IOP-Fest, May 2026

The Proximity Prize: What it is and What is Currently Known

Dan Boneh

Joint work with **Gal Arnon** and **Giacomo Fenzi**

eprint.iacr.org/2026/680

Papers We Will Discuss

First proximity gaps for general linear codes: RVW'13, AHIV'17, RZ'18, BKS'18, BGKS'20

BCIKS'20: **Proximity gaps** for Reed-Solomon codes

ZCF'23: **BaseFold**: Efficient Polynomial Commitment Schemes from Foldable Codes

ACFY'24: **WHIR**: Reed–Solomon Proximity Testing with Super-Fast Verification

Hab'24: **BaseFold in the list-decoding regime**

Hab'25: A note on mutual correlated agreement for Reed-Solomon codes

Zei'24: Khatam: Reducing the Communication Complexity of Code-Based SNARKs

GCXK'25: From List-Decodability to Proximity Gaps

BCHKS'25: **On proximity gaps for Reed–Solomon codes**

GG'25: **Optimal proximity gaps for subspace-design codes and random RS codes**

CS'25: On Reed–Solomon proximity gaps conjectures

DG'25: On the distribution of the distances of random words

FS'25: Small-field hash-based SNARGs are less sound than conjectured

KKH'26: Failure of proximity gaps close to capacity

proxember 2025

CGHLLPS'26: S-two whitepaper

ABF'26: **Open Problems in List Decoding and Correlated Agreement**

$1,000,000

initiative by
ethereum foundation

in prizes to prove (or disprove!) Reed-Solomon proximity gaps conjectures

An initiative by the Ethereum Foundation to advance the foundations of modern zkVMs.

(initiated by **Justin Drake**)

PRIZE JUDGES

Dan Boneh
Stanford University

Giacomo Fenzi
EPFL

Gal Arnon
Bocconi University

What is the proximity challenge?

A question about polynomials over finite fields

The challenge: prove that a certain conjecture is true

- Only elementary algebra (no Adeles, Schemes, Sheaves, …)

Why is the EF paying $1M for a proof ??

⇒ if true, better (provable) proof systems based on Reed-Solomon

EF's Goal: avoid conjectures, as much as possible

The grand challenges

The grand MCA challenge. We are given a Reed–Solomon code $\mathcal{C} := \mathsf{RS}[\mathbb{F}, \mathcal{L}, k]$ defined over some smooth evaluation domain $\mathcal{L} \subseteq \mathbb{F}$. The code has constant rate, and in particular the rate $\rho(\mathcal{C}) := k/|\mathcal{L}|$ is one of $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}\}$.

$$\text{determine the largest } \delta_{\mathcal{C}}^* \in [0,1] \text{ such that } \varepsilon_{\mathsf{mca}}(\mathcal{C}, \delta_{\mathcal{C}}^*) \leq 1/2^{128},$$

assuming $|\mathbb{F}|$ is sufficiently large so that such a $\delta_{\mathcal{C}}^*$ exists.

The grand list decoding challenge. We are given a Reed–Solomon code $\mathcal{C}$ as in the grand MCA challenge. For a constant m,

$$\text{determine the largest } \delta_{\mathcal{C}}^* \in [0,1] \text{ such that } |\Lambda(\mathcal{C}^{\equiv m}, \delta_{\mathcal{C}}^*)| \leq |\mathbb{F}|/2^{128},$$

assuming $|\mathbb{F}|$ is sufficiently large so that such a $\delta_{\mathcal{C}}^*$ exists.

Why is this important? What does it even mean?

Notation

Let $\mathcal{C} = RS[\mathbb{F}, \mathcal{L}, k] \subseteq \mathbb{F}^{|\mathcal{L}|}$ be a **Reed-Solomon code**

with a **smooth eval. domain** $\mathcal{L}$ (i.e. $\mathcal{L}$ is a power-of-2 coset of a subgroup of $\mathbb{F}$)

and **rate** $\rho(\mathcal{C}) \coloneqq k/|\mathcal{L}| \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$

$\delta_{\min}(\mathcal{C}) \coloneqq \frac{\Delta_{\min}(\mathcal{C})}{|\mathcal{L}|} \in [0,1]$: the **minimum relative distance** of $\mathcal{C}$

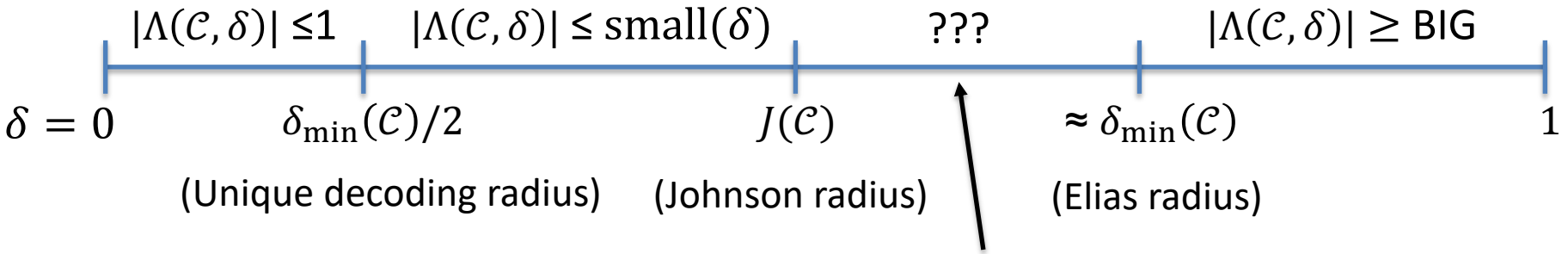
Recall: list decoding bounds for RS codes

The Johnson bound: For $\mathcal{C} \subseteq \mathbb{F}^n$ and $0 < \delta < J(\mathcal{C})$ $(= 1 - \sqrt{1 - \delta_{\min}(\mathcal{C})})$

$$\Lambda(\mathcal{C}, \delta) := \max_{w \in \mathbb{F}^n} |\{c \in \mathcal{C} : \Delta(w, c) \leq \delta\}| \leq \text{small}(\delta)$$

($\text{small}(\delta) \to \infty$ as δ approaches $J(\mathcal{C})$)

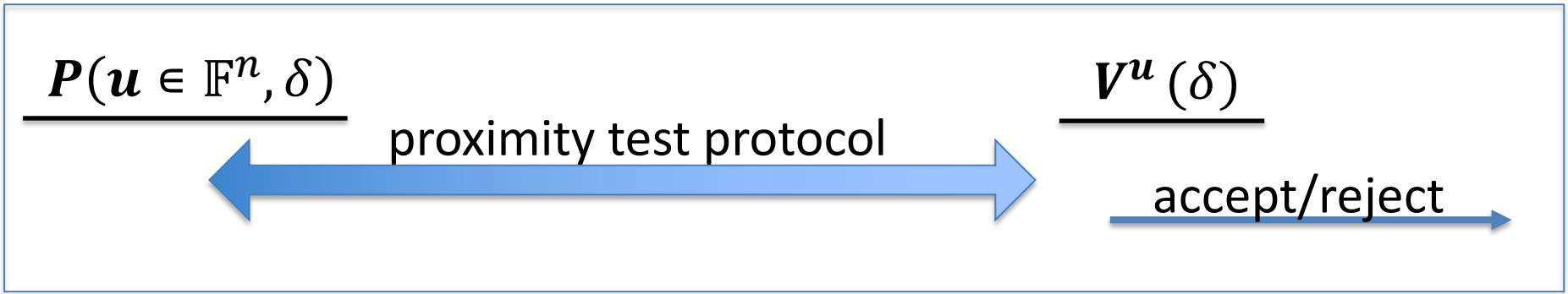
List decoding size $\Lambda(\mathcal{C}, \delta)$ as a function of $\delta \in [0,1]$:



Well behaved for Random-RS (RRS) and subspace-design codes, e.g. Folded-RS (FRS)

Proximity Testing: a SNARK ingredient

Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code, and $\delta \in [0, \delta_{\min}(\mathcal{C}))$.

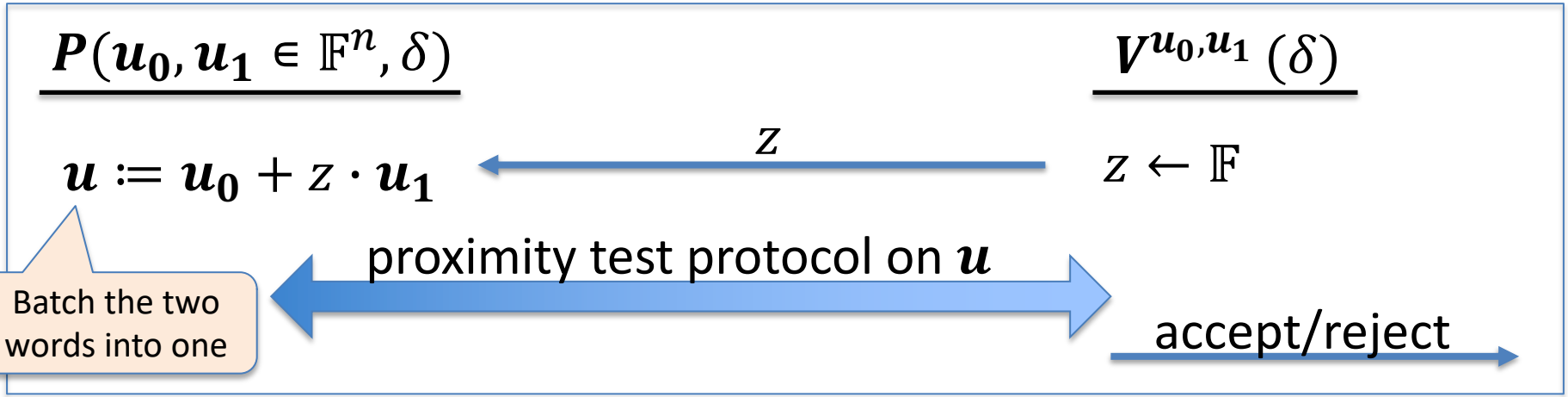


Completeness: if $u \in \mathcal{C}$ then Pr[Verifier accepts] = 1

Soundness: if u is δ-far from $\mathcal{C}$ then Pr[Verifier accepts] $< 1/2^{128}$

Batch proximity testing

Suppose we need two proximity tests. Can we batch them?



Complete: if $u_0, u_1 \in \mathcal{C}$ then $u \in \mathcal{C}$ so Pr[Verifier accepts] = 1

Is it sound? if one of u_0, u_1 is δ-far from $\mathcal{C}$, will V reject w.h.p ??

Answer: yes, if $\mathcal{C}$ has distance preservation.

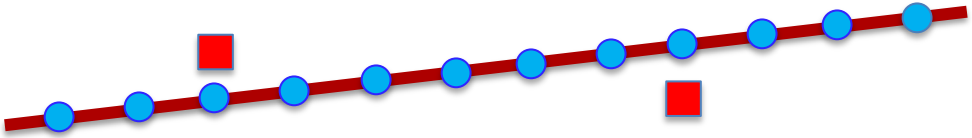
Distance preservation (proximity gap) [RVW'13]

Suppose one of $\boldsymbol{u_0}, \boldsymbol{u_1}$ is δ-far from $\mathcal{C}$.

Does a random linear combination preserve the distance?

Consider the line $L(z) = \boldsymbol{u_0} + z \cdot \boldsymbol{u_1}$

L:



we want at most few points on L to be δ-close to $\mathcal{C}$

Def: $p(L) := \Pr_{z \leftarrow \mathbb{F}}\left[\mathrm{dist}\left(L(z), \mathcal{C}\right) \leq \delta\right]$ (fraction of point on L that are δ-close to $\mathcal{C}$)

We say that $\mathcal{C}$ has **proximity gap** $\varepsilon_{\mathrm{pg}}(\mathcal{C}, \delta)$ if for every line L

$$p(L) \leq \varepsilon_{\mathrm{pg}}(\mathcal{C}, \delta) \quad \text{or} \quad p(L) = 1$$

Distance preservation (proximity gap)

For a given linear code $\mathcal{C}$,

what is the largest δ so that $\varepsilon_{\mathrm{pg}}(\mathcal{C}, \delta) < 1/2^{128}$??

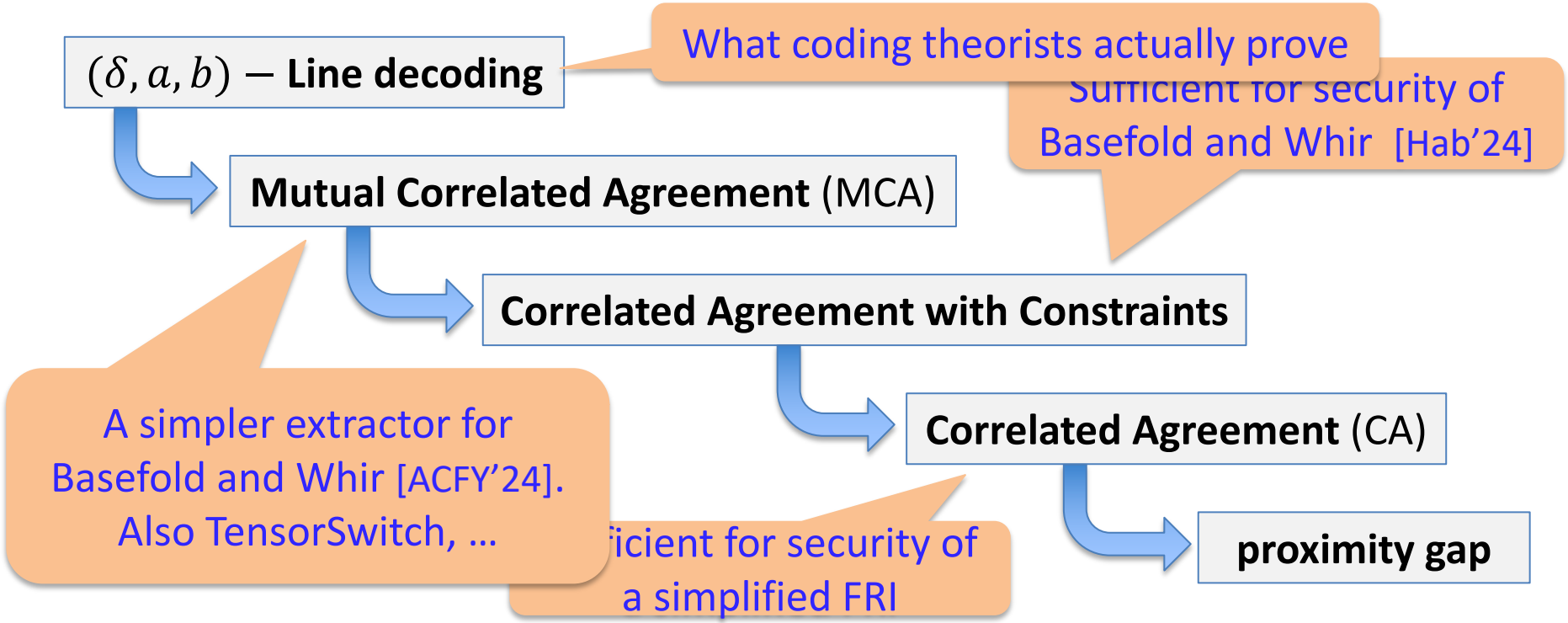
(a bigger δ results in a shorter SNARK proof --- as we will see)

A lot is known, but for for modern SNARKs

the basic proximity gap is insufficient … we need more.

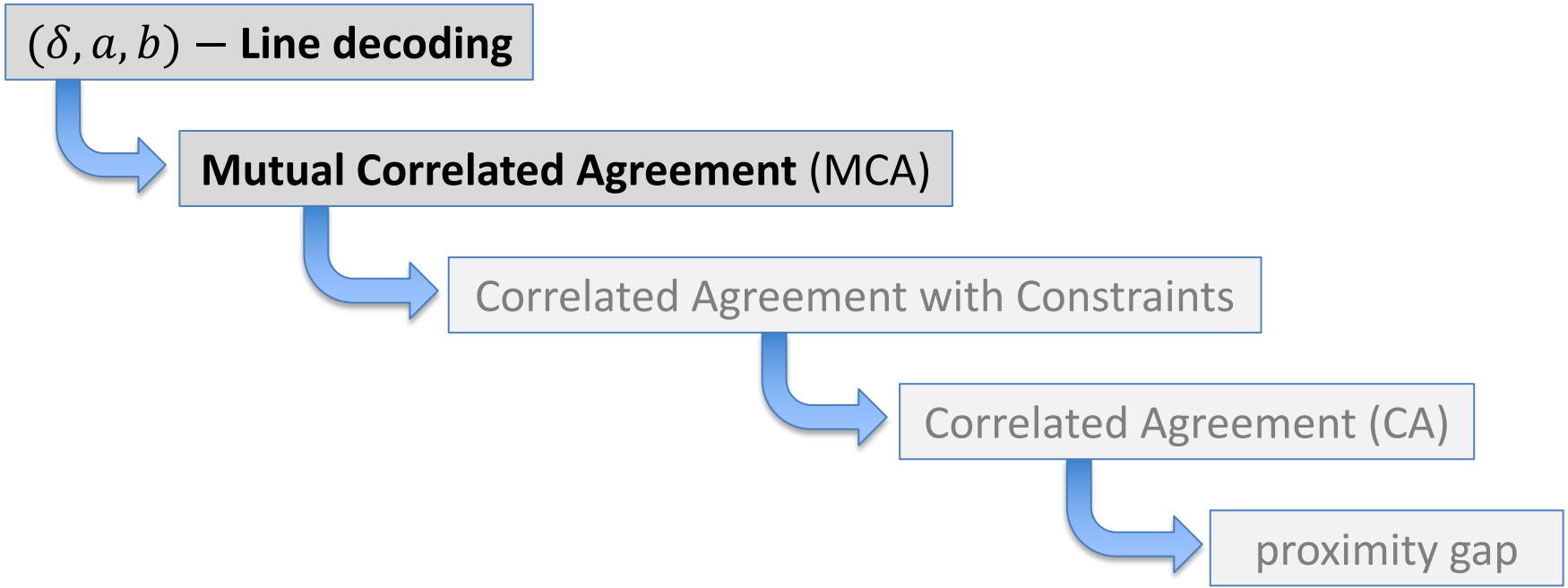
Five proximity gap notions for a linear code $\mathcal{C}$

Proximity gap notions for $\mathcal{C}$ at some distance $\delta \in [0, \delta_{\min}(\mathcal{C}))$



Five proximity gap notions for a linear code $\mathcal{C}$

Proximity gap notions for $\mathcal{C}$ at some distance $\delta \in [0, \delta_{\min}(\mathcal{C}))$

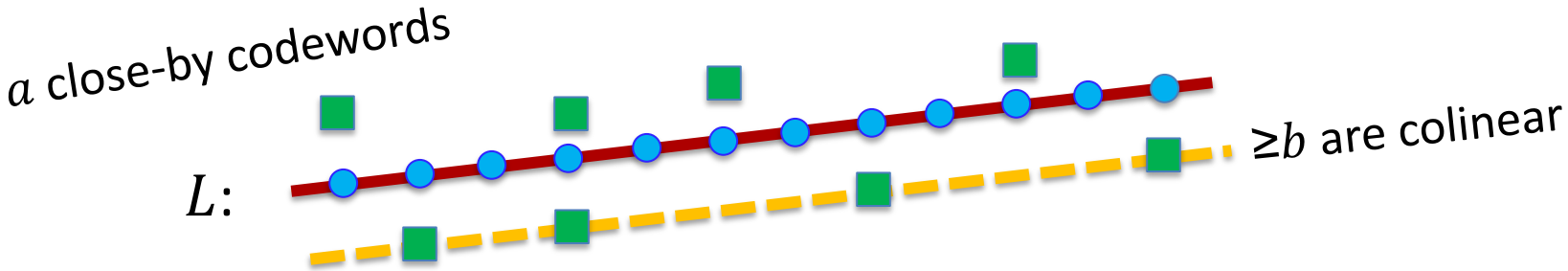


Our starting point: line decoding

Def: a linear code $\mathcal{C} \subseteq \mathbb{F}^n$ is (δ, a, b)-**line decodable** if for every line $L: \mathbb{F} \longrightarrow \mathbb{F}^n$ the following condition holds:

$\geq a$ points on L are δ-close to $\mathcal{C}$ $\Rightarrow$

≥b of the close-by codewords are colinear



(more generally, curve decodable)

GG'25: Goyal, Guruswami, eprint/2025/2054

CGH+'26: S-two paper, eprint/2026/532

Line decoding for Reed-Solomon Codes

Thm: [BCIKS'20, BCHKS'25, CGH+'26] Let $\mathcal{C} = RS[\mathbb{F}, \mathcal{L}, k] \subseteq \mathbb{F}^n$

Then $\mathcal{C}$ is (δ, a, b)-line decodable for $\delta \in \left(\frac{\delta_{min}(\mathcal{C})}{2}, J(\mathcal{C}) \right)$

unique decoding — Johnson bound

for $a \geq \eta(\delta)^5 \cdot n \ll |\mathbb{F}|$

and $b = \frac{a}{\eta(\delta)}$

where $\eta(\delta) \approx \max\left(\frac{1}{J(\mathcal{C})-\delta}, \text{ const}\right)$

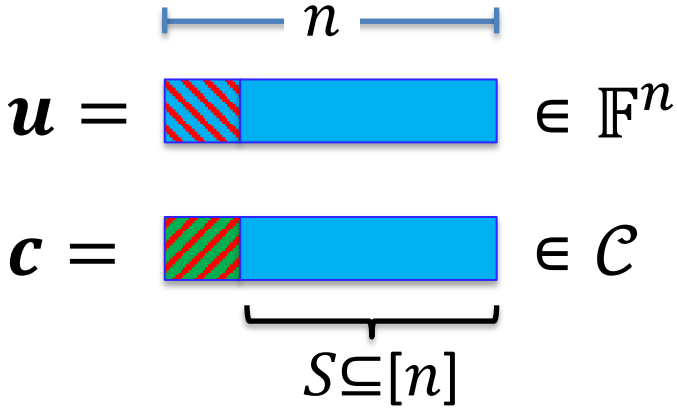
What about line decoding when $\delta > J(\mathcal{C})$? Good question!

CGH+'26: S-two paper, eprint/2026/532, Theorem 25

Mutual Correlated Agreement (MCA) [ACFY'25]

Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code, and $\delta \in [0, \delta_{\min}(\mathcal{C}))$.

Def: for $\boldsymbol{u} \in \mathbb{F}^n$ and $S \subseteq [n]$ we say that $\boldsymbol{u}$ **is** $(\boldsymbol{S}, \boldsymbol{\delta})$**-close to** $\boldsymbol{\mathcal{C}}$ if $|S| \geq (1-\delta)n$ and there is a $\boldsymbol{c} \in \mathcal{C}$ such that $\Delta_S(u, c) = 0$



$\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}^n$ are (S, δ)-close to $\mathcal{C}$

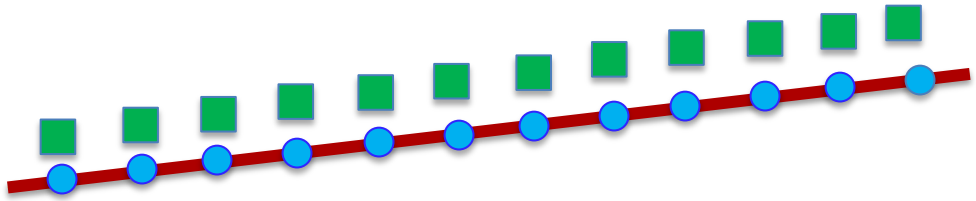
⇩

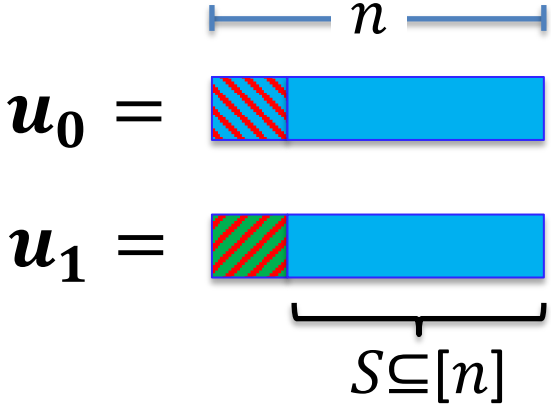
$\forall x, y \in \mathbb{F}$: $x \cdot \boldsymbol{u} + y \cdot \boldsymbol{v}$ is (S, δ)-close to $\mathcal{C}$

Mutual Correlated Agreement (MCA) [ACFY'25]

For $\boldsymbol{u_0}, \boldsymbol{u_1} \in \mathbb{F}^n$ consider the line $L(z) = \boldsymbol{u_0} + z \cdot \boldsymbol{u_1}$

L:





Clearly: if $\boldsymbol{u_0}, \boldsymbol{u_1}$ are (S, δ)-close to $\mathcal{C}$ then
$L(z)$ is (S, δ)-close to $\mathcal{C}$ for all $z \in \mathbb{F}$

MCA is the converse property:
If $L(z)$ is (S, δ)-close to $\mathcal{C}$ then w.h.p $\boldsymbol{u_0}, \boldsymbol{u_1}$ are (S, δ)-close to $\mathcal{C}$

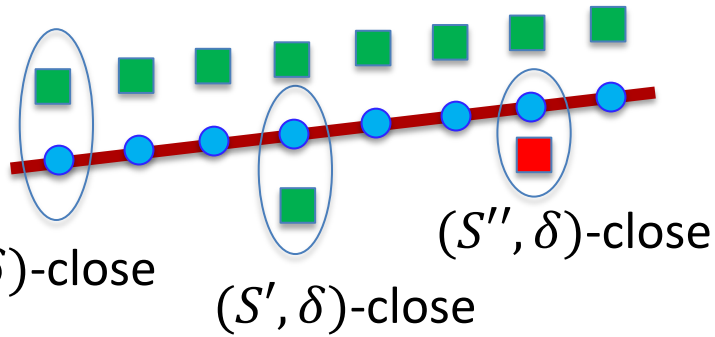
Mutual Correlated Agreement (MCA) [ACFY'25]

Def: $\mathcal{C} \subseteq \mathbb{F}^n$ has **Mutual Correlated Agreement** error $\varepsilon_{\text{mca}}(\mathcal{C}, \delta)$ if for every line $L(z) = \boldsymbol{u_0} + z \cdot \boldsymbol{u_1}$, $\boldsymbol{u_0}, \boldsymbol{u_1} \in \mathbb{F}^n$ we have:

for all $z \in \mathbb{F}$:

$L(z)$ is (S, δ)-close to $\mathcal{C}$ $\Rightarrow$ $\boldsymbol{u_0}, \boldsymbol{u_1}$ are (S, δ)-close to $\mathcal{C}$

except for $\leq \varepsilon_{\text{mca}}(\mathcal{C}, \delta)$ of the points on L (the MCA error)



$\Rightarrow$ $\boldsymbol{u_0}, \boldsymbol{u_1}$ are (S, δ)-close to $\mathcal{C}$

$\Rightarrow$ all $L(z)$ are (S, δ)-close to $\mathcal{C}$

$\Rightarrow$ $\boldsymbol{u_0}, \boldsymbol{u_1}$ are (S', δ)-close to $\mathcal{C}$

$\boldsymbol{u_0}, \boldsymbol{u_1}$ not (S'', δ)-close to $\mathcal{C}$ $\Rightarrow$ **MCA error**

From colinearity to agreement

Thm: [GG'25, Thm 3.5] for a linear code $\mathcal{C} \subseteq \mathbb{F}^n$

$$(\delta, a, n+1)\text{-line decodable} \quad \Rightarrow \quad \varepsilon_{\text{mca}}(\delta) < \frac{a}{|\mathbb{F}|}$$

Proof: Suppose some a points on the line $L(z) = u_0 + z\, u_1$ are MCA errors.

All these a points are δ-close to $\mathcal{C}$ $\Rightarrow$ $n+1$ of the close-by codewords are colinear

So: $\exists n+1$ points $\{\boldsymbol{p}_i \in L,\ (S_i, \delta)$-close to $\mathcal{C}\}$ and close-by codewords are colinear

For $n+1$ subsets of $[n]$ there must exist an i^* s.t. $S_{i^*} \subseteq \bigcup_{j \in [n+1] \setminus i^*} S_j$

By collinearity of close-by codewords: both $\boldsymbol{u_0}, \boldsymbol{u_1}$ are (S_{i^*}, δ)-close to $\mathcal{C}$.

… but then $\boldsymbol{p}_{i^*}$ is not an MCA error, contradiction! ■

The resulting MCA bound for RS codes

Let $\mathcal{C} = RS[\mathbb{F}, \mathcal{L}, k] \subseteq \mathbb{F}^{|\mathcal{L}|}$. For $\delta \in \left(\frac{\delta_{\min}(\mathcal{C})}{2}, J(\mathcal{C})\right)$

$$\varepsilon_{\text{mca}}(\mathcal{C}, \delta) < \frac{\eta(\delta)^5 \cdot |\mathcal{L}|}{|\mathbb{F}|} < 2^{-128}$$

for sufficiently large $|\mathbb{F}|$

What about ε_{mca} when $\delta > J(\mathcal{C})$???

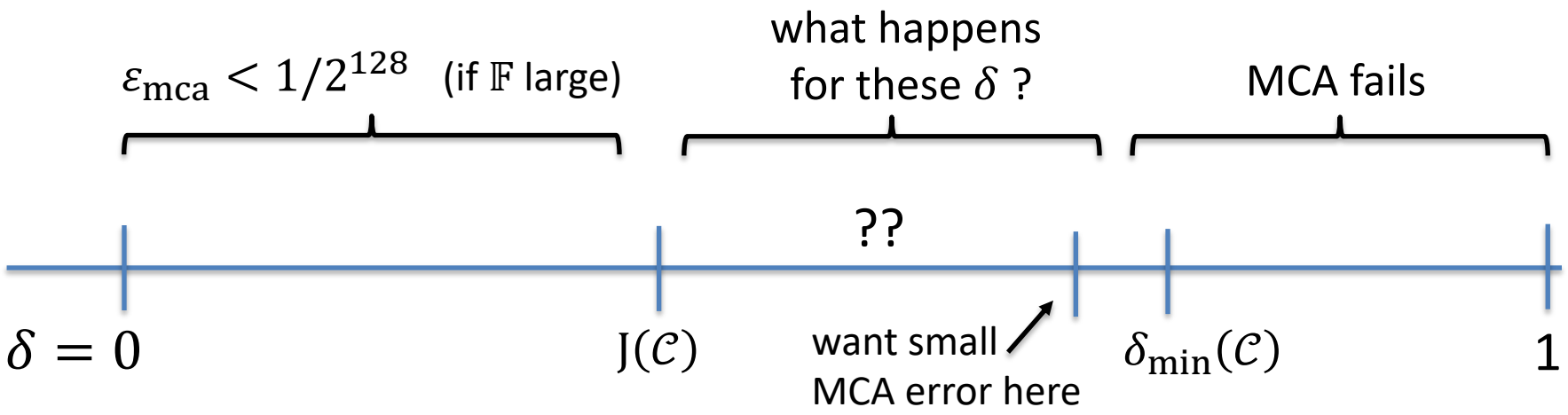
Habock'25: eprint/2025/2110

The Proximity Challenge for Constant Rate RS Codes

The grand MCA challenge. We are given a Reed–Solomon code $\mathcal{C} := \mathsf{RS}[\mathbb{F}, \mathcal{L}, k]$ defined over some smooth evaluation domain $\mathcal{L} \subseteq \mathbb{F}$. The code has constant rate, and in particular the rate $\rho(\mathcal{C}) := k/|\mathcal{L}|$ is one of $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}\}$.

determine the largest $\delta_{\mathcal{C}}^* \in [0, 1]$ such that $\varepsilon_{\mathsf{mca}}(\mathcal{C}, \delta_{\mathcal{C}}^*) \leq 1/2^{128}$,

assuming $|\mathbb{F}|$ is sufficiently large so that such a $\delta_{\mathcal{C}}^*$ exists.



Summary for RS codes over a smooth domain

distance δ	error $\varepsilon_{\mathsf{mca}}(\mathcal{C}, \delta)$	reference
$\delta = 0$	$\varepsilon_{\mathsf{mca}}(\mathcal{C}, \delta) = \frac{1}{\lvert\mathbb{F}\rvert}$	trivial
$\delta < \delta_{\mathsf{min}}(\mathcal{C})/2$ (δ below the unique decoding radius)	$\varepsilon_{\mathsf{mca}}(\mathcal{C}, \delta) \leq \frac{O(n)}{\lvert\mathbb{F}\rvert}$	[ACFY25; BCIKS20]
$\delta = J(\delta_{\mathsf{min}}(\mathcal{C})) - \eta$ (as δ approaches the Johnson radius)	$\varepsilon_{\mathsf{mca}}(\mathcal{C}, \delta) \leq \frac{n \cdot \mathrm{poly}(1/\eta)}{\lvert\mathbb{F}\rvert}$	[BCHKS25; Hab25; BCGM25; BCIKS20]

The behavior of $\varepsilon_{\mathsf{mca}}(\mathcal{C}, \delta)$ for $\delta \in [0, 1]$ and $\mathcal{C} := \mathsf{RS}[\mathbb{F}, \mathcal{L}, k]$, where $n := \lvert\mathcal{L}\rvert$.

The KKH negative result: MCA fails when δ is too big (also CS'25, DG'25, FS'25)

Krachun, Kazanin, Haböck: eprint/2026/782

Line decoding for RRS and FRS

Thm: [GG'25] The following codes are line decodable **almost up to capacity** ($\delta_{\min}(\mathcal{C})$)

- Subspace design codes, such as Folded Reed-Solomon (FRS)

  (see also JLR'26)

- The Random Reed-Solomon (RRS) code

How good are the resulting FRS-based SNARKs?

… Currently, worse than provable RS-based SNARKs due to a large alphabet

… Can one improve the alphabet size of FRS ?? (maybe not [BCDZ'25, Thm 1.11])

Why Mutual Correlated Agreement? (MCA)

What is it good for?

An example problem on a linear code $\mathcal{C} \subseteq \mathbb{F}^n$

$\boldsymbol{P(s, t_0, t_1;\ u_0, u_1;\ c_0, c_1)}$ $\qquad$ $\boldsymbol{V^{u_0,u_1}(s, t_0, t_1)}$

$$\boldsymbol{s, u_0, u_1} \in \mathbb{F}^n, \quad t_0, t_1 \in \mathbb{F}, \quad \boldsymbol{c_0, c_1} \in \mathcal{C}$$

Goal: convince the verifier that

- $\text{dist}(\boldsymbol{u_i, c_i}) \leq \delta$, and
- $\langle \boldsymbol{c_i, s} \rangle = t_i$ for $i = 0,1$

prove $\boldsymbol{u_0, u_1}$ are δ-close to codewords that satisfy a linear constraint

Captures intermediate steps in Basefold and Whir

A simple IOP for this relation

$\boldsymbol{P}(\boldsymbol{s}, t_0, t_1;\ \boldsymbol{u_0}, \boldsymbol{u_1};\ \boldsymbol{c_0}, \boldsymbol{c_1})$

$\boldsymbol{V}^{\boldsymbol{u_0}, \boldsymbol{u_1}}(\boldsymbol{s}, t_0, t_1)$

$z \in \mathbb{F}$

$z \leftarrow \mathbb{F}$

Honest prover:

$\boldsymbol{c} := \boldsymbol{c_0} + z \cdot \boldsymbol{c_1}$

$\boldsymbol{c} \in \mathbb{F}^n$

$x_1, \ldots, x_\ell \leftarrow [n]$

Batch the two codewords into one

spot checks that $\boldsymbol{c}$ was computed honestly

$\boldsymbol{V}$ accepts if $\boldsymbol{c} \in \mathcal{C}$ and

- $\langle \boldsymbol{c}, \boldsymbol{s} \rangle = t_0 + z \cdot t_1$,
- $\forall i \in [\ell]$: $\boldsymbol{c}(x_i) = \boldsymbol{u_0}(x_i) + z \cdot \boldsymbol{u_1}(x_i)$

Complete when $u_i = c_i$, but is it knowledge sound?

An extractor to prove knowledge soundness

Because of spot checks, we know that w.h.p $|S| \geq (1-\delta)n$

$\Rightarrow$ because $\boldsymbol{c} \in \mathcal{C}$, we have $\boldsymbol{u_0} + z \cdot \boldsymbol{u_1}$ is (S, δ)-close to $\mathcal{C}$

$\Rightarrow$ by MCA, w.h.p both $\boldsymbol{u_0}, \boldsymbol{u_1}$ are (S, δ)-close to $\mathcal{C}$

So: $\mathbf{Ext}$ knows error locations in $\boldsymbol{u_0}, \boldsymbol{u_1}$ and can efficiently find $\boldsymbol{c_0}, \boldsymbol{c_1}$

$\mathbf{Ext}\big((\boldsymbol{u_0}, \boldsymbol{u_1}), (z, \boldsymbol{c})\big)$ ← takes V's input and an accepting transcript

- Let $S \subseteq [n]$ be the largest set such that $\Delta_S(\boldsymbol{c},\ \boldsymbol{u_0} + z \cdot \boldsymbol{u_1}) = 0$
- Erasure correct $\boldsymbol{u_0}, \boldsymbol{u_1}$ with respect to S to get codewords $\boldsymbol{c_0}, \boldsymbol{c_1}$
- Output $\boldsymbol{c_0}, \boldsymbol{c_1}$

Extraction theorem

Thm: $\Pr\left[\begin{array}{l}\text{Ext fails to output a}\\ \text{valid witness } \boldsymbol{c_0}, \boldsymbol{c_1}\end{array}\right] \leq \max\left(\varepsilon_{\text{mca}}(\mathcal{C}, \delta) + \frac{\Lambda(\mathcal{C}^{\equiv 2}, \delta)}{|\mathbb{F}|},\ (1-\delta)^t\right)$

See paper for proof

Need MCA and list decoding bound to hold for large δ

bigger δ ⇒ smaller t ⇒ shorter proof

Note: we need both MCA and list-decoding bounds for $\delta > J(\mathcal{C})$

$1,000,000

initiative by
ethereum foundation

in prizes to prove (or disprove!) Reed-Solomon proximity gaps conjectures … bound the MCA error above the Johnson bound

An initiative by the Ethereum Foundation to advance the foundations of modern zkVMs.

PRIZE JUDGES

Dan Boneh
Stanford University

Giacomo Fenzi
EPFL

Gal Arnon
Bocconi University

Help us solve it …

$1,000,000

initiative by
ethereum foundation

in prize

proxi

bound

An initia

PRIZE J

What we don’t want:
AI slop on proximity that confuses incoming researchers

Help us solve it …

To learn more:

Open Problems in List Decoding and Correlated Agreement

Gal Arnon
galarnon42@gmail.com
Bocconi University

Dan Boneh
dabo@cs.stanford.edu
Stanford University

Giacomo Fenzi
giacomo.fenzi@epfl.ch
EPFL

eprint.iacr.org/2026/680

Lists many additional interesting questions in this area, such as:

- **Can we build shorter SNARKs from Folded-RS (FRS) ?**
- **Non-RS codes: Better MCA bounds for LDPC codes?**

(random LDPC have MCA to capacity [GG'25])