - How can tiny computers analyze massive data?
  - Only option: Design sublinear time algorithms.
    - Algorithms that take less time to analyze data, than it takes to read/write all the data.
    - Can such algorithms exist?

Yes! Polling …

- Is the majority of the population Red/Blue
  - Can find out by random sampling.
  - Sample size $\propto$ margin of error
    - Independent of size of population

- Other similar examples: (can estimate other moments …)

Recent "novel" example

- Can test for homomorphisms:
  - Given: $f: G \rightarrow H$ (G,H finite groups), is f essentially a homomorphism?
  - Test:
    - Pick x,y in G uniformly, ind. at random;
    - Verify $f(x) \cdot f(y) = f(x \cdot y)$
  - Completeness: accepts homomorphisms w.p. 1
    - (Obvious)
  - Soundness: Rejects f w.p prob. Proportional to its "distance" (margin) from homomorphisms.
    - (Not obvious)

Property Testing

- Data = a function from D to R:
  - Property $P \subseteq \{D \to R\}$
- Distance
  - $\delta(f,g) = \Pr_{x \in \mathbf{D}} [f(x) \neq g(x)]$
  - $\delta(f,P) = \min_{g \in \mathbf{P}} [\delta(f,g)]$
  - f is ε-close to g ($f \approx_\epsilon g$) iff $\delta(f,g) \leq \varepsilon$.
- Local testability:
  - P is (k, ε, δ)-locally testable if $\exists$ k-query test T
    - $f \in P \Rightarrow T^f$ accepts w.p. $1-\varepsilon$.
    - $\delta(f,P) > \delta \Rightarrow T^f$ accepts w.p. ε.
- Notes: want $k(\varepsilon, \delta) = O(1)$ for $\varepsilon,\delta = \Omega(1)$.

Brief History

- [Blum,Luby,Rubinfeld – S'90]
  - Linearity + application to program testing
- [Babai,Fortnow,Lund – F'90]
  - Multilinearity + application to PCPs (MIP).
- [Rubinfeld+S.]
  - Low-degree testing
- [Goldreich,Goldwasser,Ron]
  - Graph property testing
- Since then … many developments
  - Graph properties
  - Statistical properties
  - …
  - More algebraic properties

Graph Property Testing

- Initiated by [GoldreichGoldwasserRon]
- Initial examples:
  - Is graph bipartite?
  - Is it 3-colorable?
  - Is it triangle-free (underlying theorem dates back to 80s)?
- Many intermediate results
  - Close ties to Szemeredi's regularity lemma
- Culmination: [AlonFisherNewmanSzegedy]:
  - Characterization of all testable properties in terms of regularity.

Specific Directions in Algebraic P.T.

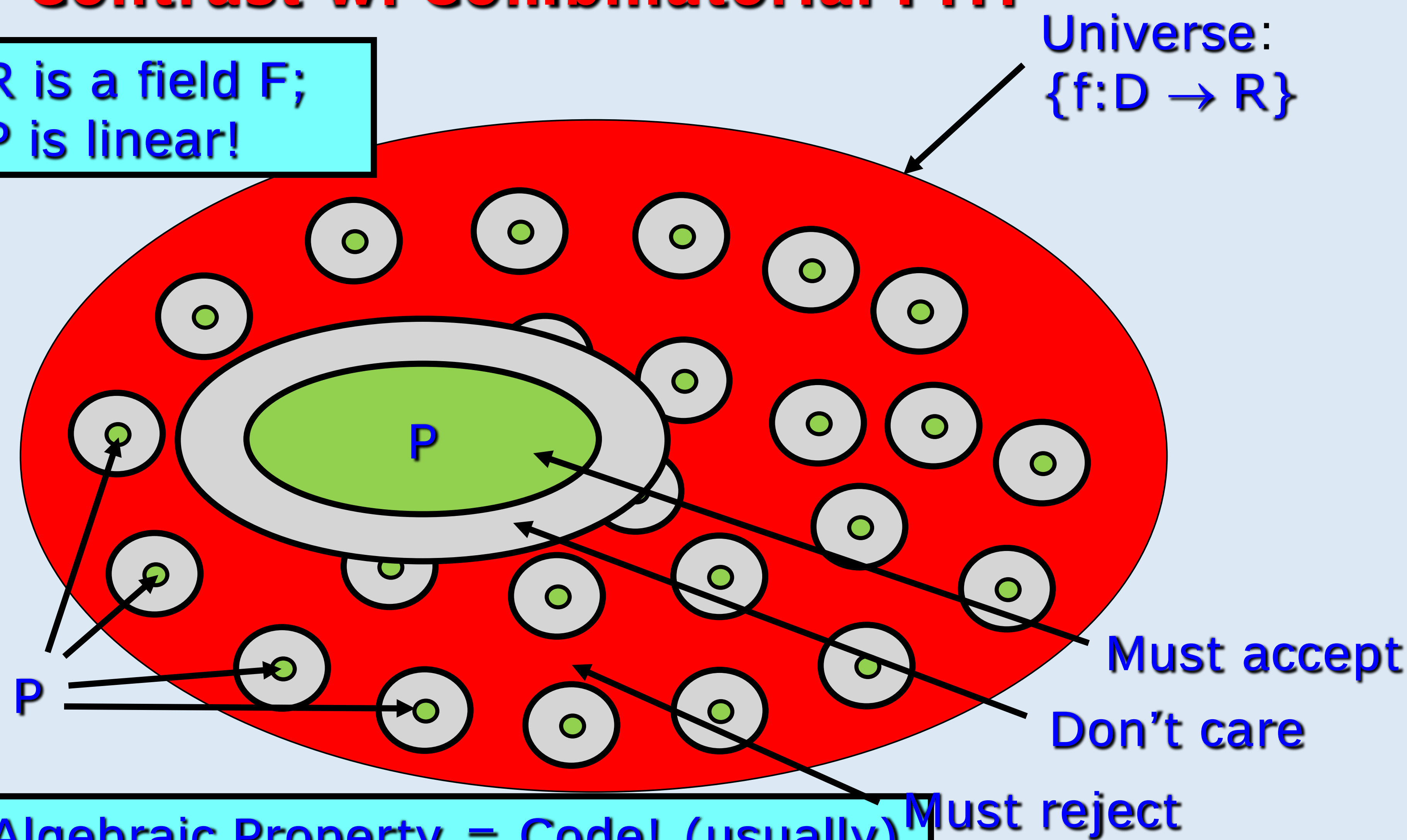
- Fewer results
- More Properties
  - Low-degree ($d < q$) functions [RS]
  - Moderate-degree ($q < d < n$) functions
    - $q=2$: [AKKLR]
    - General q: [KR, JPRZ]
  - Long code/Dictator/Junta testing [BGS,PRS]
  - BCH codes (Trace of low-deg. poly.) [KL]
- Better Parameters (motivated by PCPs).
  - #queries, high-error, amortized query complexity, reduced randomness.

My concerns …

- Relatively few results …
  - Why can’t we get “rich” class of properties that are all testable?
  - Why are proofs so specific to property being tested?
- What made Graph Property Testing so well-understood?
- What is “novel” about Property Testing, when compared to “polling”?

Contrast w. Combinatorial P.T.

R is a field F;
P is linear!



Algebraic Property = Code! (usually)

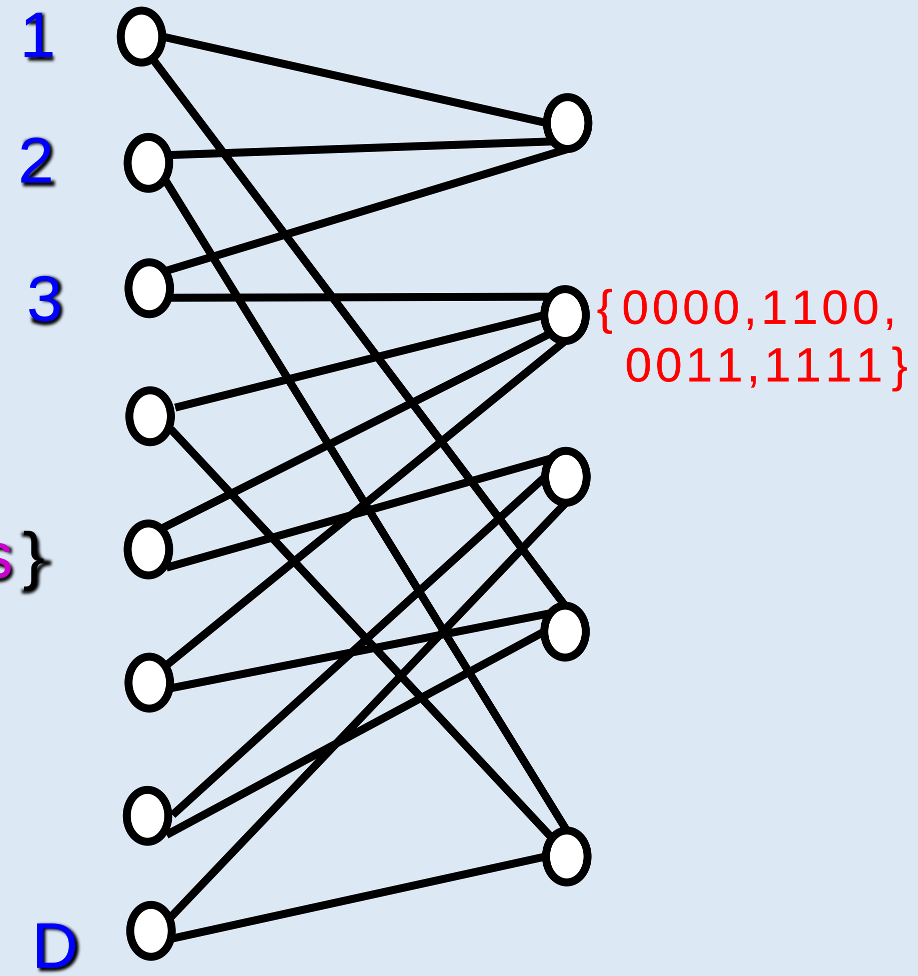
Basic Implications of Linearity [BHR]

- If P is linear, then:
  - Tester can be made non-adaptive.
  - Tester makes one-sided error
    - ($f \in P \Rightarrow$ tester always accepts).
- Motivates:
  - Constraints:
    - k-query test => constraint of size k:
      - value of f at $\alpha_1, \ldots \alpha_k$ constrained to lie in subspace.
  - Characterizations:
    - If non-members of P rejected with positive probability, then P characterized by local constraints.
      - functions satisfying all constraints are members of P.

Pictorially

- f = assgm't to left
- Right = constraints
- Characterization of P:

  P = {f sat. all constraints}



Sufficient conditions?

- Linearity + k-local characterization
  $\Rightarrow$ k-local testability?

- [BHR] No!
  - Elegant use of expansion
  - Rule out obvious test; but also <u>any</u> test … of <u>any</u> "q(k)"-locality

- Why is characterization insufficient?
  - Lack of symmetry?

Example motivating symmetry

- Conjecture (AKKLR '96):
  - Suppose property P is a vector space over F_2;
  - Suppose its "invariant group" is "2-transitive".
  - Suppose P satisfies a k-ary constraint
    - $\forall f \in P,\ f(\alpha_1) + \cdots + f(\alpha_k) = 0.$
  - Then P is $(q(k), \epsilon(k,\delta),\delta)$-locally testable.
- Inspired by "low-degree" test over F_2. Implied all previous algebraic tests (at least in weak forms).

Invariances

- Property P invariant under permutation (function) π: $D \to D$, if

$$f \in P \Rightarrow f \circ \pi \in P$$

- Property P invariant under group G if

$$\forall \pi \in G, \text{ P is invariant under } \pi.$$

- Can ask: Does invariance of P w.r.t. "nice" G leads to local testability?

Invariances are the key?

- "Polling" works well when (because) invariant group of property is the full symmetric group.
- Modern property tests work with much smaller group of invariances.
- Graph property ~ Invariant under vertex renaming.
- Algebraic Properties & Invariances?

Abstracting Algebraic Properties

- [Kaufman & S.]

- Range is a field F and P is F-linear.
- Domain is a vector space over F (or some field K extending F).

- Property is invariant under affine (sometimes only linear) transformations of domain.

- "Property characterized by single constraint, and its orbit under affine (or linear) transformations."

Invariance, Orbits and Testability

- Single constraint implies many
  - One for every permutation $\pi \in \mathrm{Aut}(P)$:
    - "Orbit of a constraint C"

$$= \{C \circ \pi \mid \pi \in \mathrm{Aut}(P)\}$$

- Extreme case:
  - Property characterized by single constraint + its orbit: "Single orbit feature"
    - Most algebraic properties have this feature.
    - W.l.o.g. if domain = vector space over small field.

Example: Degree d polynomials

- **Constraint:** When restricted to a small dimensional affine subspace, function is polynomial of degree d (or less).
  - #dimensions $\leq d/(K - 1)$
- **Characterization:** If a function satisfies above for every small dim. subspace, then it is a degree d polynomial.
- **Single orbit:** Take constraint on any one subspace of dimension $d/(K-1)$; and rotate over all affine transformations.

Some results

- If P is affine-invariant and has k-single orbit feature (characterized by orbit of single k-local constraint); then it is $(k, \delta/k^3, \delta)$-locally testable.
  - Unifies previous algebraic tests (in weak form) with single proof.

Analysis of Invariance-based test

- Property P given by $\alpha_1,\ldots,\alpha_k$; $V \in F^k$
- $P = \{f \mid f(A(\alpha_1)) \ldots f(A(\alpha_k)) \in V, \forall$ affine $A:K^n \to K^n\}$
- $Rej(f) = \text{Prob}_A\ [\ f(A(\alpha_1)) \ldots f(A(\alpha_k))$ not in $V\ ]$
- Wish to show: If $Rej(f) < 1/k^3$,
  then $\delta(f,P) = O(Rej(f))$.

BLR Analog

- $\mathrm{Rej}(f) = \Pr_{x,y}[\, f(x) + f(y) \neq f(x+y)] < \epsilon$
- Define $g(x) = \mathrm{majority}_y\ \{\mathrm{Vote}_x(y)\}$, where $\mathrm{Vote}_x(y) = f(x+y) - f(y)$.
- Step 0: Show $\delta(f,g)$ small
- Step 1: $\forall\, x$, $\Pr_{y,z}[\mathrm{Vote}_x(y) \neq \mathrm{Vote}_x(z)]$ small.
- Step 2: Use above to show g is well-defined and a homomorphism.

BLR Analysis of Step 1

- Why is $f(x+y) - f(y) = f(x+z) - f(z)$, usually?

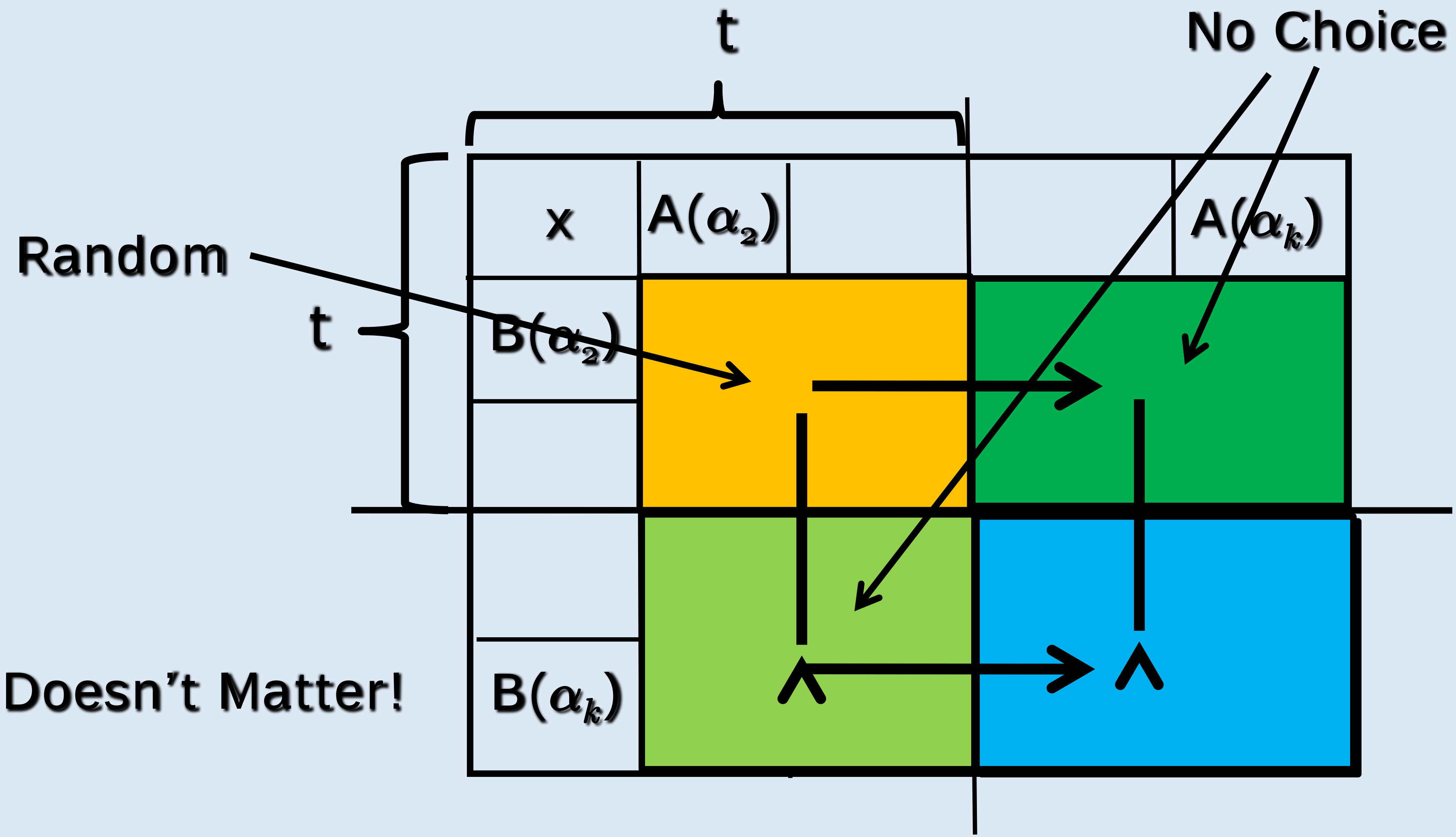
?	$f(z)$	$- f(x+z)$
$f(y)$	0	$-f(y)$
$- f(x+y)$	$-f(z)$	$f(x+y+z)$

Generalization

- $g(x) = \beta$ that maximizes, over A s.t. $A(\alpha_1) = x$,
  $\Pr_A [\beta, f(A(\alpha_2)),\ldots,f(A(\alpha_k)) \in V]$
- Step 0: $\delta(f,g)$ small.
- $\mathrm{Vote}_x(A) = \beta$ s.t. β, $f(A(\alpha_2))\ldots f(A(\alpha_k)) \in V$
  (if such β exists)
- Step 1 (key): $\forall$ x, whp $\mathrm{Vote}_x(A) = \mathrm{Vote}_x(B)$.
- Step 2: Use above to show $g \in P$.

Matrix Magic?

Say $A(\alpha_1)$ … $A(\alpha_t)$ independent; rest dependent



Some results

- If P is affine-invariant and has k-single orbit feature (characterized by orbit of single k-local constraint); then it is $(k, \delta/k^3, \delta)$-locally testable.
    - Unifies previous algebraic tests with single proof.

- If P is affine-invariant over K and has a single k-local constraint, then it is has a q-single orbit feature (for some $q = q(K,k)$)
    - (explains the AKKLR optimism)

Subsequent results

- [GrigorescuKaufmanS.; CCC08]: Counterexample to AKKLR Conjecture
- [GrigorescuKaufmanS., Random09]: Single orbit characterization of some BCH (and other) codes.
- [Ben-SassonS.]: Limitations on rate of affine-invariant codes.

- [KaufmanWigderson]: LDPC codes with invariance (not affine-invariant)
- [BhattacharyyaChenS.Xie, Shapira]: Affine-invariant non-linear properties.

Broad directions to consider

- Is every locally characterized affine-invariant property testable?
- Is every single-orbit characterized affine-invariant property testable?
- What groups of invariances lead to testability?
- In general … seek invariances

Thanks