Regularity for Polynomials and Linear Forms

Pooya Hatami

(joint work with Hamed Hatami and Shachar Lovett)

October 18, 2014
Setting

\* \( f : \mathbb{F}^n \rightarrow \mathbb{R} \)

\* \( \mathbb{F} = \mathbb{F}_p \).
\* \( p \) is a fixed prime, and \( n \) is large

\* \( e(x) := e_p(x) := e^{2\pi i x/p} \).

\( x, y, z \in \mathbb{F}^n, X, Y, Z \in (\mathbb{F}^n)^k \).
Fourier Analysis and Higher-order Fourier Analysis
Fourier Analysis

Study a function by looking at how it correlates with linear functions.

\[ f : F_n \to \mathbb{R}, \quad f(x) = \sum_{\sigma \in F_n} \hat{f} \chi_{\sigma} \]

\[ \chi_{\sigma} = e^{\langle \sigma, x \rangle} = e^{\sum_{i} \sigma_i x_i} \]

Applications

Useful in controlling several expressions regarding a given function, such as approximate Linearity, density of 3-term APs.
Fourier Analysis

Study a function by looking at how it correlates with linear functions.

\[ f : \mathbb{F}^n \to \mathbb{R}, \quad f = \sum_{\sigma \in \mathbb{F}^n} \hat{f}_\sigma \chi_\sigma. \]

\[ \chi_\sigma = e(\langle \sigma, x \rangle) = e(\sum_i \sigma_i x_i) \]
Fourier Analysis

Study a function by looking at how it correlates with linear functions.

\[ f : \mathbb{F}^n \to \mathbb{R}, \]
\[ f = \sum_{\sigma \in \mathbb{F}^n} \hat{f}_\sigma \chi_\sigma. \]

\[ \chi_\sigma = e(\langle \sigma, x \rangle) = e(\sum_i \sigma_i x_i) \]

Applications

Useful in controlling several expressions regarding a given function, such as approximate Linearity, density of 3-term APs.
Approximate Linearity (As seen in an analysis of BLR test)

\[ f : \mathbb{F}_2^n \rightarrow \{0, 1\}, \text{ letting } g(x) = (-1)^{f(x)} \]

\[
\Pr_{x,y}(f(x + y) = f(x) + f(y)) = \frac{1}{2} + \frac{1}{2} \mathbb{E}_{x,y}(g(x + y)g(x)g(y)) \\
= \frac{1}{2} + \frac{1}{2} \sum_{\sigma_1, \sigma_2, \sigma_3 \in \mathbb{F}_2^n} \hat{g}_{\sigma_1} \hat{g}_{\sigma_2} \hat{g}_{\sigma_3} \mathbb{E}_{x,y} e_2(\sigma_1^t x + \sigma_2^t y + \sigma_3^t (x + y)) \\
= \frac{1}{2} + \frac{1}{2} \sum_{\sigma \in \mathbb{F}_2^n} \hat{g}_\sigma^3 \leq \max_{\sigma} \hat{g}_\sigma
\]
Approximate Linearity (As seen in an analysis of BLR test)

\[ f : \mathbb{F}_2^n \rightarrow \{0, 1\}, \text{ letting } g(x) = (-1)^{f(x)} \]

\[
\Pr_{x,y}(f(x + y) = f(x) + f(y)) = \frac{1}{2} + \frac{1}{2} \mathbb{E}_{x,y}(g(x + y)g(x)g(y))
\]

\[
= \frac{1}{2} + \frac{1}{2} \sum_{\sigma_1, \sigma_2, \sigma_3 \in \mathbb{F}_2^n} \hat{g}_{\sigma_1} \hat{g}_{\sigma_2} \hat{g}_{\sigma_3} \mathbb{E}_{x,y}e_2(\sigma_1^t x + \sigma_2^t y + \sigma_3^t (x + y))
\]

\[
= \frac{1}{2} + \frac{1}{2} \sum_{\sigma \in \mathbb{F}_2^n} \hat{g}_{\sigma} \leq \max_{\sigma} \hat{g}_{\sigma}
\]

Correlation with characters captures approximate linearity.
3-term Arithmetic Progressions

$A \subseteq \mathbb{F}_p^n$, let $g(x) = 1_A(x)$. 

Correlation with characters can control density of 3-term APs.
3-term Arithmetic Progressions

\( A \subset \mathbb{F}_p^n \), let \( g(x) = 1_A(x) \).

\[
\Pr_{x,d}(x, x + d, x + 2d \in A) = \mathbb{E}_{x,d} g(x)g(x + d)g(x + 2d)
\]
\[
= \sum_{\sigma_1, \sigma_2, \sigma_3} \hat{g}_{\sigma_1} \hat{g}_{\sigma_2} \hat{g}_{\sigma_3} \mathbb{E}_{x,y} e_2(\sigma_1^t x + \sigma_2^t (x + d) + \sigma_3^t (x + 2d))
\]
\[
= \sum_{\sigma} |\hat{g}_\sigma|^2 \hat{g}_{-2\sigma} = \hat{g}_0^3 + \sum_{\sigma \neq 0} |\hat{g}_\sigma|^2 \hat{g}_{-2\sigma}
\]

Correlation with characters can control density of 3-term APs.
3-term Arithmetic Progressions

$A \subset \mathbb{F}_p^n$, let $g(x) = 1_A(x)$.

$$
\Pr_{x,d}(x, x + d, x + 2d \in A) = \mathbb{E}_{x,d} g(x) g(x + d) g(x + 2d) \\
= \sum_{\sigma_1, \sigma_2, \sigma_3} \hat{g}_{\sigma_1} \hat{g}_{\sigma_2} \hat{g}_{\sigma_3} \mathbb{E}_{x,y} e_2(\sigma_1 x + \sigma_2 (x + d) + \sigma_3 (x + 2d)) \\
= \sum_{\sigma} |\hat{g}_{\sigma}|^2 \hat{g}_{-2\sigma} = \hat{g}_0^3 + \sum_{\sigma \neq 0} |\hat{g}_{\sigma}|^2 \hat{g}_{-2\sigma}
$$

Correlation with characters can control density of 3-term APs.
Higher-order Fourier Analysis

- Introduce higher degree phase polynomials, $e(P(x))$ instead of characters $e(\sigma^t x)$.
- Study a function by looking at how it correlates with these higher-order terms.
Higher-order Fourier Analysis

- Introduce higher degree phase polynomials, $e(P(x))$ instead of characters $e(\sigma^t x)$.
- Study a function by looking at how it correlates with these higher-order terms.
- More complex behavior, such as 4-APs.
Higher-order Fourier Analysis

- Introduce higher degree phase polynomials, $e(P(x))$ instead of characters $e(\sigma^t x)$.
- Study a function by looking at how it correlates with these higher-order terms.
- More complex behavior, such as 4-APs.

- Need approximation of functions by a linear combination of these higher-order polynomials.
Decomposition Theorems as a result of Inverse Theorems

[Bergelson, Green, Samorodnitsky, Szegedy, Tao, Ziegler]

\[ f \approx_{U^{d+1}} \Gamma(P_1, \ldots, P_C), \]

where \( P_1, \ldots, P_C \) are degree \( \leq d \) polynomials.
Decomposition Theorems as a result of Inverse Theorems

[Bergelson, Green, Samorodnitsky, Szegedy, Tao, Ziegler]

\[ f \approx_U d+1 \Gamma(P_1, \ldots, P_C), \]

where \( P_1, \ldots, P_C \) are degree \( \leq d \) polynomials.

\[ f \approx \sum_{\sigma \in \mathbb{F}_C} \hat{\Gamma}_\sigma e\left( \sum_{i \in [C]} \sigma_i P_i \right). \]  \hspace{1cm} (1)
Decomposition Theorems as a result of Inverse Theorems

[Bergelson, Green, Samorodnitsky, Szegedy, Tao, Ziegler]

\[ f \approx_{U^{d+1}} \Gamma(P_1, \ldots, P_C), \]

where \( P_1, \ldots, P_C \) are degree \( \leq d \) polynomials.

\[ f \approx \sum_{\sigma \in \mathbb{F}_C} \hat{\Gamma}_\sigma e\left( \sum_{i \in [C]} \sigma_i P_i \right). \quad (1) \]

No Orthogonality, unlike in classical Fourier analysis!
Regularity [Green-Tao, Kaufman-Lovett]

High-rank polynomials are unbiased

- $|\mathbb{E}_x e(P(x))| < \epsilon$
- $\Pr(P = a) \approx 1/p$
High-rank polynomials are unbiased

- $|\mathbb{E}_x e(P(x))| < \epsilon$
- $\Pr(P = a) \approx 1/p$

**Near-Orthogonality**: High-rank collection of polynomials provide near-orthogonality.

**Approximate Equidistribution**: For high-rank collection of polynomials, $(P_1(x), \ldots, P_C(x))$ is distributed close to uniform on $\mathbb{F}^C$. 
High-rank polynomials are unbiased

- $|\mathbb{E}_x e(P(x))| < \epsilon$
- $\Pr(P = a) \approx 1/p$

Near-Orthogonality: High-rank collection of polynomials provide near-orthogonality.

Approximate Equidistribution: For high-rank collection of polynomials, $(P_1(x), ..., P_C(x))$ is distributed close to uniform on $\mathbb{F}^C$.

Regularization [Green-Tao, Kaufman-Lovett]

Any collection of polynomials can be refined to a high-rank collection.
High-rank polynomials are unbiased

- \(|\mathbb{E}_x e(P(x))| < \epsilon\)
- \(\Pr(P = a) \approx 1/p\)

Near-Orthogonality: High-rank collection of polynomials provide near-orthogonality.

Approximate Equidistribution: For high-rank collection of polynomials, \((P_1(x), \ldots, P_C(x))\) is distributed close to uniform on \(\mathbb{F}^C\).

Regularization [Green-Tao, Kaufman-Lovett]

Any collection of polynomials can be refined to a high-rank collection.

Can assume \(P_1, \ldots, P_C\) in \(f \approx \Gamma(P_1, \ldots, P_C)\) is a high-rank collection.
Figure: Approximation by polynomials: $\Gamma(P_1, \ldots, P_C)$
Figure: Regular refinement: $\Gamma'(Q_1, \ldots, Q_c)$
But is this sufficient for applications?
But is this sufficient for applications?

Developed in order to understand more complex averages.
Density of Linear Patterns, such as APs

\[ \mathbb{E}_{x,y \in \mathbb{F}^n} f(x)f(x+y) \cdots f(x+(k-1)y), \tag{2} \]
Density of Linear Patterns, such as APs

\[ \mathbb{E}_{X \in (\mathbb{F}_n)^k} f(L_1(X)) f(L_2(X)) \cdots f(L_m(X)), \]

(2)

\[ L_i = (\lambda_{i,1}, \ldots, \lambda_{i,k}) \in \mathbb{F}^k \text{ is a linear form and } L_i(X) = \sum_{j=1}^k \lambda_{i,j} x_j. \]
Density of Linear Patterns, such as APs

\[ \mathbb{E}_{X \in (\mathbb{F}^n)^k} f(L_1(X)) f(L_2(X)) \cdots f(L_m(X)), \quad (2) \]

\[ L_i = (\lambda_{i,1}, \ldots, \lambda_{i,k}) \in \mathbb{F}^k \text{ is a linear form and } L_i(X) = \sum_{j=1}^k \lambda_{i,j} x_j. \]

Using \( f \approx \sum_{\sigma \in \mathbb{F}^C} \hat{\Gamma}_\sigma e(\sum_{i \in [C]} \sigma_i P_i) \) we have

\[ (2) \approx \sum_{\sigma_1, \ldots, \sigma_m \in \mathbb{F}^C} C_{\sigma_1, \ldots, \sigma_m} e(\sum_{\sigma_j,i \in [m], j \in [C]} \sigma_{j,i} P_i(L_j(X))), \]
Density of Linear Patterns, such as APs

\[ \mathbb{E}_{X \in (\mathbb{F}^n)^k} f(L_1(X))f(L_2(X)) \cdots f(L_m(X)), \]  

(2)

\[ L_i = (\lambda_{i,1}, \ldots, \lambda_{i,k}) \in \mathbb{F}^k \text{ is a linear form and } L_i(X) = \sum_{j=1}^{k} \lambda_{i,j}x_i. \]

Using \( f \approx \sum_{\sigma \in \mathbb{F}^C} \hat{\Gamma}_\sigma e(\sum_{i \in [C]} \sigma_i P_i) \) we have

\[ (2) \approx \sum_{\sigma_1, \ldots, \sigma_m \in \mathbb{F}^C} C_{\sigma_1, \ldots, \sigma_m} e(\sum_{i \in [m], j \in [C]} \sigma_{j,i} P_i(L_j(X))), \]

We need stronger near-orthogonality over sets of linear forms!
Property Testing

Every locally characterizable "algebraic" property is testable.

Test "algebraic" properties of $f$ by querying it over a random subspace.

Let $L_1, \ldots, L_p$ be the points of a random $V$.

$f \approx \Gamma(P_1(x), \ldots, P_C(x)).$

We need to understand the joint distribution $(P_i(L_j(x)))_{i \in [C], j \in [p]}$. 
Property Testing

[BFL, BFHHL] Every locally characterizable “algebraic” property is testable.
Studying a function by Sampling a Subspace

Property Testing

[BFL, BFHHL] Every locally characterizable “algebraic” property is testable.

Test “algebraic” properties of $f$ by querying it over a random subspace.
Property Testing

[BFL, BFHHL] Every locally characterizable “algebraic” property is testable.

Test “algebraic” properties of $f$ by querying it over a random subspace.

- Need to analyze the distribution of $f|_V$.
- Let $L_1, \ldots, L_{p^k}$ be the points of a random $V$.
- $f \approx \Gamma(P_1(x), \ldots, P_C(x))$. 
Property Testing

[BFL, BFHHL] Every locally characterizable “algebraic” property is testable.

Test “algebraic” properties of $f$ by querying it over a random subspace.

- Need to analyze the distribution of $f|_V$.
- Let $L_1, \ldots, L_{p^k}$ be the points of a random $V$.
- $f \approx \Gamma(P_1(x), \ldots, P_C(x))$.
- We need to understand the joint distribution $(P_i(L_j(X)))_{i \in [C], j \in [p^k]}$. 
We need to understand the joint distribution of

\[
\begin{pmatrix}
P_1(L_1(X)) & \ldots & P_C(L_1(X)) \\
P_1(L_2(X)) & \ldots & P_C(L_2(X)) \\
\vdots & & \vdots \\
P_1(L_m(X)) & \ldots & P_C(L_m(X))
\end{pmatrix}
\]

\[\text{[Kaufman-Lovett, Green-Tao]}:\]

If \(P_1, \ldots, P_C\) are of "high rank", then \(P_1(X), \ldots, P_C(X)\), are almost independent.

The entries in each row are almost independent.

Cannot expect almost independence for all entries!

\[\text{e.g. } \deg(P) = 1, \text{ then } P(x+y) + P(z) = P(x) + P(y+z).\]

\[\text{e.g. } \deg(P) < d, \text{ then } \sum_{\omega \in \{0, 1\}^{d+1}} (-1)^{|\omega|} P(X + \sum_{i \in \omega} Y_i) = 0.\]
We need to understand the joint distribution of

\[
\begin{pmatrix}
P_1(L_1(X)) & \ldots & P_C(L_1(X)) \\
P_1(L_2(X)) & \ldots & P_C(L_2(X)) \\
\vdots & & \vdots \\
P_1(L_m(X)) & \ldots & P_C(L_m(X))
\end{pmatrix}
\]

- [Kaufman-Lovett, Green-Tao]:
  If $P_1, \ldots, P_C$ are of “high rank”, then $P_1(X), \ldots, P_C(X)$, are almost independent.
We need to understand the joint distribution of

\[
\begin{pmatrix}
P_1(L_1(X)) & \ldots & P_C(L_1(X)) \\
P_1(L_2(X)) & \ldots & P_C(L_2(X)) \\
\vdots & & \vdots \\
P_1(L_m(X)) & \ldots & P_C(L_m(X))
\end{pmatrix}
\]

[Kaufman-Lovett, Green-Tao]:
If \( P_1, \ldots, P_C \) are of “high rank”, then \( P_1(X), \ldots, P_C(X) \), are almost independent,
The entries in each row are almost independent.
We need to understand the joint distribution of

\[
\begin{pmatrix}
  P_1(L_1(X)) & \ldots & P_C(L_1(X)) \\
  P_1(L_2(X)) & \ldots & P_C(L_2(X)) \\
  \vdots & & \vdots \\
  P_1(L_m(X)) & \ldots & P_C(L_m(X))
\end{pmatrix}
\]

[Kaufman-Lovett, Green-Tao]:
If \( P_1, \ldots, P_C \) are of “high rank”, then \( P_1(X), \ldots, P_C(X) \), are almost independent,
The entries in each row are almost independent.

Cannot expect almost independence for all entries!
We need to understand the joint distribution of

$$
\begin{pmatrix}
P_1(L_1(X)) & \ldots & P_C(L_1(X)) \\
P_1(L_2(X)) & \ldots & P_C(L_2(X)) \\
\vdots & & \vdots \\
P_1(L_m(X)) & \ldots & P_C(L_m(X))
\end{pmatrix}
$$

- [Kaufman-Lovett, Green-Tao]:
  If $P_1, \ldots, P_C$ are of “high rank”, then $P_1(X), \ldots, P_C(X)$, are almost independent,
  The entries in each row are almost independent.

- Cannot expect almost independence for all entries!
  - e.g. $\deg(P) = 1$, then $P(x + y) + P(z) = P(x) + P(y + z)$.
  - e.g. $\deg(P) < d$, then $\sum_{\omega \in \{0,1\}^{d+1}} (-1)^{|\omega|} P(X + \sum_{i \in \omega} Y_i) = 0$. 
Theorem (Strong Regularity)

For a high-rank collection of polynomials, up to a controllable error, these degree related dependencies are the only dependencies.
Theorem (Strong Regularity)

For a high-rank collection of polynomials, up to a controlable error,

These degree related dependencies are the only dependencies

- Large values of $p$: [Hamed Hatami, Lovett 2011].
Theorem (Strong Regularity)

For a high-rank collection of polynomials, up to a controllable error,

**These degree related dependencies are the only dependencies**

- Large values of $p$: [Hamed Hatami, Lovett 2011].
- General $p$, but affine systems of linear forms: [Bhattacharyya, Fischer, Hamed Hatami, P. H., and Lovett 2013].
Theorem (Strong Regularity)

For a high-rank collection of polynomials, up to a controllable error,

These degree related dependencies are the only dependencies

- Large values of $p$: [Hamed Hatami, Lovett 2011].
- General $p$, but affine systems of linear forms: [Bhattacharyya, Fischer, Hamed Hatami, P. H., and Lovett 2013].
- General case: [H. Hatami, P.H., and Lovett, General systems of linear forms].
Columns are almost independently distributed.
Theorem. (Near Orthogonality [Hamed Hatami, P.H., Lovett])

$P_1, \ldots, P_C$ be a high-rank set of polynomials. Let

$$P_{\wedge}(X) = \sum_{i \in [C], j \in [m]} \lambda_{i,j} P_i(L_j(X)).$$

Then $P_{\wedge}(X) \equiv 0$ or $\left| \left| E_{X \in (\mathbb{F}_n^m)} e(P_{\wedge}(X)) \right| \right| < \epsilon$ if and only if the same is true for any collection of same degree polynomials.
Theorem. (Near Orthogonality [Hamed Hatami, P.H., Lovett])

Let \( P_1, \ldots, P_C \) be a high-rank set of polynomials. Let

\[
P_\Lambda(X) = \sum_{i \in [C], j \in [m]} \lambda_{i,j} P_i(L_j(X)).
\]

Then

\[
P_\Lambda \equiv 0 \quad \text{or} \quad \left| \mathbb{E}_{X \in (\mathbb{F}^n)^\ell} [e(P_\Lambda)] \right| < \epsilon
\]
Theorem. (Near Orthogonality [Hamed Hatami, P.H., Lovett])

Let \( P_1, \ldots, P_C \) be a high-rank set of polynomials. Let

\[
P_{\Lambda}(X) = \sum_{i \in [C], j \in [m]} \lambda_{i,j} P_i(L_j(X)).
\]

Then

\[
P_{\Lambda} \equiv 0 \quad \text{or} \quad \left| \mathbb{E}_{X \in (\mathbb{F}_n)^\ell} [e(P_{\Lambda})] \right| < \epsilon
\]

\( P_{\Lambda} \equiv 0 \) if and only if the same is true for any collection of same degree polynomials.
Proof Ideas.

- Reduce to the case when $|L_j| \leq \deg(P_i)$.
- For example, $Q(2x + z) = 2Q(x) + Q(z) - 2Q(x + z) - Q(2x)$.
- Reduce to the case that the polynomials are homogeneous.
- Applications of certain derivative operators $D_i$ such that $\left| E \left[ e^{\left( D_1 \cdots D_d \Lambda \left( X \right) \right)} \right] \right|_2 \leq \left| e^{\left( \sum \lambda_i P_i(L_j(X)) \right)} \right|_2^{d^2}$.

\[ P_{\Lambda}(X) = \sum_{i \in [C], j \in [m]} \lambda_{i,j} P_i(L_j(X)) \]
Proof Ideas.

- Reduce to the case when $|L_j| \leq \deg(P_i)$.
  - e.g. $Q(2x + z) = 2Q(x) + Q(z) - 2Q(x + z) - Q(2x)$. 

\[
P_{\Lambda}(X) = \sum_{i \in [C], j \in [m]} \lambda_{i,j} P_i(L_j(X))
\]
Proof Ideas.

- Reduce to the case when $|L_j| \leq \deg(P_i)$. 
  - e.g. $Q(2x + z) = 2Q(x) + Q(z) - 2Q(x + z) - Q(2x)$.

- Reduce to the case that the polynomials are homogeneous.
\[ P_L(X) = \sum_{i \in [C], j \in [m]} \lambda_{i,j} P_i(L_j(X)) \]

Proof Ideas.

- Reduce to the case when \(|L_j| \leq \deg(P_i)|. e.g. Q(2x + z) = 2Q(x) + Q(z) - 2Q(x + z) - Q(2x).
- Reduce to the case that the polynomials are homogeneous.
- Applications of certain derivative operators \(D_i\) s.t.

\[
(\|\mathbb{E}_X e(P_L(X))\|)^{2^d} \leq \mathbb{E} \left[ e(\sum(D_1 \cdots D_d P_L)(X)) \right] = \|e(\sum_{i \in C} \lambda_i P_i)\|_{U_d}^{2^d}
\]
Technical Difficulties with $|F| \leq d$

- The inverse theorems for Gowers norm is no longer true with polynomials [Lovett-Meshulam-Samorodnitsky, Green-Tao]
The inverse theorems for Gowers norm is no longer true with polynomials [Lovett-Meshulam-Samorodnitsky, Green-Tao]

The inverse theorem holds with more complex “nonclassical polynomials” [Tao-Ziegler].

- e.g. $P(x_1, x_2) = \frac{x_1^2}{p^2} \mod 1$, $\deg(P) = p$. 

Much more complex behavior.

Cannot simply assume homogeneity.

[H.Hatami, P .H., Lovett]: Define a notion of homogeneity for nonclassical polynomials, $P(cx) = \lambda c P(x)$.

Show that every degree- $d$ polynomial can be written as linear combination of homogeneous polynomials.
Technical Difficulties with $|F| \leq d$

- The inverse theorems for Gowers norm is no longer true with polynomials [Lovett-Meshulam-Samorodnitsky, Green-Tao]
- The inverse theorem holds with more complex “nonclassical polynomials” [Tao-Ziegler].
  - e.g. $P(x_1, x_2) = \frac{x_1^2}{p^2} \mod 1$, $\deg(P) = p$.
  - Much more complex behavior.
  - Cannot simply assume homogeneity.
Technical Difficulties with $|\mathbb{F}| \leq d$

- The inverse theorems for Gowers norm is no longer true with polynomials [Lovett-Meshulam-Samorodnitsky, Green-Tao]
- The inverse theorem holds with more complex “nonclassical polynomials” [Tao-Ziegler].
  - e.g. $P(x_1, x_2) = \frac{x_1^2}{p^2} \mod 1, \deg(P) = p$.
  - Much more complex behavior.
  - Cannot simply assume homogeneity.

[H.Hatami, P.H., Lovett]:
- Define a notion of homogeneity for nonclassical polynomials, $P(cx) = \lambda_c P(x)$. 
Technical Difficulties with $|F| \leq d$

- The inverse theorems for Gowers norm is no longer true with polynomials [Lovett-Meshulam-Samorodnitsky, Green-Tao]
- The inverse theorem holds with more complex “nonclassical polynomials” [Tao-Ziegler].
  - e.g. $P(x_1, x_2) = \frac{x_1^2}{p^2} \mod 1$, $\deg(P) = p$.
  - Much more complex behavior.
  - Cannot simply assume homogeneity.

[H.Hatami, P.H., Lovett]:
- Define a notion of homogeneity for nonclassical polynomials, $P(cx) = \lambda_c P(x)$.
- Show that every degree-$d$ polynomial can be written as linear combination of homogeneous polynomials.
“Application”
Do Gowers norms control density of linear patterns

\[ \mathbb{E}_X f(L_1(X)) \cdots f(L_m(X))? \]
Do Gowers norms control density of linear patterns

\[ \mathbb{E}_X f(L_1(X)) \cdots f(L_m(X)) \]

Yes: seen in proofs of Szemerédi Theorem, Green-Tao Theorem on APs in Primes.
Do Gowers norms control density of linear patterns

$$\mathbb{E}_X f(L_1(X)) \cdots f(L_m(X))?$$

Yes: seen in proofs of Szemerédi Theorem, Green-Tao Theorem on APs in Primes.

[Green-Tao] Cauchy-Schwarz Complexity

$$|\mathbb{E} f(L_1(X)) \cdots f(L_m(X))| \leq \min_{i \in [m]} \| f \|_{U^{s+1}},$$

where $s$ is the Cauchy-Schwarz complexity of $\{L_1, \ldots, L_m\}$. 
Do Gowers norms control density of linear patterns

\[ \mathbb{E}_X f(L_1(X)) \cdots f(L_m(X)) \]

Yes: seen in proofs of Szemerédi Theorem, Green-Tao Theorem on APs in Primes.

[Green-Tao] Cauchy-Schwarz Complexity

\[ |\mathbb{E} f(L_1(X)) \cdots f(L_m(X))| \leq \min_{i \in [m]} \| f \|_{U^{s+1}}, \]

where \( s \) is the Cauchy-Schwarz complexity of \( \{ L_1, \ldots, L_m \} \).

Gowers-Wolf: There are cases where CS-Complexity \( s \) is not optimal.
Do Gowers norms control density of linear patterns

\[ \mathbb{E}_X f(L_1(X)) \cdots f(L_m(X)) \]?

Yes: seen in proofs of Szemerédi Theorem, Green-Tao Theorem on APs in Primes.

[Green-Tao] Cauchy-Schwarz Complexity

\[ |\mathbb{E} f(L_1(X)) \cdots f(L_m(X))| \leq \min_{i \in [m]} \| f \|_{U^{s+1}}, \]

where \( s \) is the Cauchy-Schwarz complexity of \( \{L_1, \ldots, L_m\} \).

Gowers-Wolf: There are cases where CS-Complexity \( s \) is not optimal.

\[ \| f \|_{U^{s'}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E} f(M_1(X)) \cdots f(M_\ell(X))| \leq \epsilon \quad \text{with } s' < s + 1. \]
Define the **true complexity** of $L_1, \ldots, L_m$ to be the smallest $d$ such that

$$\|f\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E}f(L_1(X)) \cdots f(L_m(X))| \leq \epsilon$$
True Complexity [Gowers-Wolf]

Define the true complexity of $L_1, \ldots, L_m$ to be the smallest $d$ such that

$$\|f\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E}f(L_1(X)) \cdots f(L_m(X))| \leq \epsilon$$

Theorem (CS – Complexity $<$ $|\mathbb{F}|$, [Gowers-Wolf])

A characterization of true complexity for sets of linear forms.
**True Complexity [Gowers-Wolf]**

Define the true complexity of $L_1, \ldots, L_m$ to be the smallest $d$ such that

$$\|f\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E} f(L_1(X)) \cdots f(L_m(X))| \leq \epsilon$$

| Theorem (CS – Complexity $< |\mathbb{F}|$, [Gowers-Wolf]) |
|----------------------------------------------------------|
| A characterization of true complexity for sets of linear forms. |

<table>
<thead>
<tr>
<th>Conjecture ([Gowers-Wolf])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Let $d$ be the smallest such that $L_1^{d+1}$ is not in $\text{span}(L_2^{d+1}, \ldots, L_m^{d+1})$, then</td>
</tr>
</tbody>
</table>

$$\|f_1\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E} f_1(L_1(X)) \cdots f_m(L_m(X))| \leq \epsilon$$
**True Complexity [Gowers-Wolf]**

Define the true complexity of $L_1, \ldots, L_m$ to be the smallest $d$ such that

$$\|f\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E} f(L_1(X)) \cdot \cdots \cdot f(L_m(X))| \leq \epsilon$$

**Theorem (CS – Complexity < |F|, [Gowers-Wolf])**

A characterization of true complexity for sets of linear forms.

**Conjecture ([Gowers-Wolf])**

Let $d$ be the smallest such that $L_1^{d+1}$ is not in $\text{span}(L_2^{d+1}, \ldots, L_m^{d+1})$, then

$$\|f_1\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E} f_1(L_1(X)) \cdots f_m(L_m(X))| \leq \epsilon$$

- [H. Hatami-P.H.-Lovett] Verify the conjecture in its full generality.
A simple telescoping (hybrid) argument leads to:

Corollary.

Assume that $L_1^{d+1}, \ldots, L_m^{d+1}$ are linearly independent. Then $\|f - g\|_{U^{d+1}} \leq \delta(\epsilon)$ implies

$$\left| \mathbb{E}_X \left[ \prod_{i=1}^{m} f(L_i(X)) \right] - \mathbb{E}_X \left[ \prod_{i=1}^{m} g(L_i(X)) \right] \right| \leq \epsilon$$
A simple telescoping (hybrid) argument leads to:

**Corollary.**

Assume that $L^d_1, \ldots, L^d_m$ are linearly independent. Then

$$\|f - g\|_{U^{d+1}} \leq \delta(\epsilon)$$

implies

$$\left| \mathbb{E}_X \left[ \prod_{i=1}^m f(L_i(X)) \right] - \mathbb{E}_X \left[ \prod_{i=1}^m g(L_i(X)) \right] \right| \leq \epsilon$$

$$\|\mathbf{1}_A - \mathbf{1}_B\|_{U^{d+1}} \leq \delta$$

implies that the number of $d$-APs in $A$ and $B$ are similar.
Theorem (H. Hatami-P.H.-Lovett)

Let $L_1, \ldots, L_m$ be such that $L_1^{d+1}$ is not in the span of $L_2^{d+1}, \ldots, L_m^{d+1}$. Then

$$\|f_1\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E} f_1(L_1(X)) \cdots f_m(L_m(X))| \leq \epsilon.$$
Theorem (H. Hatami-P.H.-Lovett)

Let $L_1, \ldots, L_m$ be such that $L_1^{d+1}$ is not in the span of $L_2^{d+1}, \ldots, L_m^{d+1}$. Then

$$\|f_1\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |\mathbb{E}f_1(L_1(X)) \cdots f_m(L_m(X))| \leq \epsilon.$$ 

Proof steps.

- We may assume that $d$ is less than the CS-Complexity.
Theorem (H. Hatami-P.H.-Lovett)

Let \( L_1, \ldots, L_m \) be such that \( L_1^{d+1} \) is not in the span of \( L_2^{d+1}, \ldots, L_m^{d+1} \). Then

\[
\|f_1\|_{U^{d+1}} \leq \delta(\epsilon) \Rightarrow |E f_1(L_1(X)) \cdots f_m(L_m(X))| \leq \epsilon.
\]

Proof steps.

- We may assume that \( d \) is less than the CS-Complexity.
- Write \( f_i = g_i + h_i \), where
  - \( g_i = \Gamma_i(P_1, \ldots, P_C) \),
  - \( P_1, \ldots, P_C \) is a regular (high-rank) set of degree \( \leq s \) polynomials.
  - \( \|h_i\|_{U^{s+1}} < \epsilon \).
True Complexity

Theorem (H. Hatami-P.H.-Lovett)

Let \( L_1, \ldots, L_m \) be such that \( L_1^{d+1} \) is not in the span of \( L_2^{d+1}, \ldots, L_m^{d+1} \). Then

\[
\|f_1\|_{U^{d+1}} \leq \delta(\varepsilon) \implies |\mathbb{E} f_1(L_1(X)) \cdots f_m(L_m(X))| \leq \varepsilon.
\]

Proof steps.

- We may assume that \( d \) is less than the CS-Complexity.
- Write \( f_i = g_i + h_i \), where
  - \( g_i = \Gamma_i(P_1, \ldots, P_C) \),
  - \( P_1, \ldots, P_C \) is a regular (high-rank) set of degree \( \leq s \) polynomials.
  - \( \|h_i\|_{U^{s+1}} < \varepsilon \).

\[
\mathbb{E} [(g_i + h_i)(L_1(X)) \cdots (g_m + h_m)(L_m(X))] \approx \mathbb{E} [g_1(L_1(X)) \cdots g_m(L_m(X))]
\]
True Complexity

\[\mathbb{E}[g_i(L_1(X)) \cdots g_m(L_m(X))]

Two cases based on \(\deg(P_{\Lambda_1})\):

(i) \(\deg(P_{\Lambda_1}) \leq d\): The coefficients will be small since \(\hat{\Gamma}_1(\Lambda_1)\) is small.

(ii) \(\deg(P_{\Lambda_1}) \geq d + 1\): The phase polynomials will be unbiased.
True Complexity

\[(\ast)\quad E \left[ g_i(L_1(X)) \cdots g_m(L_m(X)) \right] \]

\[g_i(x) = \Lambda_i(P_1(x), \ldots, P_C(x)) = \sum_{\Lambda=(\lambda_1, \ldots, \lambda_C)} \hat{\Gamma}_i(\Lambda) e\left(\sum_{P_\Lambda} \lambda_j P_j(x)\right)\]
True Complexity

\[(\ast) \quad \mathbb{E} [g_i(L_1(X)) \cdots g_m(L_m(X))]\]

\[g_i(x) = \Lambda_i(P_1(x), \ldots, P_C(x)) = \sum_{\Lambda=(\lambda_1, \ldots, \lambda_C)} \hat{\Gamma}_i(\Lambda)e\left(\sum_{P_\Lambda} \lambda_j P_j(x)\right)\]

\[(\ast) = \sum_{\Lambda_1, \ldots, \Lambda_m} \left(\prod_{i=1}^m \hat{\Gamma}_i(\Lambda_i)\right) \cdot \mathbb{E}_X \left[e\left(\sum_{i=1}^m P_{\Lambda_i}(L_i(X))\right)\right]\]
True Complexity

$$\mathbb{E} \left[ g_i(L_1(X)) \cdots g_m(L_m(X)) \right]$$

$$g_i(x) = \Lambda_i(P_1(x), \ldots, P_C(x)) = \sum_{\Lambda=(\lambda_1, \ldots, \lambda_C)} \hat{\Gamma}_i(\Lambda) e\left( \sum_{P_{\Lambda}} \lambda_j P_j(x) \right)$$

$$\left( \ast \right) = \sum_{\Lambda_1, \ldots, \Lambda_m} \left( \prod_{i=1}^m \hat{\Gamma}_i(\Lambda_i) \right) \cdot \mathbb{E}_x \left[ e \left( \sum_{i=1}^m P_{\Lambda_i}(L_i(X)) \right) \right]$$

Two cases based on $\text{deg}(P_{\Lambda_1})$

(i) $\text{deg}(P_{\Lambda_1}) \leq d$: 
True Complexity

\[(\ast) \quad \mathbb{E} [g_i(L_1(X)) \cdots g_m(L_m(X))]\]

\[g_i(x) = \Lambda_i(P_1(x), \ldots, P_C(x)) = \sum_{\Lambda=(\lambda_1, \ldots, \lambda_C)} \hat{\Gamma}_i(\Lambda) e\left(\sum_{P_\Lambda} \lambda_j P_j(x)\right)\]

\[(\ast) = \sum_{\Lambda_1, \ldots, \Lambda_m} \left(\prod_{i=1}^m \hat{\Gamma}_i(\Lambda_i)\right) \cdot \mathbb{E}_X \left[ e \left(\sum_{i=1}^m P_{\Lambda_i}(L_i(X))\right)\right]\]

Two cases based on \(\deg(P_{\Lambda_1})\)

(i) \(\deg(P_{\Lambda_1}) \leq d\): The coefficients will be small since \(\hat{\Gamma}_1(\Lambda_1)\) is small.
True Complexity

\[ g_i(x) = \Lambda_i(P_1(x), \ldots, P_C(x)) = \sum_{\Lambda=(\lambda_1,\ldots,\lambda_C)} \hat{\Gamma}_i(\Lambda)e\left(\sum_{P_\Lambda} \lambda_j P_j(x)\right) \]

\[ (\ast) \quad \mathbb{E} [g_i(L_1(X)) \cdots g_m(L_m(X))] \]

\[ (\ast) = \sum_{\Lambda_1,\ldots,\Lambda_m} \left( \prod_{i=1}^m \hat{\Gamma}_i(\Lambda_i) \right) \cdot \mathbb{E}_x \left[ e \left( \sum_{i=1}^m P_{\Lambda_i}(L_i(X)) \right) \right] \]

Two cases based on \( \text{deg}(P_{\Lambda_1}) \)

(i) \( \text{deg}(P_{\Lambda_1}) \leq d \): The coefficients will be small since \( \hat{\Gamma}_1(\Lambda_1) \) is small.

(ii) \( \text{deg}(P_{\Lambda_1}) \geq d + 1 \):
True Complexity

\[(\ast) \quad \mathbb{E} \left[ g_i(L_1(X)) \cdots g_m(L_m(X)) \right] \]

\[g_i(x) = \Lambda_i(P_1(x), \ldots, P_C(x)) = \sum_{\Lambda=(\lambda_1, \ldots, \lambda_C)} \hat{\Gamma}_i(\Lambda)e\left(\sum_{P_\Lambda} \lambda_j P_j(x)\right)\]

\[(\ast) = \sum_{\Lambda_1, \ldots, \Lambda_m} \left( \prod_{i=1}^m \hat{\Gamma}_i(\Lambda_i) \right) \cdot \mathbb{E}_X \left[ e \left( \sum_{i=1}^m P_{\Lambda_i}(L_i(X)) \right) \right] \]

Two cases based on $\text{deg}(P_{\Lambda_1})$

(i) $\text{deg}(P_{\Lambda_1}) \leq d$: The coefficients will be small since $\hat{\Gamma}_1(\Lambda_1)$ is small.

(ii) $\text{deg}(P_{\Lambda_1}) \geq d + 1$: The phase polynomials will be unbiased.
True Complexity

\[ (*) \quad \mathbb{E} [g_i(L_1(X)) \cdots g_m(L_m(X))] \leq \epsilon \]

\[ g_i(x) = \Lambda_i(P_1(x), \ldots, P_C(x)) = \sum_{\Lambda=(\lambda_1, \ldots, \lambda_C)} \hat{\Gamma}_i(\Lambda) e\left( \sum_{P_{\Lambda}} \lambda_j P_j(x) \right) \]

\[ (*) = \sum_{\Lambda_1, \ldots, \Lambda_m} \left( \prod_{i=1}^m \hat{\Gamma}_i(\Lambda_i) \right) \cdot \mathbb{E}_X \left[ e \left( \sum_{i=1}^m P_{\Lambda_i}(L_i(X)) \right) \right] \]

**Two cases based on \( \text{deg}(P_{\Lambda_1}) \)**

(i) \( \text{deg}(P_{\Lambda_1}) \leq d \): The coefficients will be small since \( \hat{\Gamma}_1(\Lambda_1) \) is small.

(ii) \( \text{deg}(P_{\Lambda_1}) \geq d + 1 \): The phase polynomials will be unbiased. \( \square \)
Thanks!