Distance Preservation for All Polynomial Generators

Sarah Bordage and Alessandro Chiesa
EPFL

Lattices Meet Hashes
May 2, 2023

Distance Preservation for All Polynomial Generators

Sarah Bordage and Alessandro Chiesa
EPFL

Lattices Meet Hashes
May 2, 2023

Linear codes

A linear **code** $\mathscr{C}$ is a linear subspace of $\mathbb{F}^n$.

Linear codes

A linear **code** $\mathscr{C}$ is a linear subspace of $\mathbb{F}^n$.

A code $\mathscr{C}$ has relative minimum **distance** $\delta_{\text{min}} \in [0, 1]$ if

$$\forall c, c' \in \mathscr{C}, c \neq c' : \Delta(c, c') \geq \delta_{\text{min}}.$$

$\Delta(\cdot, \cdot) =$ relative Hamming distance

Linear codes

A linear **code** $\mathscr{C}$ is a linear subspace of $\mathbb{F}^n$.

A code $\mathscr{C}$ has relative minimum **distance** $\delta_{\min} \in [0,1]$ if

$$\forall c, c' \in \mathscr{C}, c \neq c' : \Delta(c, c') \geq \delta_{\min}.$$

$\Delta(\cdot,\cdot) =$ relative Hamming distance

A vector $u \in \mathbb{F}^n$ is **δ-close to** $\mathscr{C}$ if

$$\min_{c \in \mathscr{C}} \Delta(u, c) = \Delta(u, \mathscr{C}) < \delta.$$

A linear **code** $\mathscr{C}$ is a linear subspace of $\mathbb{F}^n$.

A code $\mathscr{C}$ has relative minimum **distance** $\delta_{\text{min}} \in [0, 1]$ if

$$\forall c, c' \in \mathscr{C}, c \neq c' : \Delta(c, c') \geq \delta_{\text{min}}.$$

$\Delta(\cdot, \cdot) =$ relative Hamming distance

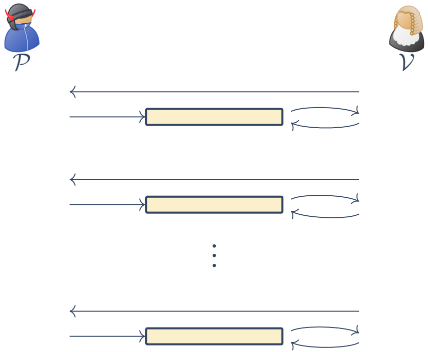
A vector $u \in \mathbb{F}^n$ is **δ-close to $\mathscr{C}$** if

$$\min_{c \in \mathscr{C}} \Delta(u, c) = \Delta(u, \mathscr{C}) < \delta.$$

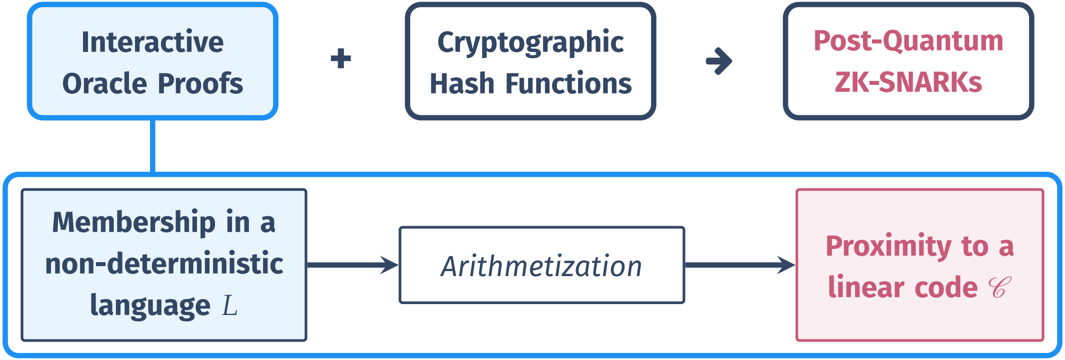
Otherwise, u is **δ-far from $\mathscr{C}$**.

Probabilistic proofs and proximity testing to codes

Interactive Oracle Proofs + **Cryptographic Hash Functions** → **Post-Quantum ZK-SNARKs**

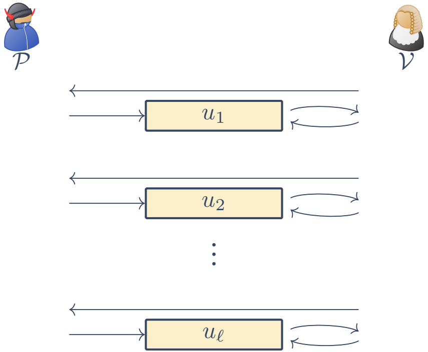


Probabilistic proofs and proximity testing to codes



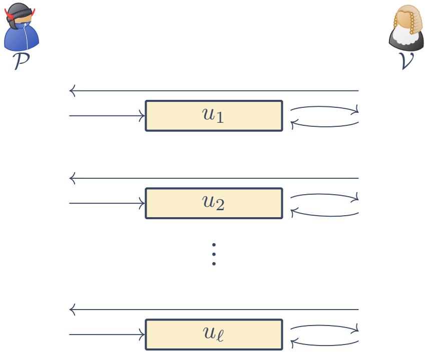
Batch Proximity Testing in Interactive Oracle Proofs

- If x $\in L$, then $\exists u_1, \ldots, u_\ell \in \mathscr{C}$ satisfying all verifier's checks.
- If x $\notin L$, then any $(u_1, \ldots, u_\ell) \in (\mathbb{F}^n)^\ell$ falsifies verifier's checks with high probability, **given that the u_i's are all close to $\mathscr{C}$**.



Batch Proximity Testing in Interactive Oracle Proofs

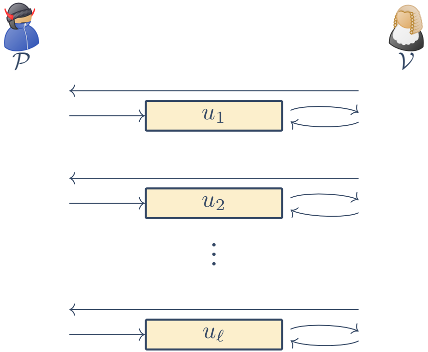
- If $\mathsf{x} \in L$, then $\exists u_1, \ldots, u_\ell \in \mathscr{C}$ satisfying all verifier's checks.
- If $\mathsf{x} \notin L$, then any $(u_1, \ldots, u_\ell) \in (\mathbb{F}^n)^\ell$ falsifies verifier's checks with high probability, **given that the u_i's are all close to $\mathscr{C}$**.



Needed: check proximity of $u_1, \ldots, u_\ell$ to $\mathscr{C}$.

Batch Proximity Testing in Interactive Oracle Proofs

- If $\mathbb{x} \in L$, then $\exists u_1, \ldots, u_\ell \in \mathscr{C}$ satisfying all verifier's checks.
- If $\mathbb{x} \notin L$, then any $(u_1, \ldots, u_\ell) \in (\mathbb{F}^n)^\ell$ falsifies verifier's checks with high probability, **given that the u_i's are all close to $\mathscr{C}$**.



Needed: check proximity of $u_1, \ldots, u_\ell$ to $\mathscr{C}$.

Proximity tests can be **expensive**, e.g. FRI protocol used in STARKs, Aurora, Ligero, Shockwave, ...

Testing Proximity to Linear Codes

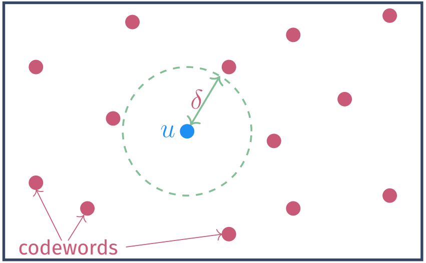
Proximity test for a single vector

Proximity test $(\mathcal{P}, \mathcal{V})$

Given:
- linear code $\mathscr{C} \subseteq \mathbb{F}^n$
- proximity parameter δ
- purported codeword $u \in \mathbb{F}^n$

$\mathcal{P}$'s inputs: $\mathscr{C}, \delta, u$.

$\mathcal{V}$'s inputs: $\mathscr{C}, \delta$ and oracle access to u.



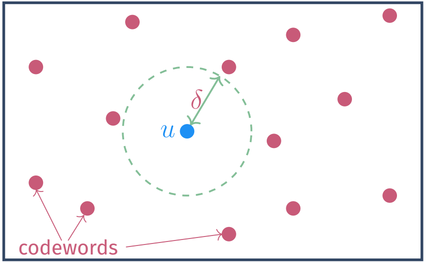
Proximity test for a single vector

Proximity test $(\mathcal{P}, \mathcal{V})$

Given:
- linear code $\mathscr{C} \subseteq \mathbb{F}^n$
- proximity parameter δ
- purported codeword $u \in \mathbb{F}^n$

$\mathcal{P}$'s inputs: $\mathscr{C}, \delta, u$.

$\mathcal{V}$'s inputs: $\mathscr{C}, \delta$ and oracle access to u.



Completeness. If $u \in \mathscr{C}$, verifier $\mathcal{V}$ accepts.

Soundness. If $\Delta(u, \mathscr{C}) \geq \delta$, verifier $\mathcal{V}$ rejects with high prob.

Batch proximity test using random linear combinations [RVW13]

Batch proximity test $(\mathcal{P}_{\text{batch}}, \mathcal{V}_{\text{batch}})$

Given:
- linear code $\mathscr{C} \subseteq \mathbb{F}^n$
- proximity parameter δ
- purported codewords $u_1, \ldots, u_\ell \in \mathbb{F}^n$ (oracles)

Batch proximity test $(\mathcal{P}_{\text{batch}}, \mathcal{V}_{\text{batch}})$

Given:
- linear code $\mathscr{C} \subseteq \mathbb{F}^n$
- proximity parameter δ
- purported codewords $u_1, \ldots, u_\ell \in \mathbb{F}^n$ (oracles)

1. $\mathcal{V}_{\text{batch}} \to \mathcal{P}_{\text{batch}} : (z_1, \ldots, z_\ell) \xleftarrow{\$} \mathbb{F}^\ell$.
2. $\mathcal{P}_{\text{batch}}$ and $\mathcal{V}_{\text{batch}}$ run $(\mathcal{P}, \mathcal{V})$ to check δ-proximity of $\sum z_i u_i$ to $\mathscr{C}$.

Key properties:

- If $u_1, \ldots, u_\ell \in \mathscr{C}$, then $\sum z_i u_i \in \mathscr{C}$.
- For every $\delta \in (0, \frac{1}{2})$, if $\max_i \Delta(u_i, \mathscr{C}) \geq \delta$, then

$$\Pr_{z_1,\ldots,z_\ell \leftarrow \mathbb{F}^\ell} \left[\Delta\left(\sum z_i u_i, \mathscr{C}\right) < 2\delta\right] \leq \frac{1}{|\mathbb{F}|}.$$

Correlated agreements

Many situations require a **stronger guarantee**.

If there are many $(z_1, \dots, z_\ell) \in \mathbb{F}^\ell$ such that $\sum z_i u_i$ is close to $\mathscr{C}$, it must be because $u_1, \dots, u_\ell \in \mathbb{F}^n$ have large **correlated agreement** with the code $\mathscr{C}$:

$$\exists T \subseteq [n], \exists c_1, \dots, c_\ell \in \mathscr{C} \text{ s.t. } \begin{cases} |T| > (1-\delta)n, \\ \forall i \in [\ell], u_{i\,|T} = c_{i\,|T}. \end{cases}$$

Correlated agreements

Many situations require a **stronger guarantee**.

If there are many $(z_1, \ldots, z_\ell) \in \mathbb{F}^\ell$ such that $\sum z_i u_i$ is close to $\mathscr{C}$, it must be because $u_1, \ldots, u_\ell \in \mathbb{F}^n$ have large **correlated agreement** with the code $\mathscr{C}$:

$$\exists T \subseteq [n], \exists c_1, \ldots, c_\ell \in \mathscr{C} \text{ s.t. } \begin{cases} |T| > (1-\delta)n, \\ \forall i \in [\ell], u_{i\,|T} = c_{i\,|T}. \end{cases}$$

- **Example 1.** Soundness of IOP system requires oracles $u_1, \ldots, u_\ell$ to be close to different codes $\mathscr{C}_1, \ldots, \mathscr{C}_\ell$ with different rates.
  - e.g. Reed-Solomon codes with different degree bounds.

Correlated agreements

Many situations require a **stronger guarantee**.

If there are many $(z_1, \ldots, z_\ell) \in \mathbb{F}^\ell$ such that $\sum z_i u_i$ is close to $\mathscr{C}$, it must be because $u_1, \ldots, u_\ell \in \mathbb{F}^n$ have large **correlated agreement** with the code $\mathscr{C}$:

$$\exists T \subseteq [n], \exists c_1, \ldots, c_\ell \in \mathscr{C} \text{ s.t. } \begin{cases} |T| > (1-\delta)n, \\ \forall i \in [\ell], u_{i\,|T} = c_{i\,|T}. \end{cases}$$

- **Example 1.** Soundness of IOP system requires oracles $u_1, \ldots, u_\ell$ to be close to different codes $\mathscr{C}_1, \ldots, \mathscr{C}_\ell$ with different rates.
  - e.g. Reed-Solomon codes with different degree bounds.
- **Example 2.** Soundness analysis of IOPs of Proximity for linear codes. [BBHR18, BKS18, BGKS20, BCIKS20, BCG20, ABN22, **B**LNR22]

Correlated agreement = proximity to interleaved code

Vectors $u_1, \ldots, u_\ell \in \mathbb{F}^n$ have large **correlated agreement** with $\mathscr{C}$:

$$\exists T \subseteq [n], \exists c_1, \ldots, c_\ell \in \mathscr{C} \text{ s.t. } \begin{cases} |T| > (1-\delta)n, \\ \forall i \in [\ell], u_{i\,|T} = c_{i\,|T}. \end{cases}$$

Interleaved code

$$\mathscr{C}^\ell := \left\{ \boldsymbol{C} = \begin{pmatrix} -c_1- \\ \vdots \\ -c_\ell- \end{pmatrix} \in \mathbb{F}^{\ell \times n} : \forall i \in [\ell], c_i \in \mathscr{C} \right\}$$

Correlated agreement = proximity to interleaved code

Vectors $u_1, \ldots, u_\ell \in \mathbb{F}^n$ have large **correlated agreement** with $\mathscr{C}$:

$$\exists T \subseteq [n], \exists c_1, \ldots, c_\ell \in \mathscr{C} \text{ s.t. } \begin{cases} |T| > (1-\delta)n, \\ \forall i \in [\ell], u_{i\,|T} = c_{i\,|T}. \end{cases}$$

Interleaved code

$$\mathscr{C}^\ell := \left\{ \boldsymbol{C} = \begin{pmatrix} -c_1- \\ \vdots \\ -c_\ell- \end{pmatrix} \in \mathbb{F}^{\ell \times n} : \forall i \in [\ell], c_i \in \mathscr{C} \right\} \qquad \boldsymbol{U} := \begin{pmatrix} -u_1- \\ \vdots \\ -u_\ell- \end{pmatrix} \in \mathbb{F}^{\ell \times n}$$

Correlated agreement = proximity to interleaved code

Vectors $u_1, \ldots, u_\ell \in \mathbb{F}^n$ have large **correlated agreement** with $\mathscr{C}$:

$$\exists T \subseteq [n], \exists c_1, \ldots, c_\ell \in \mathscr{C} \text{ s.t. } \begin{cases} |T| > (1-\delta)n, \\ \forall i \in [\ell], u_{i\,|T} = c_{i\,|T}. \end{cases}$$

Interleaved code

$$\mathscr{C}^\ell := \left\{ \boldsymbol{C} = \begin{pmatrix} -c_1- \\ \vdots \\ -c_\ell- \end{pmatrix} \in \mathbb{F}^{\ell \times n} : \forall i \in [\ell], c_i \in \mathscr{C} \right\} \qquad \boldsymbol{U} := \begin{pmatrix} -u_1- \\ \vdots \\ -u_\ell- \end{pmatrix} \in \mathbb{F}^{\ell \times n}$$

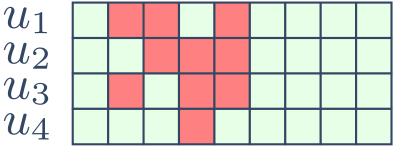
$$\text{Correlated agreement} \quad \longleftrightarrow \quad \Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) < \delta$$

Maximum distance vs column-wise distance

$$\mathscr{C} \subseteq \mathbb{F}^9 \qquad \boldsymbol{U} := \begin{pmatrix} -u_1- \\ -u_2- \\ -u_3- \\ -u_4- \end{pmatrix} \in \mathbb{F}^{4\times 9}$$

Green = correct

Red = error



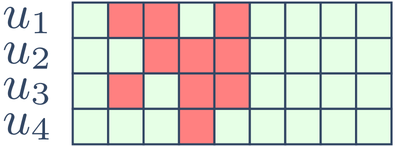
$$\max_i \Delta(u_i, \mathscr{C}) =$$

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) =$$

Maximum distance vs column-wise distance

$$\mathscr{C} \subseteq \mathbb{F}^9 \qquad \boldsymbol{U} := \begin{pmatrix} -u_1- \\ -u_2- \\ -u_3- \\ -u_4- \end{pmatrix} \in \mathbb{F}^{4\times 9}$$

Green = correct
Red = error



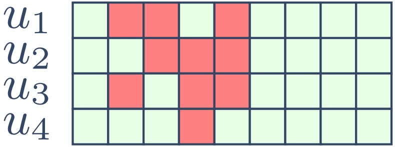
$$\max_i \Delta(u_i, \mathscr{C}) = 3/9$$
$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) =$$

Maximum distance vs column-wise distance

$$\mathscr{C} \subseteq \mathbb{F}^9 \qquad \boldsymbol{U} := \begin{pmatrix} -u_1- \\ -u_2- \\ -u_3- \\ -u_4- \end{pmatrix} \in \mathbb{F}^{4\times 9}$$

Green = correct

Red = error



$$\max_i \Delta(u_i, \mathscr{C}) = 3/9$$

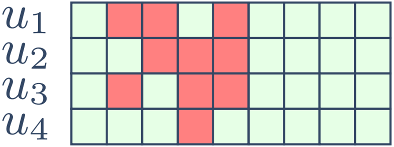
$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) = 4/9$$

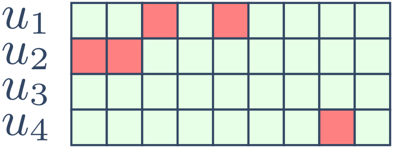
Maximum distance vs column-wise distance

$\mathscr{C} \subseteq \mathbb{F}^9$ $\quad\quad$ $\boldsymbol{U} := \begin{pmatrix} -u_1- \\ -u_2- \\ -u_3- \\ -u_4- \end{pmatrix} \in \mathbb{F}^{4\times 9}$

Green = correct

Red = error





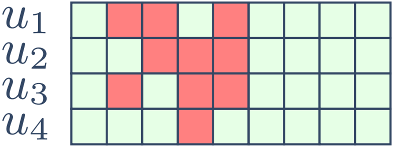
$$\max_i \Delta(u_i, \mathscr{C}) = 3/9$$
$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) = 4/9$$

Maximum distance vs column-wise distance

$$\mathscr{C} \subseteq \mathbb{F}^9 \qquad \boldsymbol{U} := \begin{pmatrix} -u_1- \\ -u_2- \\ -u_3- \\ -u_4- \end{pmatrix} \in \mathbb{F}^{4\times 9}$$

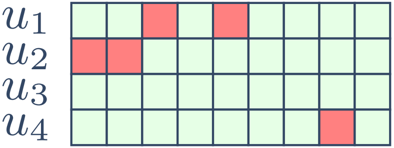
Green = correct

Red = error



$$\max_i \Delta(u_i, \mathscr{C}) = 3/9$$

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) = 4/9$$



$$\max_i \Delta(u_i, \mathscr{C}) = 2/9$$

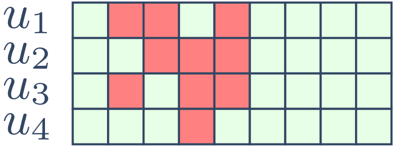
Maximum distance vs column-wise distance

$\mathscr{C} \subseteq \mathbb{F}^9$

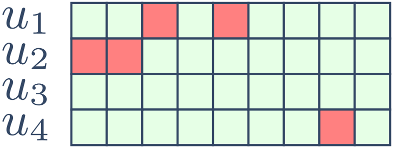
$$\boldsymbol{U} := \begin{pmatrix} -u_1- \\ -u_2- \\ -u_3- \\ -u_4- \end{pmatrix} \in \mathbb{F}^{4\times 9}$$

Green = correct

Red = error



$$\max_i \Delta(u_i, \mathscr{C}) = 3/9$$
$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) = 4/9$$



$$\max_i \Delta(u_i, \mathscr{C}) = 2/9$$
$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) = 5/9$$

Distance Preservation to Interleaved Codes

Distance preservation with random linear combinations

Distance preservation. There exists Λ s.t. for every $\delta \in (0, \Lambda)$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{z} \leftarrow \mathbb{F}^\ell}\left[\Delta\left(\boldsymbol{z} \cdot \boldsymbol{U}, \mathscr{C}\right) < \sigma(\delta)\right] \leq \tau.$$

$(\sigma(\delta) \approx \delta)$

Distance preservation with random linear combinations

Distance preservation. There exists Λ s.t. for every $\delta \in (0, \Lambda)$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{z} \leftarrow \mathbb{F}^\ell}[\Delta(\boldsymbol{z} \cdot \boldsymbol{U}, \mathscr{C}) < \sigma(\delta)] \leq \tau.$$

$$(\sigma(\delta) \approx \delta)$$

Proximity range Λ **New distance** $\sigma(\delta)$ **Error** τ

Distance preservation with random linear combinations

Distance preservation. There exists Λ s.t. for every $\delta \in (0, \Lambda)$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{z} \leftarrow \mathbb{F}^\ell}\left[\Delta\left(\boldsymbol{z} \cdot \boldsymbol{U}, \mathscr{C}\right) < \sigma(\delta)\right] \leq \tau.$$

$(\sigma(\delta) \approx \delta)$

	Proximity range Λ	New distance $\sigma(\delta)$	Error τ	
[AHIV17]	$\frac{\delta_{\min}}{4}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	Unique-decoding
[RZ17]	$\frac{\delta_{\min}}{3}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	

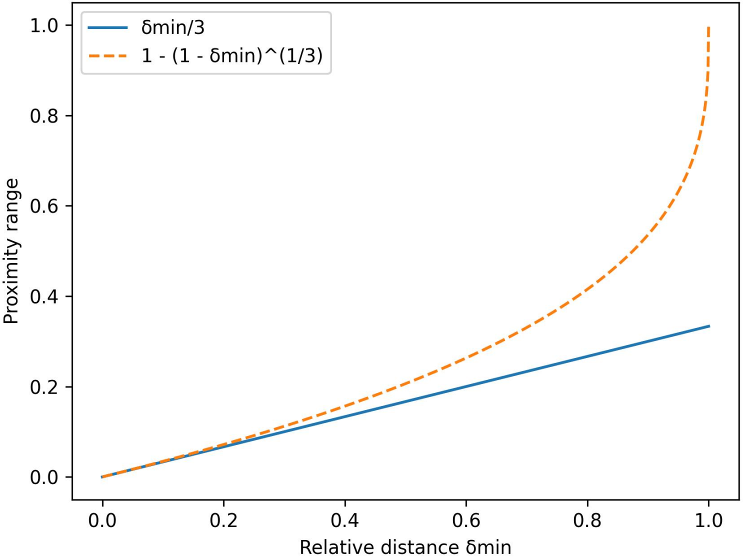
Distance preservation with random linear combinations

Distance preservation. There exists Λ s.t. for every $\delta \in (0, \Lambda)$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{z} \leftarrow \mathbb{F}^\ell}[\Delta(\boldsymbol{z} \cdot \boldsymbol{U}, \mathscr{C}) < \sigma(\delta)] \leq \tau.$$

$(\sigma(\delta) \approx \delta)$

	Proximity range Λ	**New distance** $\sigma(\delta)$	**Error** τ	
[AHIV17]	$\frac{\delta_{\min}}{4}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	Unique-decoding
[RZ17]	$\frac{\delta_{\min}}{3}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	Unique-decoding
[BKS18]	$1 - \sqrt[4]{1 - \delta_{\min} + \eta}$	$\delta - \eta$	$\frac{2}{\eta^3 \lvert\mathbb{F}\rvert}$	List-decoding
[BGKS20]	$1 - \sqrt[3]{1 - \delta_{\min} + \eta}$	$\delta - \eta$	$\frac{2}{\eta^2 \lvert\mathbb{F}\rvert}$	List-decoding


δmin/3
1 - (1 - δmin)^(1/3)
Proximity range
Relative distance δmin
0.0
0.2
0.4
0.6
0.8
1.0

Distance preservation with random linear combinations

Distance preservation. There exists Λ s.t. for every $\delta \in (0, \Lambda)$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{z} \leftarrow \mathbb{F}^\ell}[\Delta(\boldsymbol{z} \cdot \boldsymbol{U}, \mathscr{C}) < \sigma(\delta)] \leq \tau.$$

$(\sigma(\delta) \approx \delta)$

	Proximity range Λ	New distance $\sigma(\delta)$	Error τ	
[AHIV17]	$\frac{\delta_{\min}}{4}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	Unique-decoding
[RZ17]	$\frac{\delta_{\min}}{3}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	
[BKS18]	$1 - \sqrt[4]{1 - \delta_{\min} + \eta}$	$\delta - \eta$	$\frac{2}{\eta^3 \lvert\mathbb{F}\rvert}$	List-decoding
[BGKS20]	$1 - \sqrt[3]{1 - \delta_{\min} + \eta}$	$\delta - \eta$	$\frac{2}{\eta^2 \lvert\mathbb{F}\rvert}$	

- $\Lambda = 1 - \sqrt[3]{1 - \delta_{\min} + \eta}$ is sharp for some codes with linear-size alphabet.

Distance preservation with random linear combinations

Distance preservation. There exists Λ s.t. for every $\delta \in (0, \Lambda)$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{z} \leftarrow \mathbb{F}^\ell}\left[\Delta\left(\boldsymbol{z} \cdot \boldsymbol{U}, \mathscr{C}\right) < \sigma(\delta)\right] \leq \tau.$$

$$(\sigma(\delta) \approx \delta)$$

	Proximity range Λ	**New distance** $\sigma(\delta)$	**Error** τ	
[AHIV17]	$\frac{\delta_{\min}}{4}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	Unique-decoding
[RZ17]	$\frac{\delta_{\min}}{3}$	δ	$\frac{\delta n}{\lvert\mathbb{F}\rvert}$	Unique-decoding
[BKS18]	$1 - \sqrt[4]{1 - \delta_{\min} + \eta}$	$\delta - \eta$	$\frac{2}{\eta^3 \lvert\mathbb{F}\rvert}$	List-decoding
[BGKS20]	$1 - \sqrt[3]{1 - \delta_{\min} + \eta}$	$\delta - \eta$	$\frac{2}{\eta^2 \lvert\mathbb{F}\rvert}$	List-decoding

- $\Lambda = 1 - \sqrt[3]{1 - \delta_{\min} + \eta}$ is sharp for some codes with linear-size alphabet.
- Better parameters for **specific** family of codes (Reed-Solomon) [BCIKS20].

What about distribution $\neq$ uniform?

Possible to sample coefficients from distribution $\neq$ uniform?

What about distribution $\neq$ uniform?

Possible to sample coefficients from distribution $\neq$ uniform?

Example. For every $\eta \in (0,1)$ and every $0 < \delta < 1 - \sqrt[\ell]{1 - \delta_{\min} + \eta}$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \overset{\text{[BKS18]}}{\Longrightarrow} \Pr_{x \leftarrow \mathbb{F}}\left[\Delta\left((1, x, x^2, \ldots, x^{\ell-1}) \cdot \boldsymbol{U}, \mathscr{C}\right) < \delta - \eta\right] \leq \left(\frac{2}{\eta}\right)^{\ell+1} \cdot \frac{\ell - 1}{|\mathbb{F}|}$$

What about distribution $\neq$ uniform?

Possible to sample coefficients from distribution $\neq$ uniform?

Example. For every $\eta \in (0,1)$ and every $0 < \delta < 1 - \sqrt[\ell]{1 - \delta_{\min} + \eta}$,

$$\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \overset{\text{[BKS18]}}{\Longrightarrow} \Pr_{x \leftarrow \mathbb{F}}\left[\Delta\left((1, x, x^2, \ldots, x^{\ell-1}) \cdot \boldsymbol{U}, \mathscr{C}\right) < \delta - \eta\right] \leq \left(\frac{2}{\eta}\right)^{\ell+1} \cdot \frac{\ell - 1}{|\mathbb{F}|}$$

Why reduce randomness complexity?

- concrete efficiency of IOPs used in real-world (e.g. FRI, STARKs)
- sometimes necessary, e.g. IOPs with linear-time prover [BCL22, BCGL22]

(Recall) Goal : Batch proximity testing

We are looking for generators $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ that allow randomness-efficient batch proximity testing.

(Recall) Goal : Batch proximity testing

We are looking for generators $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ that allow randomness-efficient batch proximity testing.

Batch proximity test $(\mathcal{P}_{\text{batch}}, \mathcal{V}_{\text{batch}})$

Given:
- linear code $\mathscr{C} \subseteq \mathbb{F}^n$
- proximity parameter δ
- purported codewords $u_1, \ldots, u_\ell \in \mathbb{F}^n$ (oracles)

1. $\mathcal{V}_{\text{batch}} \to \mathcal{P}_{\text{batch}} : \boldsymbol{x} \xleftarrow{\$} \mathbb{F}^s$.
2. $\mathcal{P}_{\text{batch}}$ and $\mathcal{V}_{\text{batch}}$ run $(\mathcal{P}, \mathcal{V})$ to check δ-proximity of $\sum G(\boldsymbol{x})_i u_i$ to $\mathscr{C}$.

Distance-Preserving Generators

Warm-up: Epsilon-biased generators

Warm-up: Epsilon-biased generators

Parameters: $\ell \geq s \geq 1$ integers, $\varepsilon \in (0,1)$.

A function $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ is an ε**-biased generator** for $\mathbb{F}^\ell$ if

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell \times n}, \qquad \boldsymbol{U} \neq \mathbf{0}^{\ell \times n} \implies \Pr_{\boldsymbol{x} \leftarrow \mathbb{F}^s} [G(\boldsymbol{x}) \cdot \boldsymbol{U} = \mathbf{0}^n] \leq \varepsilon.$$

Warm-up: Epsilon-biased generators

Parameters: $\ell \geq s \geq 1$ integers, $\varepsilon \in (0,1)$.

A function $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is an ε**-biased generator** for $\mathbb{F}^\ell$ if

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell \times n}, \qquad \boldsymbol{U} \neq \boldsymbol{0}^{\ell \times n} \implies \Pr_{\boldsymbol{x} \leftarrow \mathbb{F}^s}[G(\boldsymbol{x}) \cdot \boldsymbol{U} = \boldsymbol{0}^n] \leq \varepsilon.$$

Numerous applications in theoretical computer science (derandomization, error-correcting codes, probabilistic proofs, ...).

Warm-up: Epsilon-biased generators

Parameters: $\ell \geq s \geq 1$ integers, $\varepsilon \in (0,1)$.

A function $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ is an ε**-biased generator** for $\mathbb{F}^\ell$ if

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell \times n}, \qquad \boldsymbol{U} \neq \mathbf{0}^{\ell \times n} \implies \Pr_{\boldsymbol{x} \leftarrow \mathbb{F}^s}[G(\boldsymbol{x}) \cdot \boldsymbol{U} = \mathbf{0}^n] \leq \varepsilon.$$

Numerous applications in theoretical computer science (derandomization, error-correcting codes, probabilistic proofs, ...).

Seed space	**Generator**	**Bias** ε
$\mathbb{F}^\ell$	$G(\boldsymbol{x}) = \boldsymbol{x}$	$\frac{1}{\lvert\mathbb{F}\rvert}$

Warm-up: Epsilon-biased generators

Parameters: $\ell \geq s \geq 1$ integers, $\varepsilon \in (0,1)$.

A function $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is an ε**-biased generator** for $\mathbb{F}^\ell$ if

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell \times n}, \qquad \boldsymbol{U} \neq \mathbf{0}^{\ell \times n} \implies \Pr_{\boldsymbol{x} \leftarrow \mathbb{F}^s}[G(\boldsymbol{x}) \cdot \boldsymbol{U} = \mathbf{0}^n] \leq \varepsilon.$$

Numerous applications in theoretical computer science (derandomization, error-correcting codes, probabilistic proofs, ...).

Seed space	**Generator**	**Bias** ε
$\mathbb{F}^\ell$	$G(\boldsymbol{x}) = \boldsymbol{x}$	$\frac{1}{\lvert\mathbb{F}\rvert}$
$\mathbb{F}$	$G(x) = (1, x, \ldots, x^{\ell-1})$	$\frac{\ell-1}{\lvert\mathbb{F}\rvert}$

Warm-up: Epsilon-biased generators

Parameters: $\ell \geq s \geq 1$ integers, $\varepsilon \in (0,1)$.

A function $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is an ε**-biased generator** for $\mathbb{F}^\ell$ if

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell \times n}, \qquad \boldsymbol{U} \neq \mathbf{0}^{\ell \times n} \implies \Pr_{\boldsymbol{x} \leftarrow \mathbb{F}^s}[G(\boldsymbol{x}) \cdot \boldsymbol{U} = \mathbf{0}^n] \leq \varepsilon.$$

Numerous applications in theoretical computer science (derandomization, error-correcting codes, probabilistic proofs, ...).

Seed space	Generator	Bias ε
$\mathbb{F}^\ell$	$G(\boldsymbol{x}) = \boldsymbol{x}$	$\frac{1}{\lvert\mathbb{F}\rvert}$
$\mathbb{F}$	$G(x) = (1, x, \ldots, x^{\ell-1})$	$\frac{\ell-1}{\lvert\mathbb{F}\rvert}$
$\mathbb{F}^s, 2^s = \ell$	$G(\boldsymbol{x}) = (\prod_i x_i^{b_i})_{b \in \{0,1\}^s}$	$\frac{s}{\lvert\mathbb{F}\rvert}$

Distance-preserving generators

Parameters: $\Lambda \in (0,1)$, $\sigma\colon (0,1) \to (0,1)$ non-increasing fct, $\tau \in (0,1)$.

A function $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is a (Λ,σ,τ)-**distance-preserving generator** if for every code $\mathscr{C} \subseteq \mathbb{F}^n$ and every $\delta \in (0,\Lambda)$:

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell\times n}, \quad \Delta_{\mathbb{F}^\ell}(\boldsymbol{U},\mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{x}\leftarrow\mathbb{F}^s}\left[\Delta\left(G(\boldsymbol{x})\cdot\boldsymbol{U},\mathscr{C}\right) < \sigma(\delta)\right] \leq \tau.$$

Distance-preserving generators

Parameters: $\Lambda \in (0,1)$, $\sigma\colon (0,1) \to (0,1)$ non-increasing fct, $\tau \in (0,1)$.

A function $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is a (Λ, σ, τ)**-distance-preserving generator** if for every code $\mathscr{C} \subseteq \mathbb{F}^n$ and every $\delta \in (0, \Lambda)$:

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell \times n}, \quad \Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{x} \leftarrow \mathbb{F}^s}\left[\Delta\left(G(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}\right) < \sigma(\delta)\right] \leq \tau.$$

t

Seed space	Generator	Bias ε	Dist. preserving?
$\mathbb{F}^\ell$	$G(\boldsymbol{x}) = \boldsymbol{x}$	$\frac{1}{\lvert\mathbb{F}\rvert}$	✔
$\mathbb{F}$	$G(x) = (1, x, \ldots, x^{\ell-1})$	$\frac{\ell-1}{\lvert\mathbb{F}\rvert}$	✔ [BKS18]
$\mathbb{F}^s, 2^s = \ell$	$G(\boldsymbol{x}) = (\prod_i x_i^{b_i})_{b \in \{0,1\}^s}$	$\frac{s}{\lvert\mathbb{F}\rvert}$	✔ [ABN22]

From prior work: known distance-preserving generators are in particular biased.

Distance-preserving generators

Parameters: $\Lambda \in (0,1)$, $\sigma\colon (0,1) \to (0,1)$ non-increasing fct, $\tau \in (0,1)$.

A function $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is a (Λ,σ,τ)**-distance-preserving generator** if for every code $\mathscr{C} \subseteq \mathbb{F}^n$ and every $\delta \in (0,\Lambda)$:

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell \times n}, \quad \Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{x} \leftarrow \mathbb{F}^s}\left[\Delta\left(G(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}\right) < \sigma(\delta)\right] \leq \tau.$$

Easy fact: G is (Λ,σ,τ)-distance-preserving $\implies$ G is τ-biased.
(because G preserves distance to $\{0^n\}$.)

Distance-preserving generators

Parameters: $\Lambda \in (0,1)$, $\sigma\colon (0,1) \to (0,1)$ non-increasing fct, $\tau \in (0,1)$.

A function $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is a (Λ,σ,τ)**-distance-preserving generator** if for every code $\mathscr{C} \subseteq \mathbb{F}^n$ and every $\delta \in (0,\Lambda)$:

$$\forall \boldsymbol{U} \in \mathbb{F}^{\ell\times n}, \quad \Delta_{\mathbb{F}^\ell}(\boldsymbol{U},\mathscr{C}^\ell) \geq \delta \implies \Pr_{\boldsymbol{x}\leftarrow\mathbb{F}^s}\left[\Delta\left(G(\boldsymbol{x})\cdot\boldsymbol{U},\mathscr{C}\right) < \sigma(\delta)\right] \leq \tau.$$

Easy fact: G is (Λ,σ,τ)-distance-preserving $\implies$ G is τ-biased.
(because G preserves distance to $\{0^n\}$.)

Question: Do all biased generators preserve distance?

Polynomial Generators Preserve Distance

Polynomial generators

Let s, ℓ, d be positive integers such that $d \leq |\mathbb{F}|$ and $\max(s, 2) \leq \ell \leq \binom{s+d}{s}$.

Polynomial generator

A function $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ is a *degree-d generator* if there exist ℓ linearly independent polynomials $P_1, \ldots, P_\ell \in \mathbb{F}[X_1, \ldots, X_s]$ of total degree at most d such that

$$\forall \boldsymbol{x} \in \mathbb{F}^s, \qquad G(\boldsymbol{x}) = (P_i(\boldsymbol{x}))_{1 \leq i \leq \ell}.$$

Polynomial generators

Let s, ℓ, d be positive integers such that $d \leq |\mathbb{F}|$ and $\max(s, 2) \leq \ell \leq \binom{s+d}{s}$.

Polynomial generator

A function $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ is a *degree-d generator* if there exist ℓ linearly independent polynomials $P_1, \ldots, P_\ell \in \mathbb{F}[X_1, \ldots, X_s]$ of total degree at most d such that

$$\forall \boldsymbol{x} \in \mathbb{F}^s, \qquad G(\boldsymbol{x}) = (P_i(\boldsymbol{x}))_{1 \leq i \leq \ell}.$$

- Any degree-d generator is ε–biased with $\varepsilon = \frac{d}{|\mathbb{F}|}$. (Schwartz-Zippel)

Polynomial generators

Let s, ℓ, d be positive integers such that $d \leq |\mathbb{F}|$ and $\max(s, 2) \leq \ell \leq \binom{s+d}{s}$.

Polynomial generator

A function $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ is a *degree-d generator* if there exist ℓ linearly independent polynomials $P_1, \ldots, P_\ell \in \mathbb{F}[X_1, \ldots, X_s]$ of total degree at most d such that

$$\forall \boldsymbol{x} \in \mathbb{F}^s, \qquad G(\boldsymbol{x}) = (P_i(\boldsymbol{x}))_{1 \leq i \leq \ell}.$$

- Any degree-d generator is ε–biased with $\varepsilon = \frac{d}{|\mathbb{F}|}$. (Schwartz-Zippel)
- Distance-preserving generators from literature are special cases of polynomial generators.

Main result

Theorem

Any degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)–distance-preserving.

	Proximity range Λ	**New distance** $\sigma(\delta)$	**Error** τ
Unique-decoding	$\frac{\delta_{\min}}{d+2}$	δ	$\delta n \cdot \frac{d}{\lvert\mathbb{F}\rvert}$
List-decoding	$1 - \sqrt[d+2]{1 - \delta_{\min} + \eta}$	δ	$\delta n \cdot \frac{\ell+1}{\eta} \cdot \frac{d}{\lvert\mathbb{F}\rvert}$

Main result

Theorem

Any degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)–distance-preserving.

	Proximity range Λ	**New distance** $\sigma(\delta)$	**Error** τ
Unique-decoding	$\frac{\delta_{\min}}{d+2}$	δ	$\delta n \cdot \frac{d}{\lvert\mathbb{F}\rvert}$
List-decoding	$1 - \sqrt[d+2]{1 - \delta_{\min} + \eta}$	δ	$\delta n \cdot \frac{\ell+1}{\eta} \cdot \frac{d}{\lvert\mathbb{F}\rvert}$

- **Implies prior results** about distance-preserving generators

Main result

Theorem

Any degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)–distance-preserving.

	Proximity range Λ	**New distance** $\sigma(\delta)$	**Error** τ
Unique-decoding	$\frac{\delta_{\min}}{d+2}$	δ	$\delta n \cdot \frac{d}{\lvert\mathbb{F}\rvert}$
List-decoding	$1 - \sqrt[d+2]{1 - \delta_{\min} + \eta}$	δ	$\delta n \cdot \frac{\ell+1}{\eta} \cdot \frac{d}{\lvert\mathbb{F}\rvert}$

- **Implies prior results** about distance-preserving generators
- **Improves prior results**
  - For $G(x) = (x^i)_{0 \le i < \ell}$, **remove** from τ the **exponential dependence** in ℓ from [BKS18]
  - **Exact** distance preservation (instead of *approximate*)

Application: Proximity gaps for all linear codes

Theorem $\implies$ Proximity gaps for all linear codes

Let $\delta \in (0, \Lambda)$. Let $\mathscr{C}$ be a linear code and let $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ be a polynomial generator. Exactly one of the following two statements holds:

$$(1)\, \Pr\left[\Delta\left(G(\boldsymbol{x})^\top \cdot \boldsymbol{U}, \mathscr{C}\right) < \delta\right] = 1 \quad \textbf{OR} \quad (2)\, \Pr\left[\Delta\left(G(\boldsymbol{x})^\top \cdot \boldsymbol{U}, \mathscr{C}\right) < \delta\right] \leq \tau.$$

Previous work on proximity gaps:

- All linear codes – uniform coefficients, $\delta < \frac{\delta_{\min}}{3}$ [AHIV17, RZ17]
- RS codes – uniform coefficients & powers, $\delta < 1 - \sqrt{1 - \delta_{\min}}$ [BCIKS20]

Application: Proximity gaps for all linear codes

Theorem $\Longrightarrow$ Proximity gaps for all linear codes

Let $\delta \in (0, \Lambda)$. Let $\mathscr{C}$ be a linear code and let $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ be a polynomial generator. Exactly one of the following two statements holds:

$$(1)\ \Pr\left[\Delta\left(G(\boldsymbol{x})^\top \cdot \boldsymbol{U}, \mathscr{C}\right) < \delta\right] = 1 \quad \textbf{OR} \quad (2)\ \Pr\left[\Delta\left(G(\boldsymbol{x})^\top \cdot \boldsymbol{U}, \mathscr{C}\right) < \delta\right] \leq \tau.$$

Previous work on proximity gaps:

- All linear codes – uniform coefficients, $\delta < \frac{\delta_{\min}}{3}$ [AHIV17, RZ17]
- RS codes – uniform coefficients & powers, $\delta < 1 - \sqrt{1 - \delta_{\min}}$ [BCIKS20]

In fact, nearly all combinations are at the same distance.

If $\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell) \in (0, \Lambda)$, then $\Pr\left[\Delta(G(\boldsymbol{x})^\top \cdot \boldsymbol{U}, \mathscr{C}) \neq \Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathscr{C}^\ell)\right] \leq \tau$.

Technical Overview

Proof overview

Theorem

Any degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

	Proximity range Λ	New distance σ	Error τ
Unique-decoding	$\frac{\delta_{\min}}{d+2}$	δ	$\delta n \cdot \frac{d}{\lvert\mathbb{F}\rvert}$
List-decoding	$1 - \sqrt[d+2]{1 - \delta_{\min} + \eta}$	δ	$\delta n \cdot \frac{\ell+1}{\eta} \cdot \frac{d}{\lvert\mathbb{F}\rvert}$

Proof overview

Any **univariate** degree-d generator $G \colon \mathbb{F} \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

⬇

Theorem

Any degree-d generator $G \colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

	Proximity range Λ	New distance σ	Error τ
Unique-decoding	$\frac{\delta_{\min}}{d+2}$	δ	$\delta n \cdot \frac{d}{\lvert\mathbb{F}\rvert}$
List-decoding	$1 - \sqrt[d+2]{1 - \delta_{\min} + \eta}$	δ	$\delta n \cdot \frac{\ell+1}{\eta} \cdot \frac{d}{\lvert\mathbb{F}\rvert}$

Proof overview

Any **univariate** degree-d generator $G\colon \mathbb{F} \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

⬇

Any **multivariate** degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

Proof overview

Generators from MDS codes are distance-preserving.

⬇

Any **univariate** degree-d generator $G\colon \mathbb{F} \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

⬇

Any **multivariate** degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

Proof overview

Generators from MDS codes are distance-preserving.

⬇

Any **univariate** degree-d generator $G\colon \mathbb{F} \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

⬇

Any **multivariate** degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is (Λ, σ, τ)-distance-preserving.

Epsilon-biased generators from codes with good distance

Generators from linear codes

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Epsilon-biased generators from codes with good distance

Generators from linear codes

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Consider the evaluation map $\mathrm{ev}\colon \begin{array}{ll} \mathcal{L} & \to \mathbb{F}^N \\ f & \mapsto (f(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{F}^s) \end{array}$, where $N := |\mathbb{F}|^s$.

Epsilon-biased generators from codes with good distance

Generators from linear codes

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Consider the evaluation map $\mathrm{ev}\colon \begin{array}{ll} \mathcal{L} & \to \mathbb{F}^N \\ f & \mapsto (f(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{F}^s) \end{array}$, where $N := |\mathbb{F}|^s$.

We have:

- $\mathscr{D} = \mathrm{ev}(\mathcal{L})$ is a $[N, \ell]$-code.

Epsilon-biased generators from codes with good distance

Generators from linear codes

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Consider the evaluation map $\mathrm{ev}\colon \begin{array}{ll} \mathcal{L} & \to \mathbb{F}^N \\ f & \mapsto (f(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{F}^s) \end{array}$, where $N := |\mathbb{F}|^s$.

We have:

- $\mathscr{D} = \mathrm{ev}(\mathcal{L})$ is a $[N, \ell]$-code.
- If $\delta_{\min}(\mathscr{D}) \geq 1 - \varepsilon$, then $G_{\mathscr{D}}\colon \begin{array}{ll} \mathbb{F}^s & \to \mathbb{F}^\ell \\ \boldsymbol{x} & \mapsto (f_i(\boldsymbol{x}))_{i \in [\ell]} \end{array}$ is ε-biased.

A small-biased generator from a Reed-Solomon code

Example

Let $\ell \leq |\mathbb{F}|$. Consider the encoding map $\mathrm{ev} : \begin{array}{ll} \mathbb{F}[x]_{<\ell} & \to \mathbb{F}^{|\mathbb{F}|} \\ f & \mapsto (f(x) : x \in \mathbb{F}) \end{array}$.

A small-biased generator from a Reed-Solomon code

Example

Let $\ell \leq |\mathbb{F}|$. Consider the encoding map $\mathrm{ev}: \begin{array}{ll} \mathbb{F}[x]_{<\ell} & \to \mathbb{F}^{|\mathbb{F}|} \\ f & \mapsto (f(x) : x \in \mathbb{F}) \end{array}$.

- $\mathscr{D}$ is a Reed-Solomon code with parameters $[|\mathbb{F}|, \ell]$.

A small-biased generator from a Reed-Solomon code

Example

Let $\ell \leq |\mathbb{F}|$. Consider the encoding map $\mathrm{ev} : \begin{array}{ll} \mathbb{F}[x]_{<\ell} & \to \mathbb{F}^{|\mathbb{F}|} \\ f & \mapsto (f(x) : x \in \mathbb{F}) \end{array}$.

- $\mathscr{D}$ is a Reed-Solomon code with parameters $[|\mathbb{F}|, \ell]$.
- It has relative distance $\delta_{\min}(\mathscr{D}) = 1 - \frac{\ell - 1}{\mathbb{F}}$.

A small-biased generator from a Reed-Solomon code

Example

Let $\ell \leq |\mathbb{F}|$. Consider the encoding map $\mathrm{ev} \colon \begin{array}{ll} \mathbb{F}[x]_{<\ell} & \to \mathbb{F}^{|\mathbb{F}|} \\ f & \mapsto (f(x) : x \in \mathbb{F}) \end{array}$.

- $\mathscr{D}$ is a Reed-Solomon code with parameters $[|\mathbb{F}|, \ell]$.
- It has relative distance $\delta_{\min}(\mathscr{D}) = 1 - \frac{\ell-1}{\mathbb{F}}$.
- Let $(f_i)_{i\in[\ell]}$ be a basis of $\mathbb{F}[x]_{<\ell}$.

Then $G_{\mathscr{D}} \colon \begin{array}{ll} \mathbb{F} & \to \mathbb{F}^{\ell} \\ x & \mapsto (f_i(x))_{i\in[\ell]} \end{array}$ is $\frac{\ell-1}{\mathbb{F}}$-biased.

Generators from MDS codes preserve distance

Key Lemma

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Assume that $\mathscr{D} = \mathrm{ev}(\mathcal{L})$ is **MDS**, meaning $\delta_{\min}(\mathscr{D}) = 1 - \frac{\ell - 1}{N}$. $N := |\mathbb{F}^s|$

Generators from MDS codes preserve distance

Key Lemma

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Assume that $\mathscr{D} = \mathrm{ev}(\mathcal{L})$ is **MDS**, meaning $\delta_{\min}(\mathscr{D}) = 1 - \frac{\ell-1}{N}$. $N := |\mathbb{F}^s|$

Then $G_{\mathscr{D}} : \begin{array}{ll} \mathbb{F}^s & \to \mathbb{F}^\ell \\ \boldsymbol{x} & \mapsto (f_i(\boldsymbol{x}))_{i \in [\ell]} \end{array}$ is $\begin{cases} \text{1. } \varepsilon\text{-biased for } \varepsilon = \frac{\ell-1}{N}, \\ \end{cases}$

Generators from MDS codes preserve distance

Key Lemma

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Assume that $\mathscr{D} = \mathrm{ev}(\mathcal{L})$ is **MDS**, meaning $\delta_{\min}(\mathscr{D}) = 1 - \frac{\ell-1}{N}$. $N := |\mathbb{F}^s|$

Then $G_{\mathscr{D}}: \begin{array}{ll} \mathbb{F}^s & \to \mathbb{F}^\ell \\ \boldsymbol{x} & \mapsto (f_i(\boldsymbol{x}))_{i \in [\ell]} \end{array}$ is

1. ε-biased for $\varepsilon = \frac{\ell-1}{N}$,
2. (Λ, σ, τ)-distance-preserving.

Generators from MDS codes preserve distance

Key Lemma

Let $\mathcal{L} \subseteq \{\mathbb{F}^s \to \mathbb{F}\}$ be a $\mathbb{F}$-linear space and let $\{f_1, \ldots, f_\ell\}$ be a basis of $\mathcal{L}$.

Assume that $\mathscr{D} = \mathrm{ev}(\mathcal{L})$ is **MDS**, meaning $\delta_{\min}(\mathscr{D}) = 1 - \frac{\ell-1}{N}$. $\quad N := |\mathbb{F}^s|$

Then $G_{\mathscr{D}} : \begin{array}{ll} \mathbb{F}^s & \to \mathbb{F}^\ell \\ \boldsymbol{x} & \mapsto (f_i(\boldsymbol{x}))_{i \in [\ell]} \end{array}$ is $\begin{cases} \text{1. } \varepsilon\text{-biased for } \varepsilon = \frac{\ell-1}{N}, \\ \text{2. } (\Lambda, \sigma, \tau)\text{-distance-preserving.} \end{cases}$

	Proximity range Λ	**New distance $\sigma(\delta)$**	**Error τ**
Unique-decoding	$\frac{\delta_{\min}}{\ell+1}$	δ	$\delta n \cdot \varepsilon$
List-decoding	$1 - \sqrt[\ell+1]{1 - \delta_{\min} + \eta}$	δ	$\delta n \cdot \frac{\ell+1}{\eta} \cdot \varepsilon$

Proof of Key Lemma — Unique-decoding regime

Unique-decoding regime: $\delta < \frac{\delta_{\min}}{\ell+1}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Proof of Key Lemma — Unique-decoding regime

Unique-decoding regime: $\delta < \frac{\delta_{\min}}{\ell+1}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Step 1. Find $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

Proof of Key Lemma — Unique-decoding regime

Unique-decoding regime: $\delta < \frac{\delta_{\min}}{\ell+1}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Step 1. Find $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

- For each $\boldsymbol{x} \in A$, consider $c_{\boldsymbol{x}} \in \mathscr{C}$ that is δ-close to $G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}$.

Proof of Key Lemma — Unique-decoding regime

Unique-decoding regime: $\delta < \frac{\delta_{\min}}{\ell+1}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A$, $\Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Step 1. Find $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

- For each $\boldsymbol{x} \in A$, consider $c_{\boldsymbol{x}} \in \mathscr{C}$ that is δ-close to $G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}$.
- Take ℓ distinct $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_\ell \in A$.

Proof of Key Lemma — Unique-decoding regime

Unique-decoding regime: $\delta < \frac{\delta_{\min}}{\ell+1}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Step 1. Find $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

- For each $\boldsymbol{x} \in A$, consider $c_{\boldsymbol{x}} \in \mathscr{C}$ that is δ-close to $G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}$.
- Take ℓ distinct $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_\ell \in A$.
- Since $\mathscr{D}$ is MDS*, compute $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall i \in [\ell], c_{\boldsymbol{s}_i} = G_{\mathscr{D}}(\boldsymbol{s}_i) \cdot \mathbf{C}$.

* $[N, \ell]$-code $\mathscr{D}$ is MDS iff for any $S \subseteq \mathbb{F}^s$, $|S| = \ell$, $\{G_{\mathscr{D}}(\boldsymbol{s}) : \boldsymbol{s} \in S\}$ is linearly independent.

Proof of Key Lemma — Unique-decoding regime

Unique-decoding regime: $\delta < \frac{\delta_{\min}}{\ell+1}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A$, $\Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Step 1. Find $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

- For each $\boldsymbol{x} \in A$, consider $c_{\boldsymbol{x}} \in \mathscr{C}$ that is δ-close to $G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}$.
- Take ℓ distinct $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_\ell \in A$.
- Since $\mathscr{D}$ is MDS, compute $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall i \in [\ell], c_{\boldsymbol{s}_i} = G_{\mathscr{D}}(\boldsymbol{s}_i) \cdot \mathbf{C}$.
- Using $\delta < \frac{\delta_{\min}}{\ell+1}$, prove that $\forall \boldsymbol{x} \in A, c_{\boldsymbol{x}} = G_{\mathscr{D}}(\boldsymbol{x}) \cdot C$.

Proof of Key Lemma — Unique-decoding regime

Unique-decoding regime: $\delta < \frac{\delta_{\min}}{\ell+1}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A$, $\Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Step 1. Find $\mathbf{C} \in \mathscr{C}^{\ell}$ s.t. $\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

- For each $\boldsymbol{x} \in A$, consider $c_{\boldsymbol{x}} \in \mathscr{C}$ that is δ-close to $G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}$.
- Take ℓ distinct $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{\ell} \in A$.
- Since $\mathscr{D}$ is MDS, compute $\mathbf{C} \in \mathscr{C}^{\ell}$ s.t. $\forall i \in [\ell], c_{\boldsymbol{s}_i} = G_{\mathscr{D}}(\boldsymbol{s}_i) \cdot \mathbf{C}$.
- Using $\delta < \frac{\delta_{\min}}{\ell+1}$, prove that $\forall \boldsymbol{x} \in A, c_{\boldsymbol{x}} = G_{\mathscr{D}}(\boldsymbol{x}) \cdot C$.

Step 2. Prove that $\Delta_{\mathbb{F}^{\ell}}(\boldsymbol{U}, \mathbf{C}) < \delta$. (Follows from bias of $G_{\mathscr{D}}$)

~~Unique-decoding~~ List-decoding regime: $\delta < 1 - \sqrt[\ell+1]{1 - \delta_{\min}}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

Step 1. Find $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

- For each $\boldsymbol{x} \in A$, consider $c_{\boldsymbol{x}} \in \mathscr{C}$ that is δ-close to $G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}$.
- Take ℓ distinct $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_\ell \in A$.
- Since $\mathscr{D}$ is MDS, compute $\mathbf{C} \in \mathscr{C}^\ell$ s.t. $\forall i \in [\ell], c_{\boldsymbol{s}_i} = G_{\mathscr{D}}(\boldsymbol{s}_i) \cdot \mathbf{C}$.
- Using $\delta < \frac{\delta_{\min}}{\ell+1}$, prove that $\forall \boldsymbol{x} \in A, c_{\boldsymbol{x}} = G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathrm{C}$. $\leftarrow$ **FAIL**

Step 2. Prove that $\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathbf{C}) < \delta$. (Follows from bias of $G_{\mathscr{D}}$)

Proof of Key Lemma — ~~Unique-decoding regime~~ List-decoding regime

~~Unique-decoding~~ List-decoding regime: $\delta < 1 - \sqrt[\ell+1]{1 - \delta_{\min}}$

Assume $\exists A \subseteq \mathbb{F}^s, |A| > \tau \cdot |\mathbb{F}^s|$ s.t.
$\forall \boldsymbol{x} \in A, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, \mathscr{C}) < \delta$.

New Step 1. Find a large subset $B \subseteq A$ and $\mathbf{C} \in \mathscr{C}^\ell$ such that
$\forall \boldsymbol{x} \in B, \Delta(G_{\mathscr{D}}(\boldsymbol{x}) \cdot \boldsymbol{U}, G_{\mathscr{D}}(\boldsymbol{x}) \cdot \mathbf{C}) < \delta$.

More challenging because codewords are very noisy.

Step 2. Prove that $\Delta_{\mathbb{F}^\ell}(\boldsymbol{U}, \mathbf{C}) < \delta$. (Follows from bias of $G_{\mathscr{D}}$)

Proof overview

Generators from MDS codes are distance-preserving.

⬇

Any **univariate** degree-d generator $G\colon \mathbb{F} \to \mathbb{F}^{\ell}$ is distance-preserving.

Proof overview

Generators from MDS codes are distance-preserving.

⬇

Any **univariate** degree-d generator $G\colon \mathbb{F} \to \mathbb{F}^{\ell}$ is distance-preserving.

◇ Consider $G_{\mathscr{D}}\colon \mathbb{F} \to \mathbb{F}^{d+1}$ where $\mathscr{D}$ is a RS code of dimension $d+1$.

Proof overview

Generators from MDS codes are distance-preserving.

⬇

Any **univariate** degree-d generator $G\colon \mathbb{F} \to \mathbb{F}^{\ell}$ is distance-preserving.

- Consider $G_{\mathscr{D}}\colon \mathbb{F} \to \mathbb{F}^{d+1}$ where $\mathscr{D}$ is a RS code of dimension $d+1$.
- $G_{\mathscr{D}}$ preserves distance $\implies$ ditto for $G\colon \mathbb{F} \to \mathbb{F}^{\ell}$

Proof overview

Generators from MDS codes are distance-preserving.

⬇

Any **univariate** degree-d generator $G\colon \mathbb{F} \to \mathbb{F}^\ell$ is distance-preserving.

⬇

Any **multivariate** degree-d generator $G\colon \mathbb{F}^s \to \mathbb{F}^\ell$ is distance-preserving.

◇ By induction on the number of variables s.

Conclusion

Summary

Any polynomial generator is distance-preserving.

- Our proof covers all previously known distance-preserving generators, and leads to **improved parameters** and **proximity gaps**.

Conclusion

Summary

Any polynomial generator is distance-preserving.

- Our proof covers all previously known distance-preserving generators, and leads to **improved parameters** and **proximity gaps**.

Future work

- Larger proximity range Λ? Smaller error probability τ?
  - τ is sharp in some settings, e.g. $G(x) = (x^i)_{0 \le i < \ell}$ when $\delta < \delta_{\min}/2$.

Conclusion

Summary

Any polynomial generator is distance-preserving.

- Our proof covers all previously known distance-preserving generators, and leads to **improved parameters** and **proximity gaps**.

Future work

- Larger proximity range Λ? Smaller error probability τ?
  - τ is sharp in some settings, e.g. $G(x) = (x^i)_{0 \le i < \ell}$ when $\delta < \delta_{\min}/2$.
- New distance-preserving generators? (Yes [AGHP92])
- Are all biased generators also distance-preserving generators?

Thanks!