Higher-order Fourier analysis and an application

FSTTCS ‘15 Workshop

Arnab Bhattacharyya
Indian Institute of Science

December 19, 2015

Roadmap

- Preliminaries and review of Fourier analysis
- What is “higher-order” Fourier analysis?
- An application to locally correctable codes

No historical account!

Some Preliminaries

Setting

$\mathbb{F}$ = finite field of **fixed prime** order

- For example, $\mathbb{F} = \mathbb{F}_2$ or $\mathbb{F} = \mathbb{F}_{97}$
- Theory can be extended to extensions of prime fields **[B.-Bhowmick '15]**

Functions

Functions are always multivariate,
on n variables

$$f\colon \mathbb{F}^n \to \mathbb{C} \qquad (|f| \leq 1)$$

and $$P\colon \mathbb{F}^n \to \mathbb{F}$$

Current bounds aim to be efficient wrt *n*

Polynomial

Polynomial of degree d is of the form:

$$\sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}$$

where each $c_{i_1,\ldots,i_n} \in \mathbb{F}$ and $i_1 + \cdots + i_n \leq d$

Phase Polynomial

Phase polynomial of degree d is a function $f\colon \mathbb{F}^n \to \mathbb{C}$ of the form $f(x) = \mathrm{e}(P(x))$ where:

1. $P\colon \mathbb{F}^n \to \mathbb{F}$ is a polynomial of degree d
2. $\mathrm{e}(k) = e^{2\pi i k/|\mathbb{F}|}$

Inner Product

The **inner product** of two functions $f, g\colon \mathbb{F}^n \to \mathbb{C}$ is:

$$\langle f, g \rangle = \mathbb{E}_{x \in \mathbb{F}^n}\left[f(x) \cdot \overline{g(x)}\right]$$

Magnitude captures correlation between f and g

Derivatives

Additive derivative in direction $h \in \mathbb{F}^n$ of function $P\colon\ \mathbb{F}^n \to \mathbb{F}$ is:

$$D_h P(x) = P(x+h) - P(x)$$

Derivatives

Multiplicative derivative in direction $h \in \mathbb{F}^n$ of function $f\colon \mathbb{F}^n \to \mathbb{C}$ is:

$$\Delta_h f(x) = f(x+h) \cdot \overline{f(x)}$$

Polynomial Factor

Factor of degree d and order m is a tuple of polynomials $\mathcal{B} = (P_1, P_2, \ldots, P_m)$, each of degree d.

As shorthand, write:

$$\mathcal{B}(x) = (P_1(x), \ldots, P_m(x))$$

Fourier Analysis over $\mathbb{F}$

Fourier Representation

Every function $f\colon \mathbb{F}^n \to \mathbb{C}$ is a linear combination of **linear** phases:

$$f(x) = \sum_{\alpha \in \mathbb{F}^n} \hat{f}(\alpha)\, \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

Linear Phases

- The inner product of two linear phases is:

$$\langle \mathrm{e}\left(\sum_i \alpha_i x_i\right), e\left(\sum_i \beta_i x_i\right)\rangle = \mathbb{E}_x\left[\mathrm{e}\left(\sum_i (\alpha_i - \beta_i) x_i\right)\right] = 0$$

if $\alpha \neq \beta$ and is 1 otherwise.

- So:

$\hat{f}(\alpha) = \langle f, \mathrm{e}(\sum_i \alpha_i x_i)\rangle$= correlation with linear phase

Random functions

With high probability, a **random** function $f\colon \mathbb{F}^n \to \mathbb{C}$ with $|f| = 1$ has each $\hat{f}(\alpha) \to 0$.

Decomposition Theorem

$$f(x) = g(x) + h(x)$$

where:

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

Decomposition Theorem

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

Every Fourier coefficient of h is less than ϵ, so h is "pseudorandom".

Decomposition Theorem

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

g has only $1/\epsilon^2$ nonzero Fourier coefficients

Decomposition Theorem

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

The nonzero Fourier coefficients of g can be found in poly time **[Goldreich-Levin ‘89]**

Elements of Higher-Order Fourier Analysis

Higher-order Fourier analysis is the interplay between three different notions of pseudorandomness for functions and factors.

1. Bias
2. Gowers norm
3. Rank

Bias

Bias

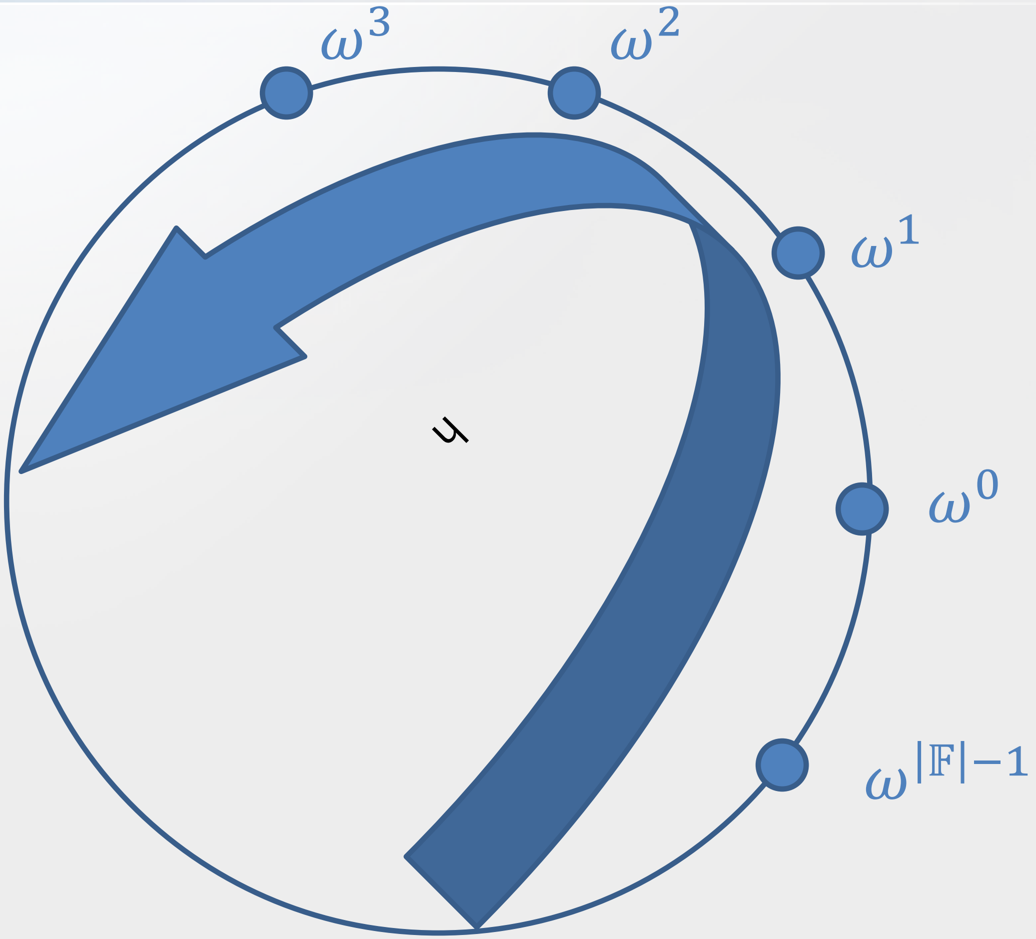
For $f\colon \mathbb{F}^n \to \mathbb{C}$,

$$\mathrm{bias}(f) = |\mathbb{E}_x[f(x)]|$$

For $P\colon \mathbb{F}^n \to \mathbb{F}$,

$$\mathrm{bias}(P) = |\mathbb{E}_x\left[\mathrm{e}\left(P(x)\right)\right]|$$

[…, Naor-Naor '89, …]



How well is P equidistributed?

Bias of Factor

A factor $\mathcal{B} = (P_1, \ldots, P_k)$ is **$\boldsymbol{\alpha}$-unbiased** if every nonzero linear combination of $P_1, \ldots, P_k$ has bias less than α:

$$\text{bias}\left(\sum_{i=1}^{k} c_i P_i\right) < \alpha$$
$$\forall (c_1, \ldots, c_k) \in \mathbb{F}^k \setminus \{0\}$$

Bias implies equidistribution

Lemma: If $\mathcal{B}$ is α-unbiased and of order k, then for any $c \in \mathbb{F}^k$:

$$\Pr[\mathcal{B}(x) = c] = \frac{1}{|\mathbb{F}|^k} \pm \alpha$$

Bias implies equidistribution

Lemma: If $\mathcal{B}$ is α-unbiased and of order k, then for any $c \in \mathbb{F}^k$:

$$\Pr[\mathcal{B}(x) = c] = \frac{1}{|\mathbb{F}|^k} \pm \alpha$$

Corollary: If $\mathcal{B}$ is α-unbiased and $\alpha < \frac{1}{|\mathbb{F}|^k}$, then $\mathcal{B}$ maps onto $\mathbb{F}^k$.

Gowers Norm

Gowers Norm

Given $f\colon \mathbb{F}^n \rightarrow \mathbb{C}$, its **Gowers norm of order d** is:

$$U^d(f) = |\mathbb{E}_{x,h_1,h_2,\ldots,h_d} \Delta_{h_1} \Delta_{h_2} \cdots \Delta_{h_d} f(x)|^{1/2^d}$$

[Gowers '01]

Gowers Norm

Given $f\colon \mathbb{F}^n \to \mathbb{C}$, its **Gowers norm of order d** is:

$$U^d(f) = |\mathbb{E}_{x,h_1,h_2,\ldots,h_d} \Delta_{h_1}\Delta_{h_2}\cdots\Delta_{h_d} f(x)|^{1/2^d}$$

Observation: If $f = \mathrm{e}(P)$ is a phase poly, then:

$$U^d(f) = |\mathbb{E}_{x,h_1,h_2,\ldots,h_d} \mathrm{e}\big(D_{h_1}D_{h_2}\cdots D_{h_d}P(x)\big)|^{1/2^d}$$

Gowers norm for phase polys

- If f is a phase poly of degree d, then:

$$U^{d+1}(f) = 1$$

- Converse is true when $d < |\mathbb{F}|$.

Other Observations

- $U^1(f) = \sqrt{|\mathbb{E}[f]|^2} = \text{bias}(f)$

- $U^2(f) = \sqrt[4]{\sum_\alpha \hat{f}^4(\alpha)}$

- $U^1(f) \leq U^2(f) \leq U^3(f) \leq \cdots$ (C.-S.)

Pseudorandomness

- For random $f\colon \mathbb{F}^n \to \mathbb{C}$ and fixed d,
$$U^d(f) \to 0$$
- By monotonicity, low Gowers norm implies low bias and low Fourier coefficients.

Correlation with Polynomials

Lemma: $U^{d+1}(f) \geq \max |\langle f, \mathrm{e}(P) \rangle|$ where max is over all polynomials P of degree d.

Proof: For any poly P of degree d:

$$\begin{aligned} \left|\mathbb{E}\left[f(x) \cdot \mathrm{e}\big(-P(x)\big)\right]\right| &= U^{1}\big(f \cdot \mathrm{e}(-P)\big) \\ &\leq U^{d+1}\big(f \cdot \mathrm{e}(-P)\big) \\ &= U^{d+1}(f) \end{aligned}$$

Gowers Inverse Theorem

Theorem: If $d < |\mathbb{F}|$, for all $\epsilon > 0$, there exists $\delta = \delta(\epsilon, d, \mathbb{F})$ such that if $U^{d+1}(f) > \epsilon$, then $|\langle f, \mathrm{e}(P)\rangle| > \delta$ for some poly P of degree d.

Proof:

- **[Green-Tao '09]** Combinatorial for phase poly f (c.f. Madhur's talk later).
- **[Bergelson-Tao-Ziegler '10]** Ergodic theoretic proof for arbitrary f.

Small Fields

Consider $f\colon \mathbb{F}_2^1 \to \mathbb{C}$ with:

$$f(0) = 1$$
$$f(1) = i$$

f not a phase poly but $U^3(f) = 1$!

Small fields: worse news

Consider $f = \mathrm{e}(P)$ where $P: \mathbb{F}_2^n \to \mathbb{F}_2$ is symmetric polynomial of degree 4.

$$U^4(f) = \Omega(1)$$

but:

$$|\langle f, \mathrm{e}(C) \rangle| = \exp(-n)$$

for all cubic poly C.

[Lovett-Meshulam-Samorodnitsky ’08, Green-Tao ‘09]

Nevertheless…

Just ***define*** **non-classical phase polynomials of degree d** to be functions $f\colon \mathbb{F}^n \to \mathbb{C}$ such that $|f| = 1$ and

$$\Delta_{h_1}\Delta_{h_2}\cdots\Delta_{h_{d+1}} f(x) = 1$$

for all $x, h_1, \ldots, h_{d+1} \in \mathbb{F}^n$

Inverse Theorem for small fields

Theorem: For all $\epsilon > 0$, there exists $\delta = \delta(\epsilon, d, \mathbb{F})$ such that if $U^{d+1}(f) > \epsilon$, then $|\langle f, g \rangle| > \delta$ for some non-classical phase poly g of degree d.

Proof:
- [Tao-Ziegler] Combinatorial for phase poly f .
- [Tao-Ziegler] Nonstandard proof for arbitrary f.

Pseudorandomness & Counting

Theorem: If $L_1, \ldots, L_m$ are m linear forms $(L_j(X_1, \ldots, X_k) = \sum_{i=1}^{k} \ell_{i,j} X_i)$, then:

$$\mathbb{E}_{X_1,\ldots,X_k \in \mathbb{F}^n}\left[\prod_{j=1}^{m} f(L_j(X_1, \ldots, X_k)\right] \leq U^t(f)$$

if $f\colon\ \mathbb{F}^n \to \mathbb{C}$ and t is the *complexity* of the linear forms $L_1, \ldots, L_m$.

[Gowers-Wolf '10]

Examples

- Given $f\colon \mathbb{F}^n \to \mathbb{R}$ and we want to "count" the number of 3-term AP's:

$$\mathbb{E}_{X,Y}[f(X) \cdot f(X+Y) \cdot f(X+2Y)] \leq \sum_{\alpha} \hat{f}^3(\alpha)$$

- Similarly, number of 4-term AP's controlled by 3rd order Gowers norm of f.

Rank

Rank

Given a polynomial $P\colon \mathbb{F}^n \to \mathbb{F}$ of degree d, its **rank** is the smallest integer r such that:

$$P(x) = \Gamma(Q_1(x), \dots, Q_r(x)) \quad \forall x \in \mathbb{F}^n$$

where $Q_1, \dots, Q_r$ are polys of degree $d-1$ and $\Gamma\colon \mathbb{F}^r \to \mathbb{F}$ is arbitrary.

Pseudorandomness

- For random poly P of fixed degree d,
$$\mathrm{rank}(P) = \omega(1)$$

- High rank is pseudorandom behavior

Rank & Gowers Norm

If $P\colon\ \mathbb{F}^n \to \mathbb{F}$ is a poly of degree d, P has high rank **if and only if** $\mathsf{e}(P)$ has low Gowers norm of order d!

Low rank implies large Gowers norm

Lemma: If $P(x) = \Gamma(Q_1(x), \ldots, Q_k(x))$ where $Q_1, \ldots, Q_k$ are polys of deg $d-1$, then $U^d\big(\mathrm{e}(P)\big) \geq \frac{1}{|\mathbb{F}|^{k/2}}$.

Low rank implies large Gowers norm

Lemma: If $P(x) = \Gamma(Q_1(x), \dots, Q_k(x))$ where $Q_1, \dots, Q_k$ are polys of deg $d-1$, then $U^d\big(\mathrm{e}(P)\big) \geq \frac{1}{|\mathbb{F}|^{k/2}}$.

Proof: By (linear) Fourier analysis:

$$\mathrm{e}\big(P(x)\big) = \sum_{\alpha} \hat{\Gamma}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i \cdot Q_i\,(x)\right)$$

Therefore:

$$|\mathbb{E}_x \sum_{\alpha} \hat{\Gamma}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i \cdot Q_i(x) - P(x)\right)| = 1$$

Then, there's an α such that

$$\langle \mathrm{e}(P), \mathrm{e}(\textstyle\sum_i \alpha_i Q_i) \rangle \geq |\mathbb{F}|^{-k/2}.$$

Inverse theorem for polys

Theorem: For all ϵ and d, there exists $R = R(\epsilon, d, \mathbb{F})$ such that if P is a poly of degree d and $U^d\big(\mathrm{e}(P)\big) > \epsilon$, then rank($P$)$< R$.

[Tao-Ziegler '11]

Bias-rank theorem

Theorem: For all ϵ and d, there exists $R = R(\epsilon, d, \mathbb{F})$ such that if P is a poly of degree d and bias$(P) > \epsilon$, then rank(P)< R.

[Green-Tao ‘09, Kaufman-Lovett ‘08]

Decomposition Theorem

For any $\epsilon > 0$ and integer $r > 1$, there is a k so that any bounded $f: \mathbb{F}^n \to \mathbb{C}$ has a decomposition:

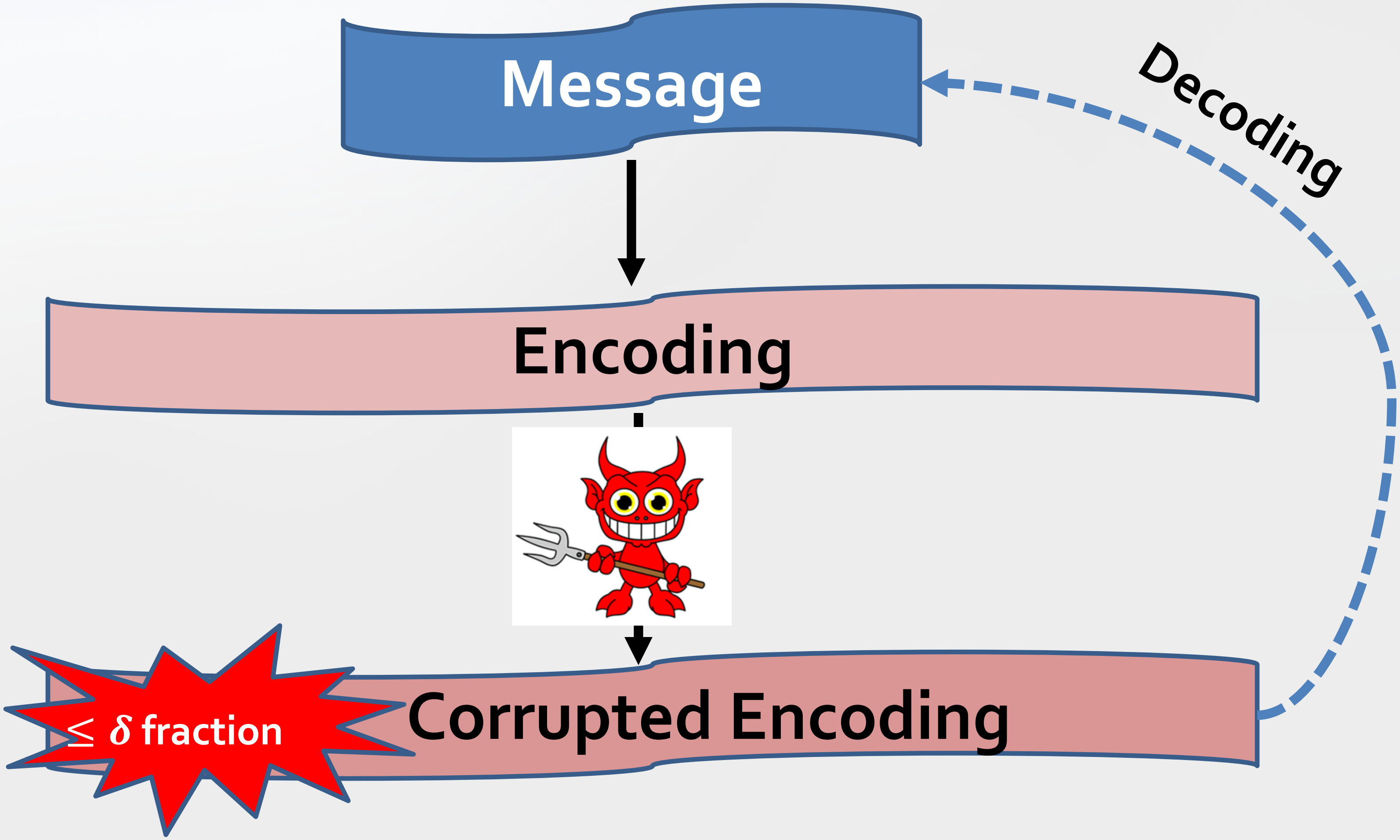
$$f = g + h$$

where $g = \Gamma(P_1, \ldots, P_k)$ for degree $< r$ non-classical polynomials $P_1, \ldots, P_k$ and $U^r(h) < \epsilon$.

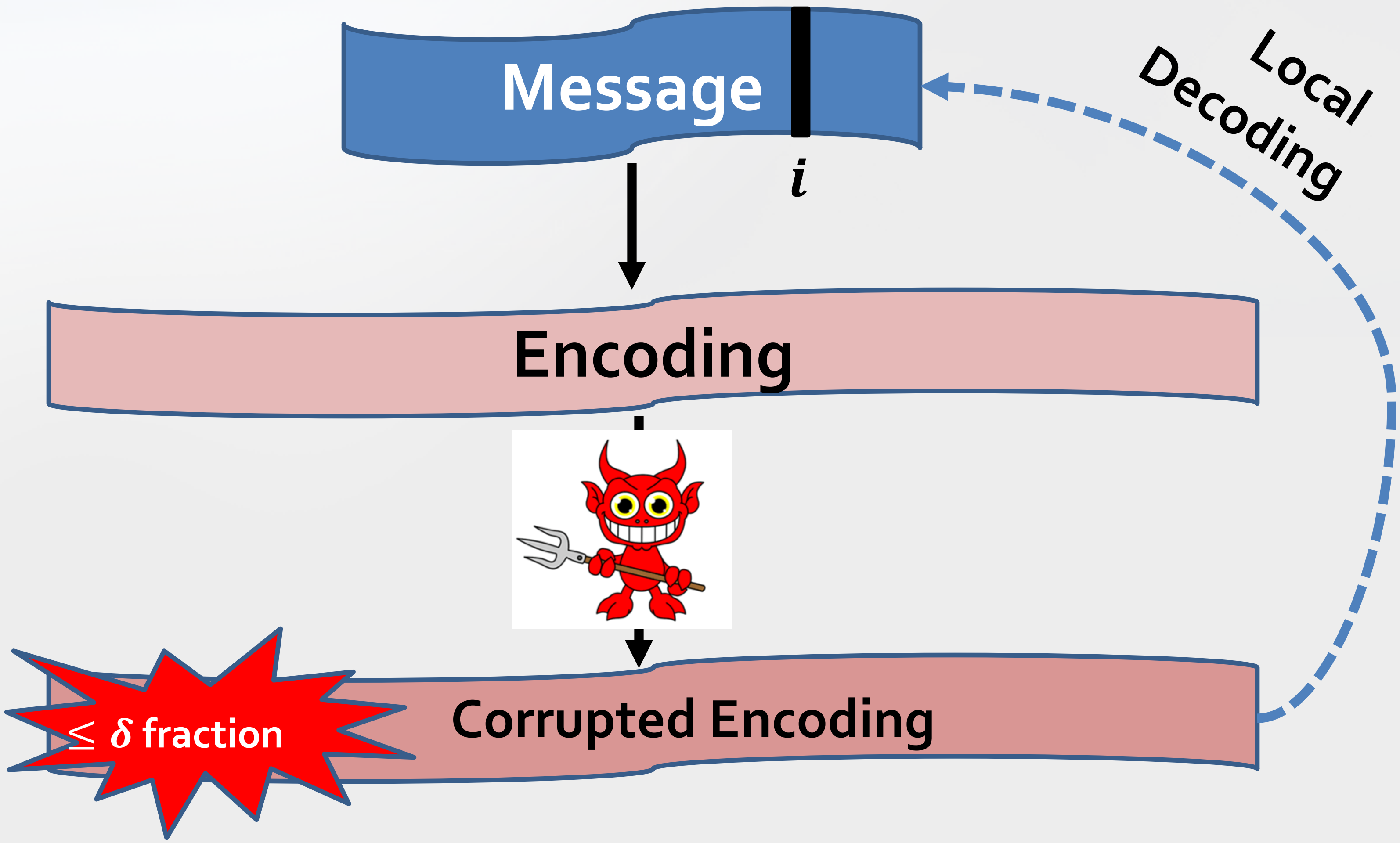
An Application: Locally Correctable Codes

[B.-Gopi '15]

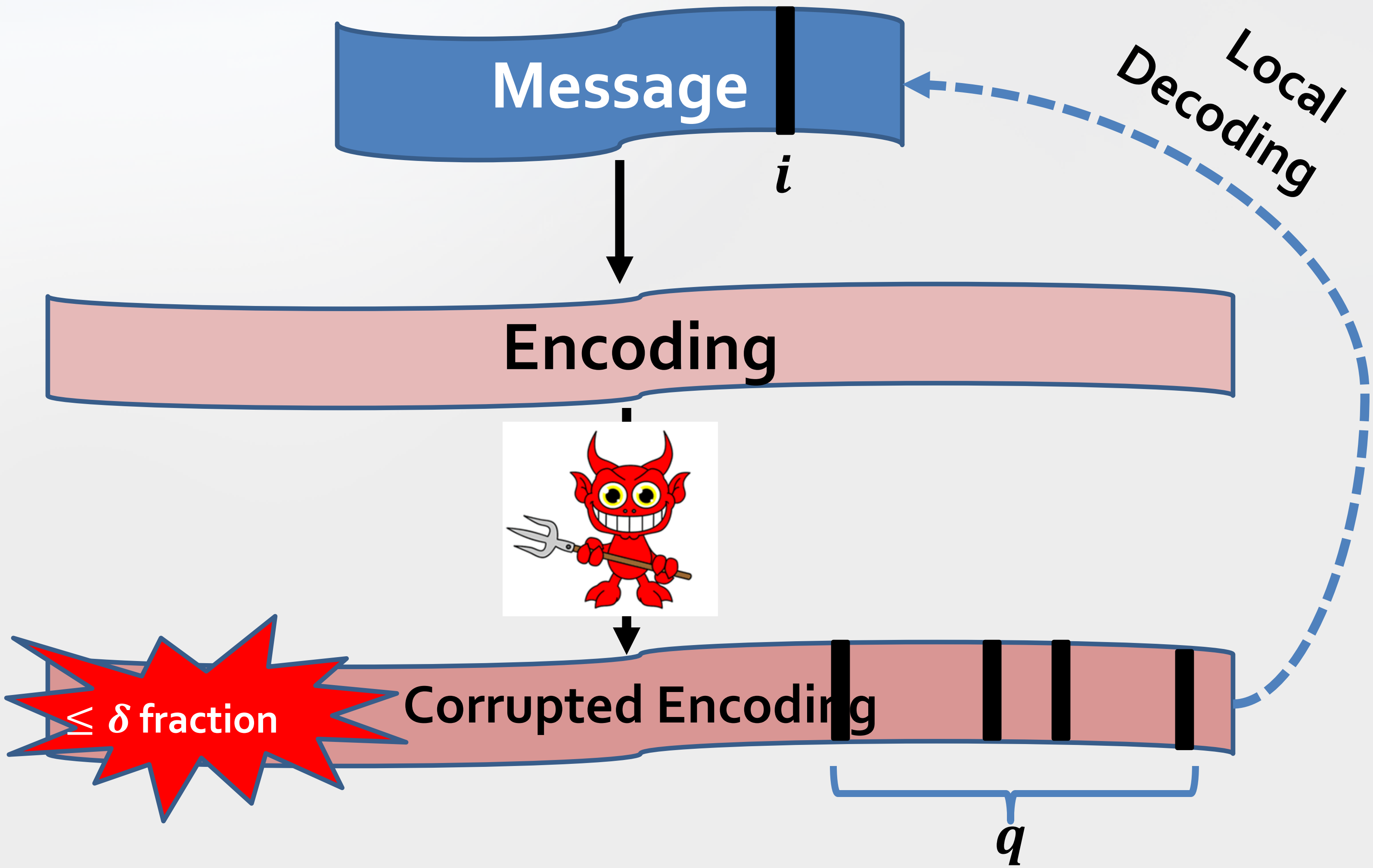
Tackling Adversarial Errors



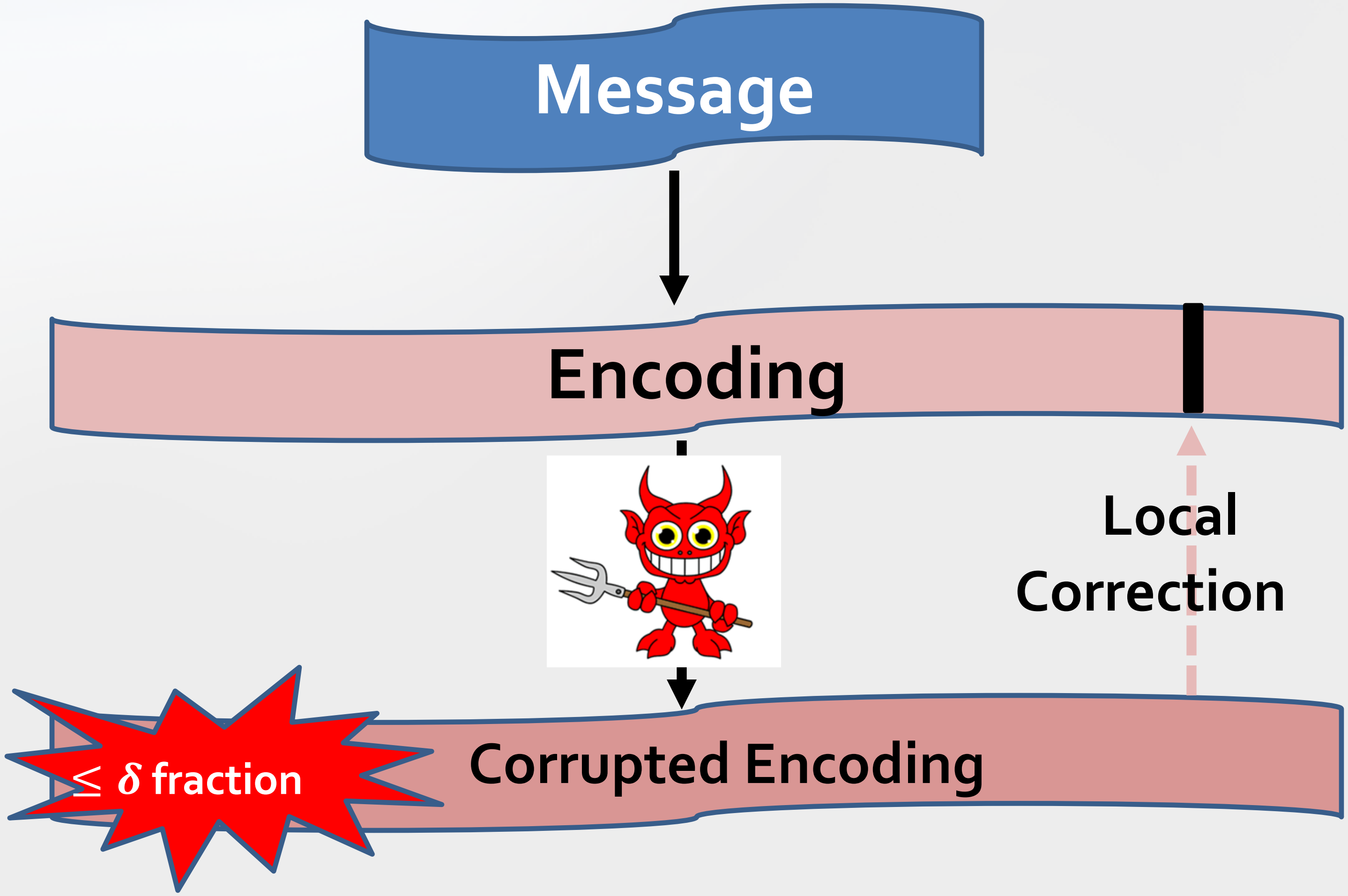
Locally Decodable Codes



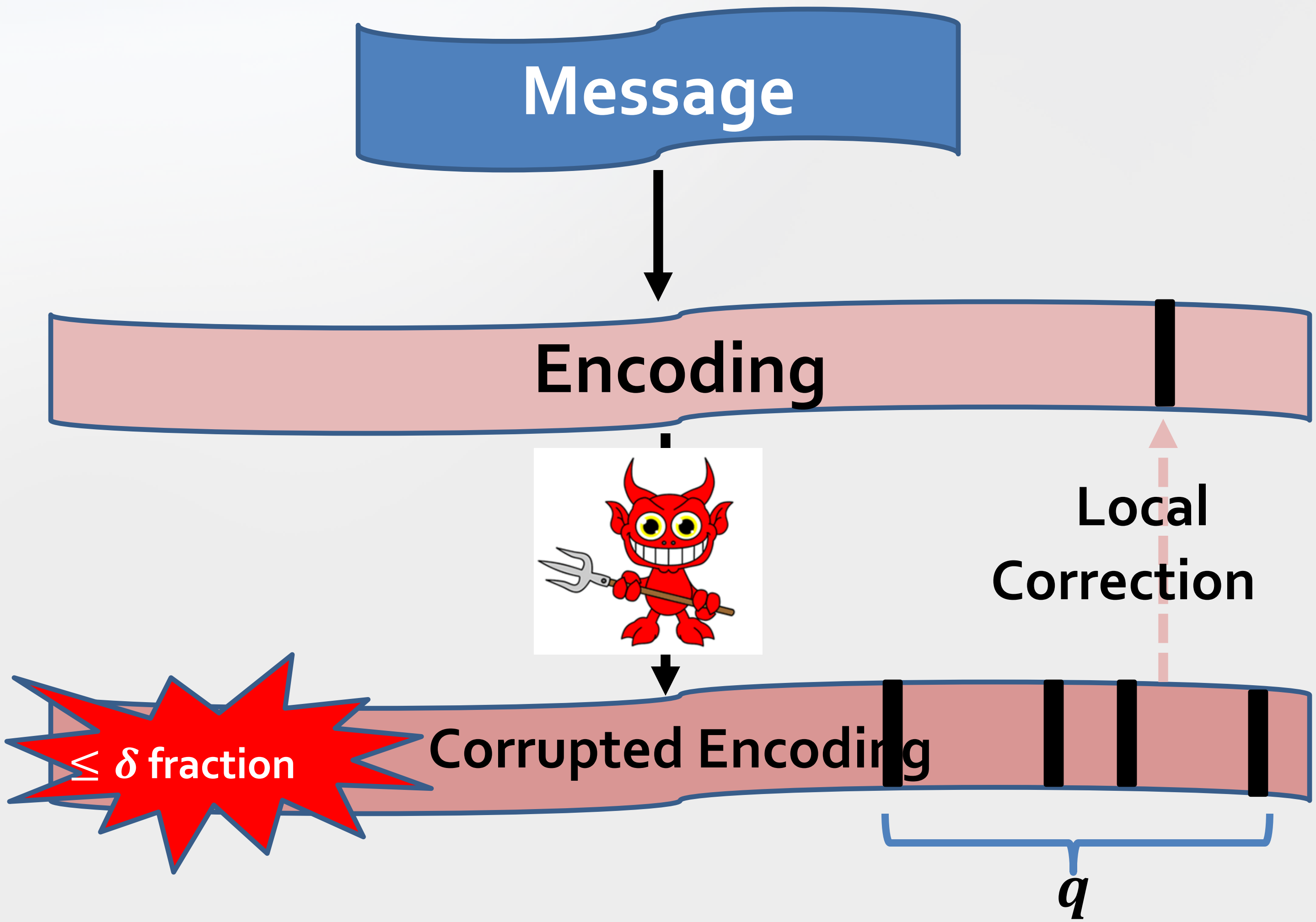
(q, δ)-Locally Decodable Codes



Locally Correctable Codes



(q, δ)-Locally Correctable Code



Locally Correctable Code (LCC)

Any encoding $y \in \Sigma^n$ by a $(\boldsymbol{q}, \boldsymbol{\delta})$**-LCC** has the property that for every δ-corruption y' of y and for every $i \in [n]$, with probability at least 90%, one can recover $y[i]$ by looking at q symbols in y'.

LDC/LCC Applications

- Private Information Retrieval (PIR) schemes
- Secure Multiparty Computation
- Complexity theoretic applications:
  - Arithmetic circuit lower bounds, Average-case complexity, Derandomization

LCC example

- Hadamard code $H \subseteq \{0,1\}^{2^n}$
- Interpret n-bit message $(a_1, \ldots, a_n)$ as linear form $H(x) = \sum_i a_i x_i$ and write evaluations of H on all $\{0,1\}^n$
- To recover $H(x)$, choose random y and output $H(x+y) - H(y)$

Current Status: Construction

- If q is a constant, current shortest LCC is Reed-Muller code of order $q-1$ (evaluation table of a polynomial of degree $q-1$ on a field of size $> q$)
  - To recover $P(x)$, pass line ℓ in random direction thru x, evaluate on q points on line to interpolate P_ℓ and evaluate $P_\ell(x)$

Current Status: Construction

- Same length also achieved by the "lifted codes" of **[Guo-Kopparty-Sudan '13]**.

Current Status: Lower bounds

- Hadamard code known to be optimal for 2 queries (for constant alphabet)

- For larger number of queries, only very weak bounds known

Our Result

- Reed-Muller (and **[GKS '13]**) optimal q-query LCC among **affine-invariant** codes

Affine invariance

- For a codeword $w \in \Sigma^{\mathbb{F}^n}$, we can view it as a function $w: \mathbb{F}^n \to \Sigma$

- Code C is **affine-invariant** if for any $w \in C$, $w \circ A \in C$ for any affine transformation $A: \mathbb{F}^n \to \mathbb{F}^n$.

Why affine invariance?

- “Generic” way to introduce many constraints among codeword positions.
- Affine-invariance natural for algebraically defined error-correcting codes
- Study of connection between correctability and invariances formally initiated by **[Kaufman-Sudan ‘08].**

Previous work

- **[Ben Sasson-Sudan ‘11]** showed that Reed-Muller is optimal among all **linear**, affine-invariant codes.
  - Their result does not assume fixed field size as ours does

Key Lemma

The metric induced by the $|| \cdot ||_{U^q}$-norm on the space of all bounded functions has an ϵ-net of size $\exp(O(n^{q-1}))$.

Proof of Key Lemma

- Net consists of all functions of the form $\Gamma(P_1, \ldots, P_k)$ where $P_1, \ldots, P_k$ are degree $< q$, non-classical polynomials, k is a constant, and Γ arbitrary.
- By decomposition theorem, such a function approximates given f!
- Can discretize Γ without affecting error too much.

Proof of Theorem

- Take two codewords f and g.

- If decoder runs on $f \circ A$ for random position y and any affine map A, it must with good prob give different answer than $g(A(y))$.

- On the other hand, if f and g close in U^q norm, then for any y and queried positions $y_1, \ldots, y_q$,

$$\mathbb{E}[\langle f \circ A(y) - g \circ A(y), \mathcal{D}\left(f \circ A(y_1), \ldots, f \circ A\left(y_q\right)\right)]$$

is small over random A.

Uses counting lemma

- Contradiction, so f and g lie in different cells of ϵ-net for U^q-norm.

More applications

- List-decoding radius for Reed-Muller codes [Bhowmick-Lovett ‘14, ‘15]

- New algorithms for factoring and decomposing polynomials [B. ‘14]

- New testers for algebraic properties [B.-Fischer-Hatami-Hatami-Lovett ‘13]

- …?

Thanks!