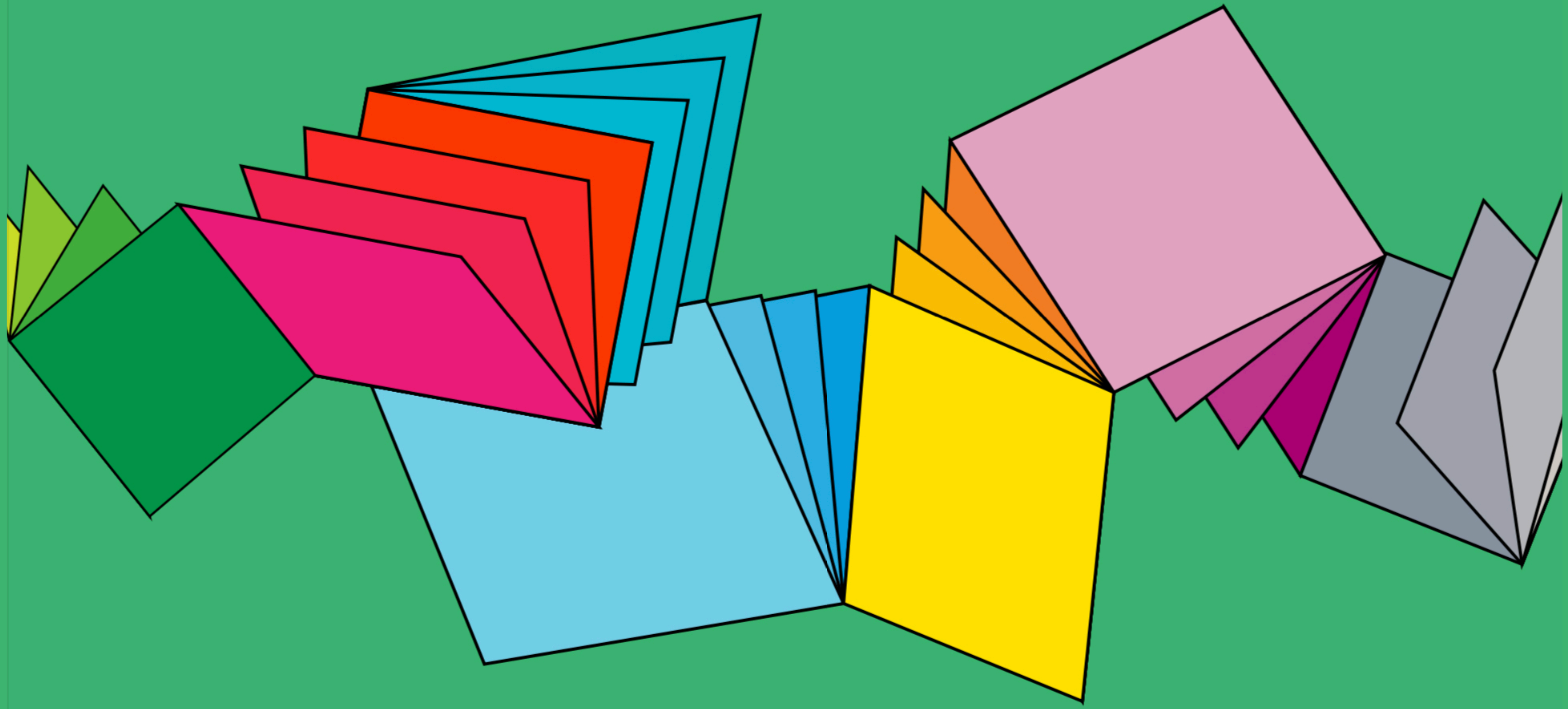
Recent Advances in Locally Testable Codes

c^3-LTC constructions



Simons Bootcamp

Prahladh Harsha

tifr

c^3 Locally Testable Codes

Theorem [Dinur-Evra-Livne-Lubotzky-Mozes and Pantaleev-Kalachev 2022]

For every $0 < r < 1$ there exist $\delta > 0$ and $q \in \mathbb{N}$ and an explicit construction of an infinite family of error-correcting codes $\{C_n\}_n$ with rate $\geq r$, distance $\geq \delta$ and locally testable with q queries.

c^3-LTC : Constant query, Constant fractional distance and constant rate

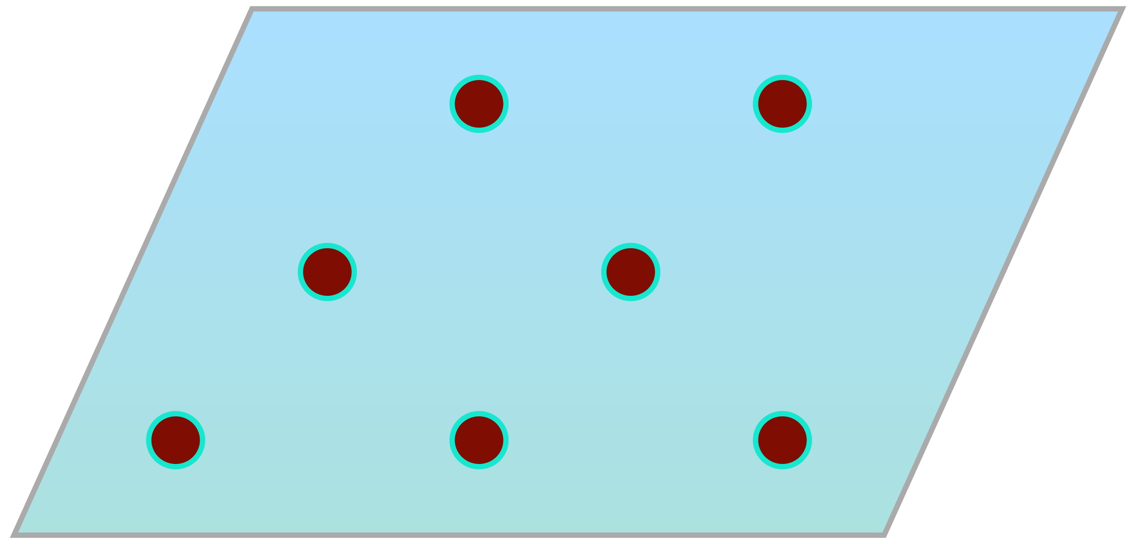
Talk Outline

1. Locally Testable Codes - quick recap
2. Existing Constructions (Hadamard, Reed-Muller, …)
3. Attempts at c^3-LTC construction
4. DELLM construction
   - Square Complex: Left-right Cayley complex
   - Code on the square complex
   - Proof Sketch of Testability

Locally Testable Codes

A linear error-correcting code is a linear subspace $C \subseteq \{0,1\}^n$

Rate = $\dfrac{dim(C)}{n}$, Distance = $min_{w \in C \setminus \{0\}} \dfrac{|\{i : w_i \neq 0\}|}{n}$



A code C is locally testable with q queries if there is a tester T that has query access to a given word w, reads q randomized bits from w and accepts / rejects, such that

- If $w \in C$ then Pr[T accepts] $= 1$
- If $w \notin C$ then Pr[T rejects] $\geq const \cdot dist(w, C)$

q = the locality of the tester

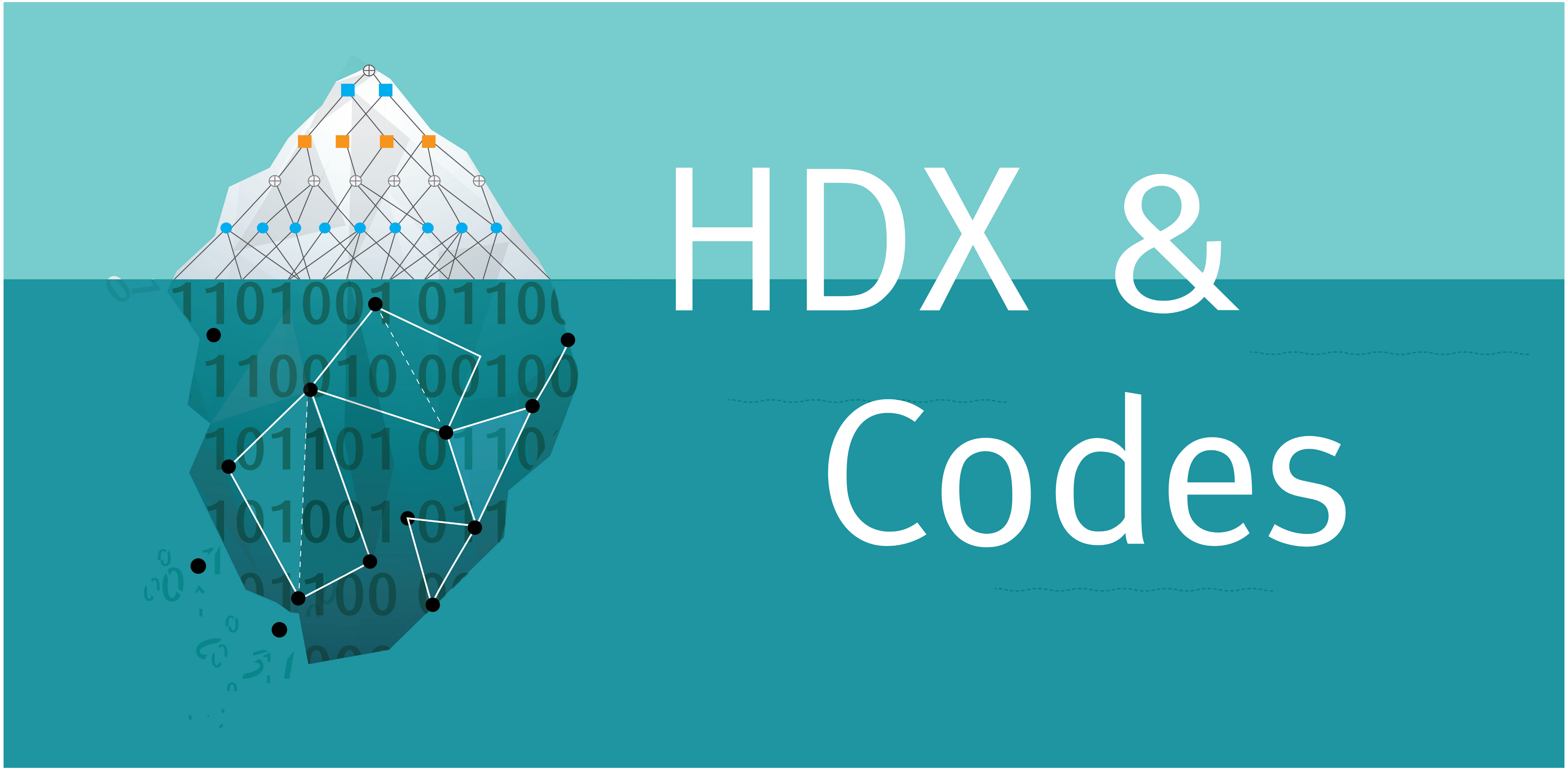
Historical background

- LTCs were studied implicitly in early PCP works [BlumLubyRubinfeld 1990, BabaiFortnowLund 1990, ..]
- Formally defined in works on low degree tests [Friedl-Sudan, Rubinfeld-Sudan] ~ 1995
- Spielman [1996 thesis]: useful in practice- can check “on the fly” if many errors occurred, and if so request re-transmission
- A systematic study initiated by Goldreich and Sudan in 2002.
  “what is the highest possible rate of an LTC?”

Historical background

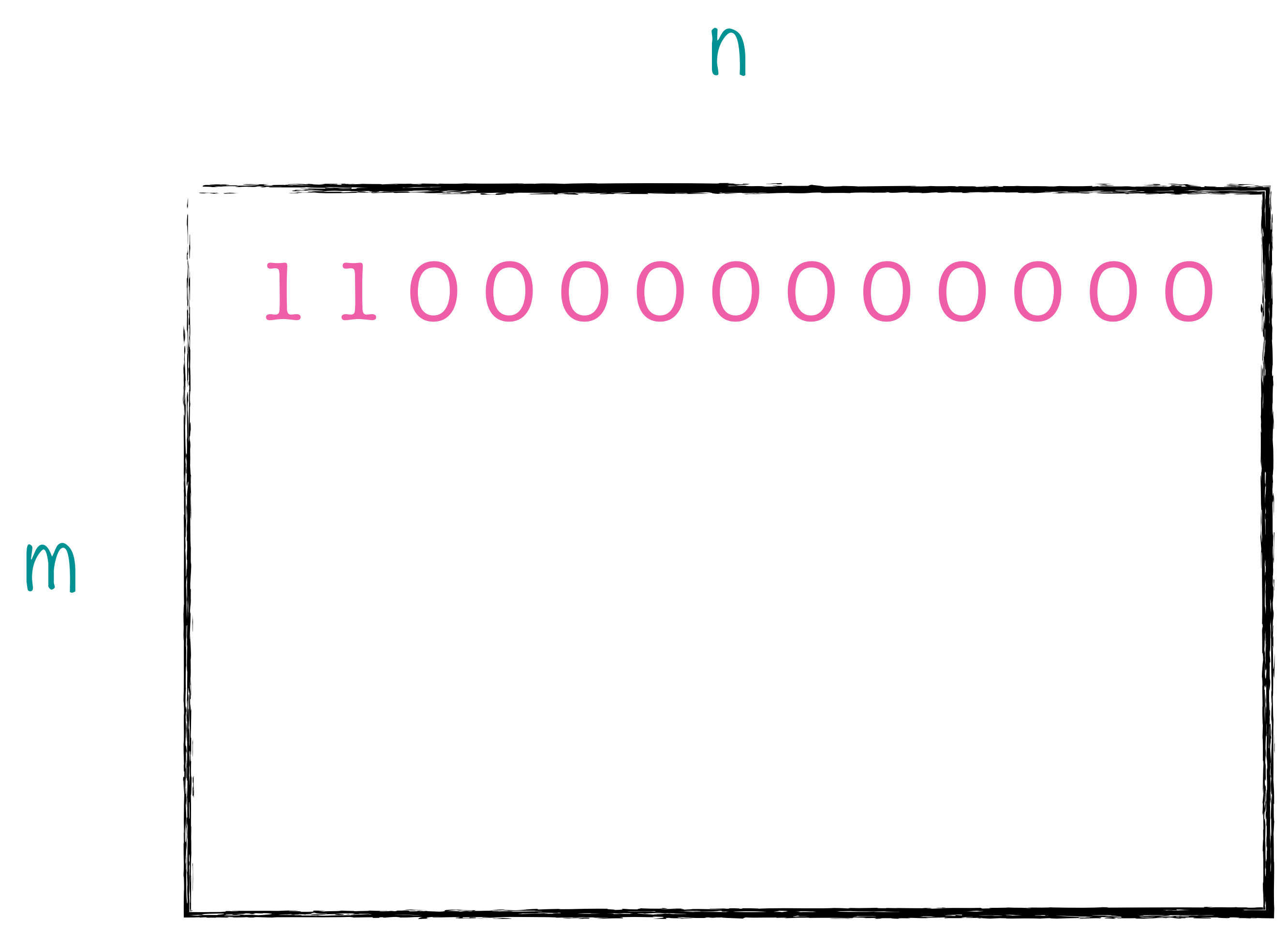
- Sequence of works (BenSasson-Sudan-Vadhan-Wigderson 2003, BenSasson-Goldreich-H.-Sudan-Vadhan 2004, Ben-Sasson-Sudan 2005, Dinur 2005) achieved rate = 1/polylog & constant locality+distance
- "c^3 LTCs" (constant rate, constant distance, constant locality) - experts doubt existence. Restricted lower bounds are shown [BenSasson-H-Rashkhodnikova 2003, Babai-Shpilka-Stefankovic 2005, BenSasson-Guruswami-Kaufman-Sudan-Viderman 2010, Dinur-Kaufman2011]
- Fix rate to constant, get locality $(\log n)^{\log\log n}$: [Kopparty-Meir-RonZewi-Saraf 2017, Gopi-Kopparty-Oliveira-RonZewi-Saraf 2018] (forget about PCPs, inject expanders)
- Affine invariance [Kaufman-Sudan 2007,…]: what makes properties testable?
- High dimensional expansion: local to global features [Garland 1973, Kaufman-Lubotzky 2013, Kaufman-Kazhdan-Lubotzky 2014, Evra-Kaufman 2016, Oppenheim 2017, Dinur-Kaufman 2017, Dinur-H.-Kaufman-LivniNavon-TaShma 2019, Dikstein-Dinur-H.-Kaufman-RonZewi 2019, Anari-Liu-OveisGharan-Vinzant 2019]

We even had a summer cluster at the Simons Institute in 2019



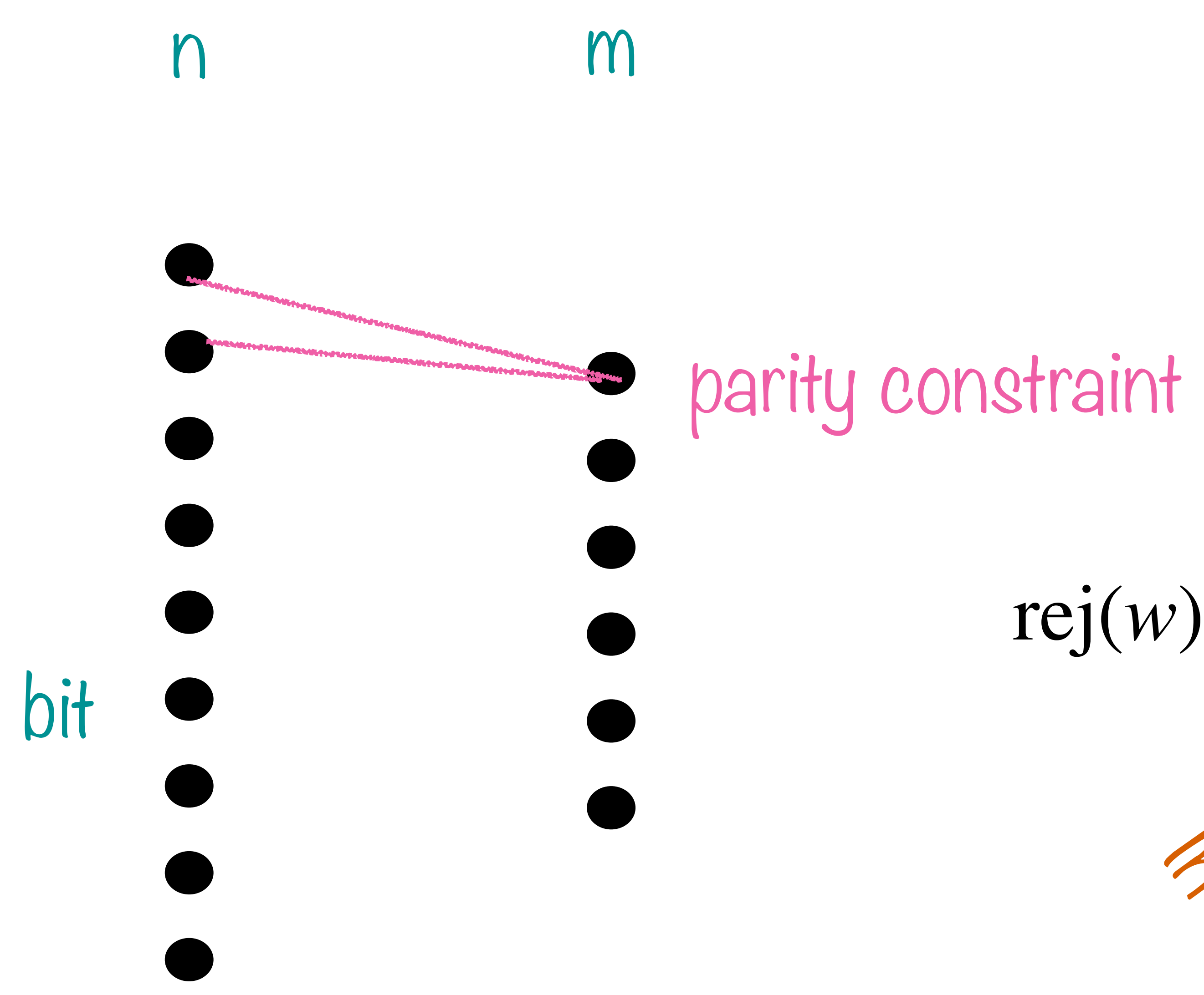
Low density parity check (LDPC) codes [Gallager '1963]

A (linear) locally testable code is necessarily an LDPC



H - parity check matrix

$$\mathscr{C} = \mathrm{Ker}(H) = \{w \in \{0,1\}^n : Hw = 0\}$$



factor graph

$$\mathrm{rej}(w) \geq \rho \cdot \mathrm{dist}(w, \mathscr{C})$$

$$\mathscr{C} = \{w \in \{0,1\}^n : \forall v \in [m], \textstyle\sum_{i \sim v} w_i = 0 \mod 2\}$$

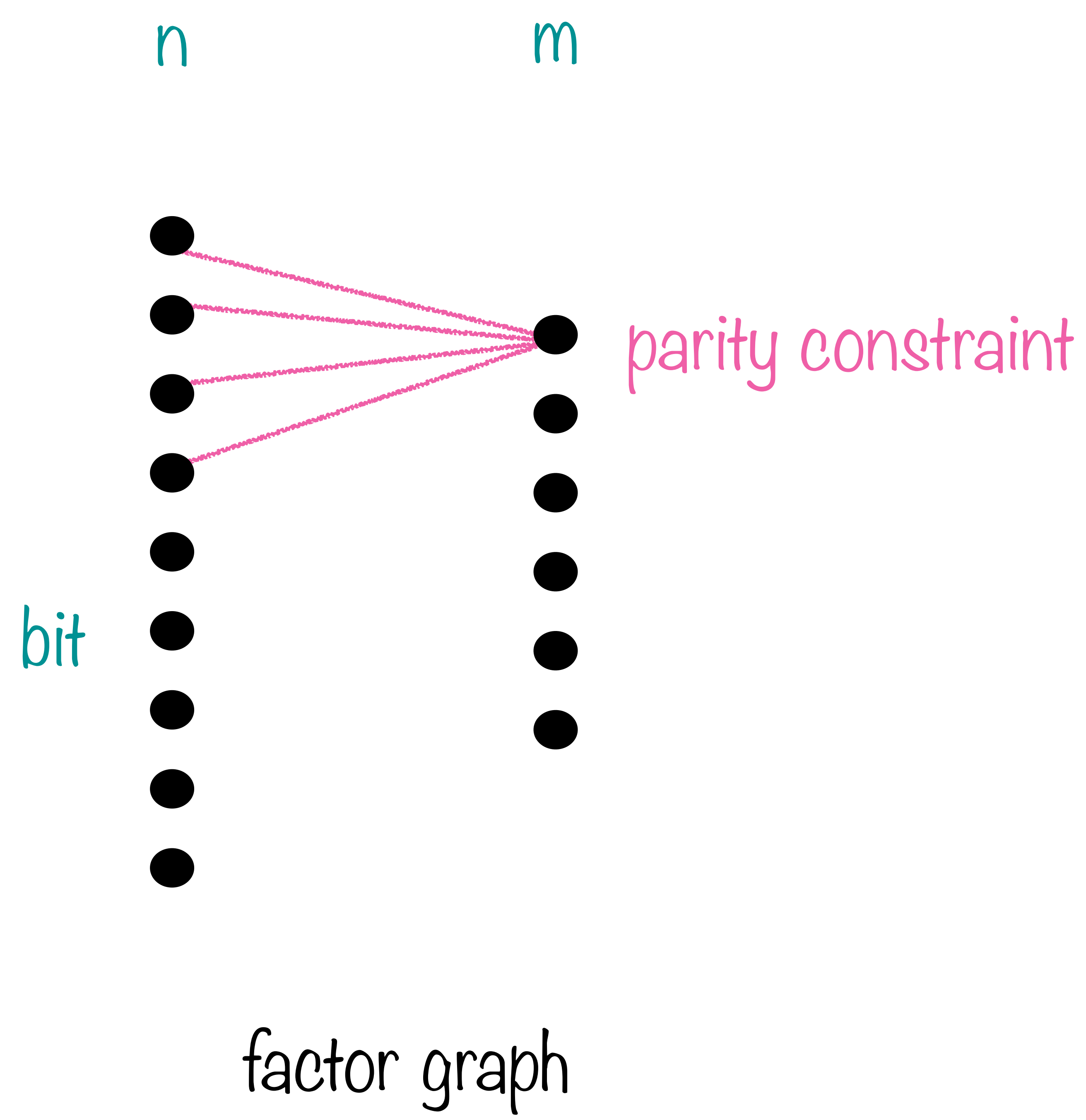
Two measures for a word w ∈ {0,1}ⁿ

1. $\mathrm{dist}(w, \mathscr{C})$ - distance to closest codeword
2. $\mathrm{rej}(w)$ - fraction of rejecting constraints

In an LTC these measures are related!

Expander Codes

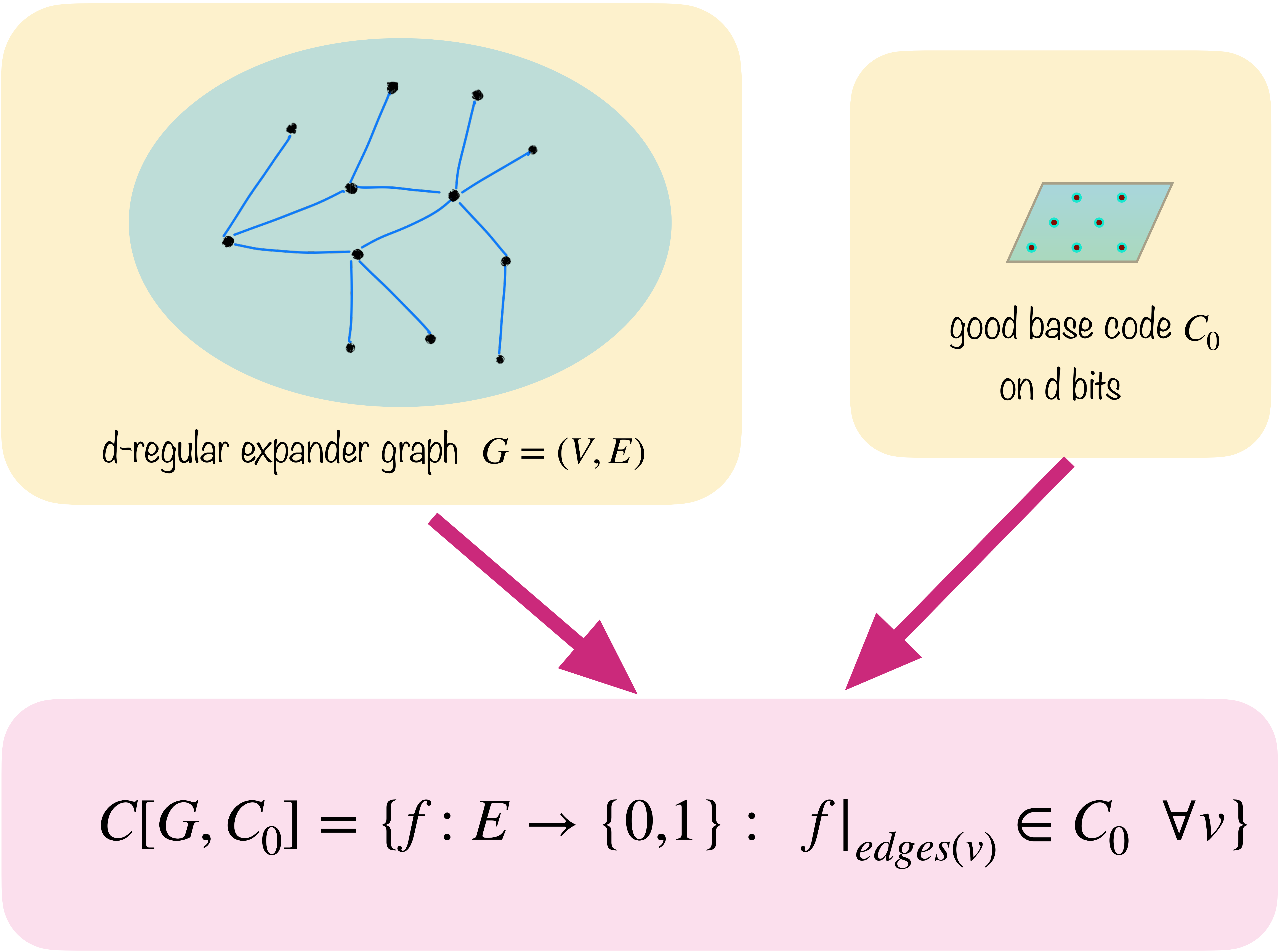
- Gallager (1963): A random LDPC code has good rate & distance
- Tanner (1981): Place a small base-code $C_0 \subseteq \{0,1\}^d$ on each constraint node. Consider various bipartite graph structures
- Sipser & Spielman (1996): Explicit expander-codes: Tanner codes using edges of an (explicit) expander



$$\mathscr{C} = \{w \in \{0,1\}^n : \forall v \in [m], \textstyle\sum_{i \sim v} w_i = 0 \mod 2\}$$

$$\mathscr{C} = \{w \in \{0,1\}^n : \forall v \in [m], w|_{\text{nbrs}(v)} \in \mathscr{C}_0\}$$

Expander Codes [Sipser & Spielman 1996]



d-regular expander graph $G = (V, E)$

good base code C_0 on d bits

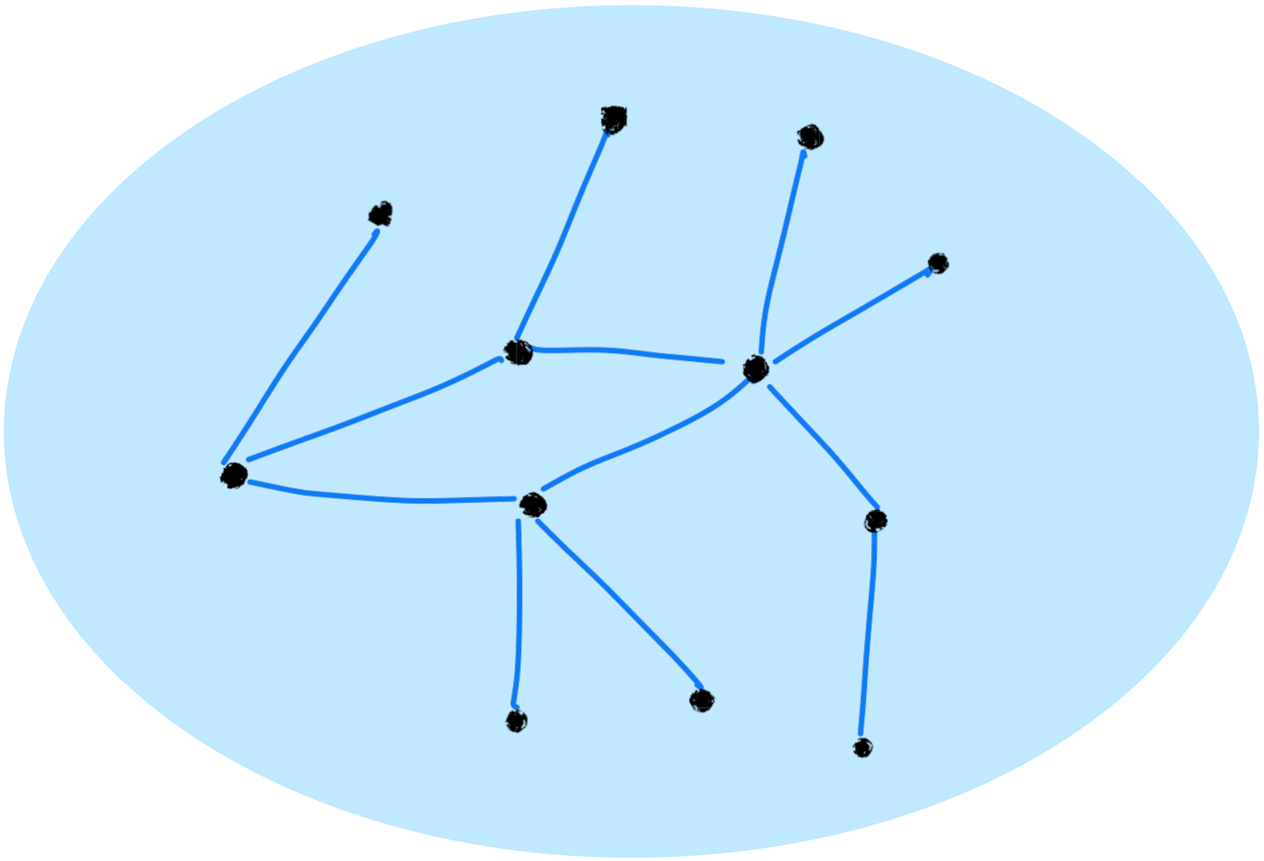
$$C[G, C_0] = \{f : E \to \{0,1\} : \ f|_{edges(v)} \in C_0 \ \ \forall v\}$$

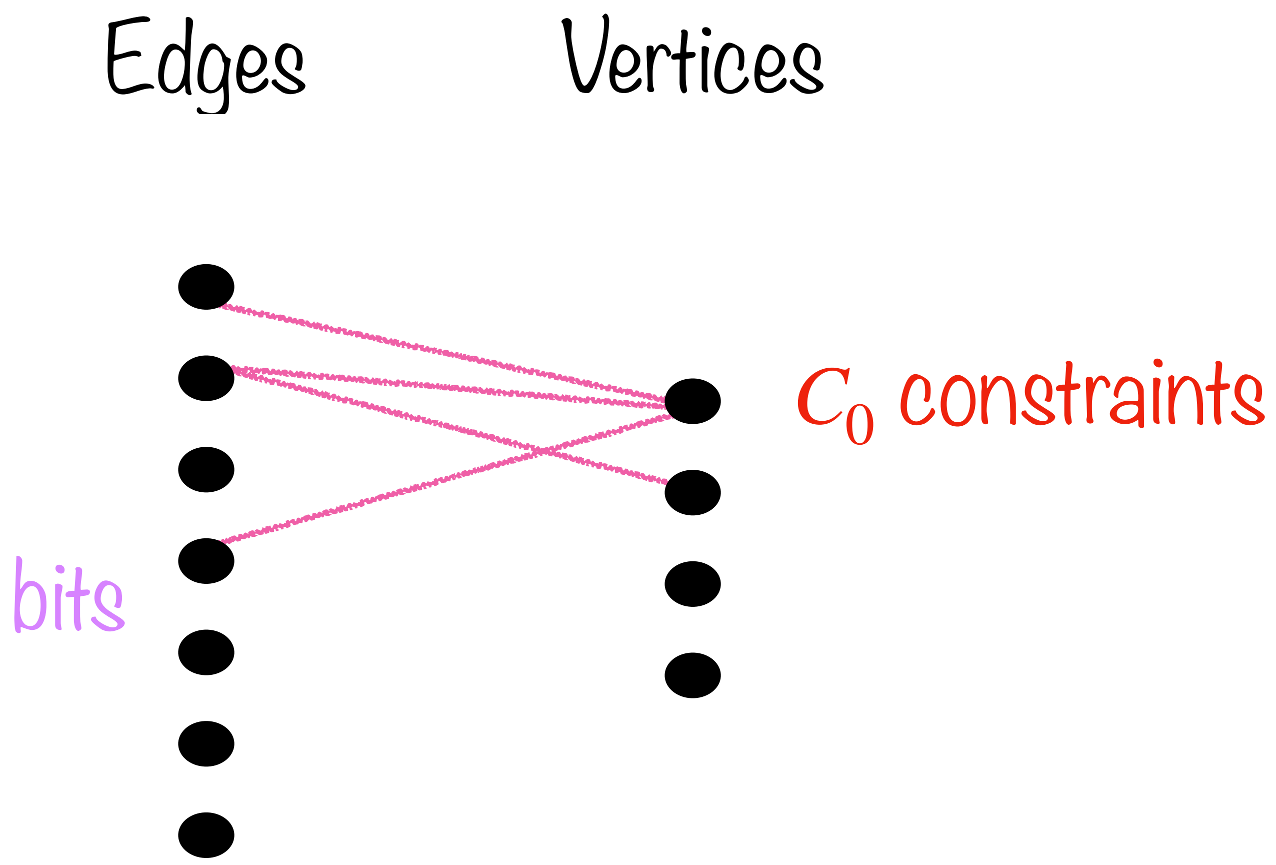
Expander Codes [Sipser & Spielman 1996]

Given

1. A d-regular $\lambda-$expander graph G on n vertices
2. A base code $C_0 \subseteq \{0,1\}^d$ with rate r_0, distance δ_0

Let $C[G, C_0] = \{f : E \to \{0,1\} : \forall v, f|_{edges(v)} \in C_0\}$



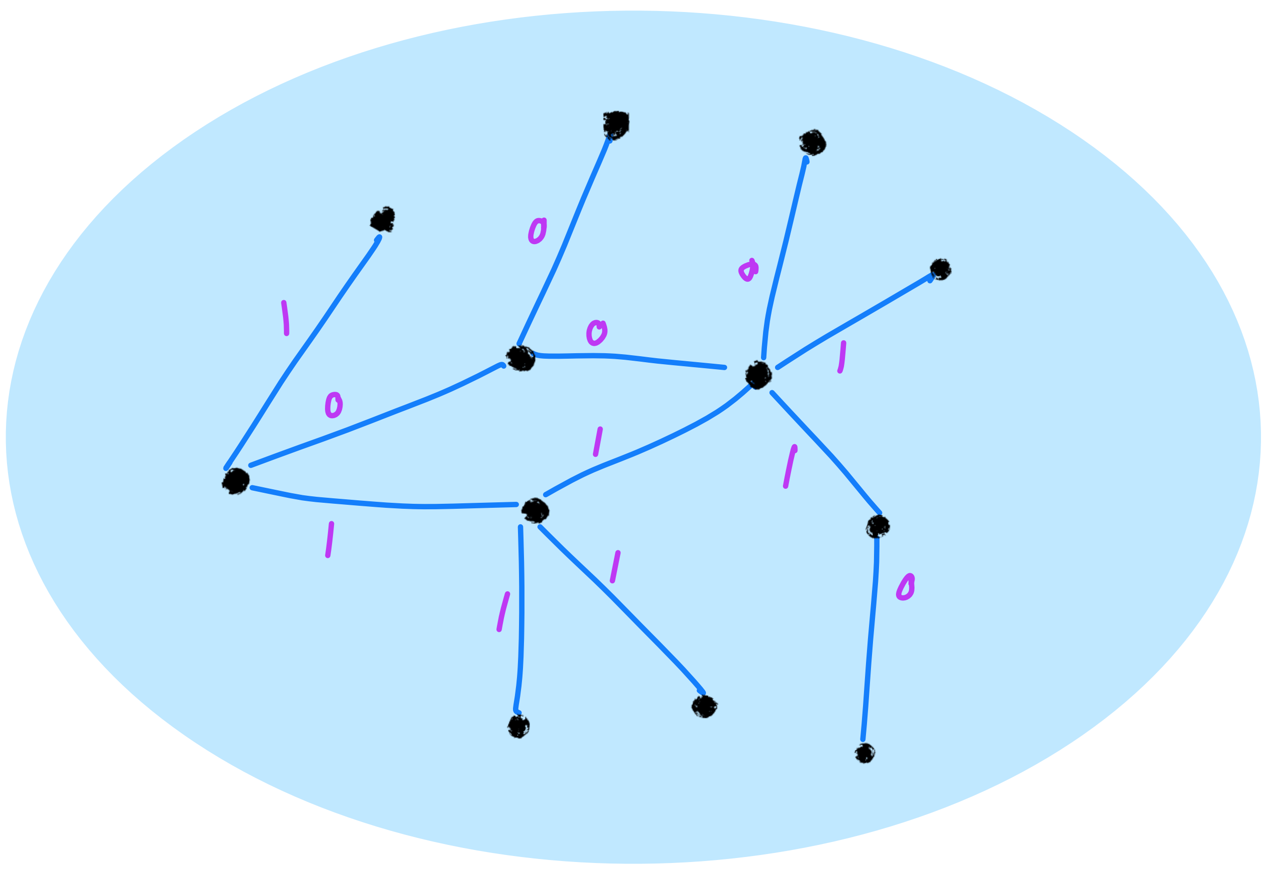


Expander Codes [Sipser & Spielman 1996]

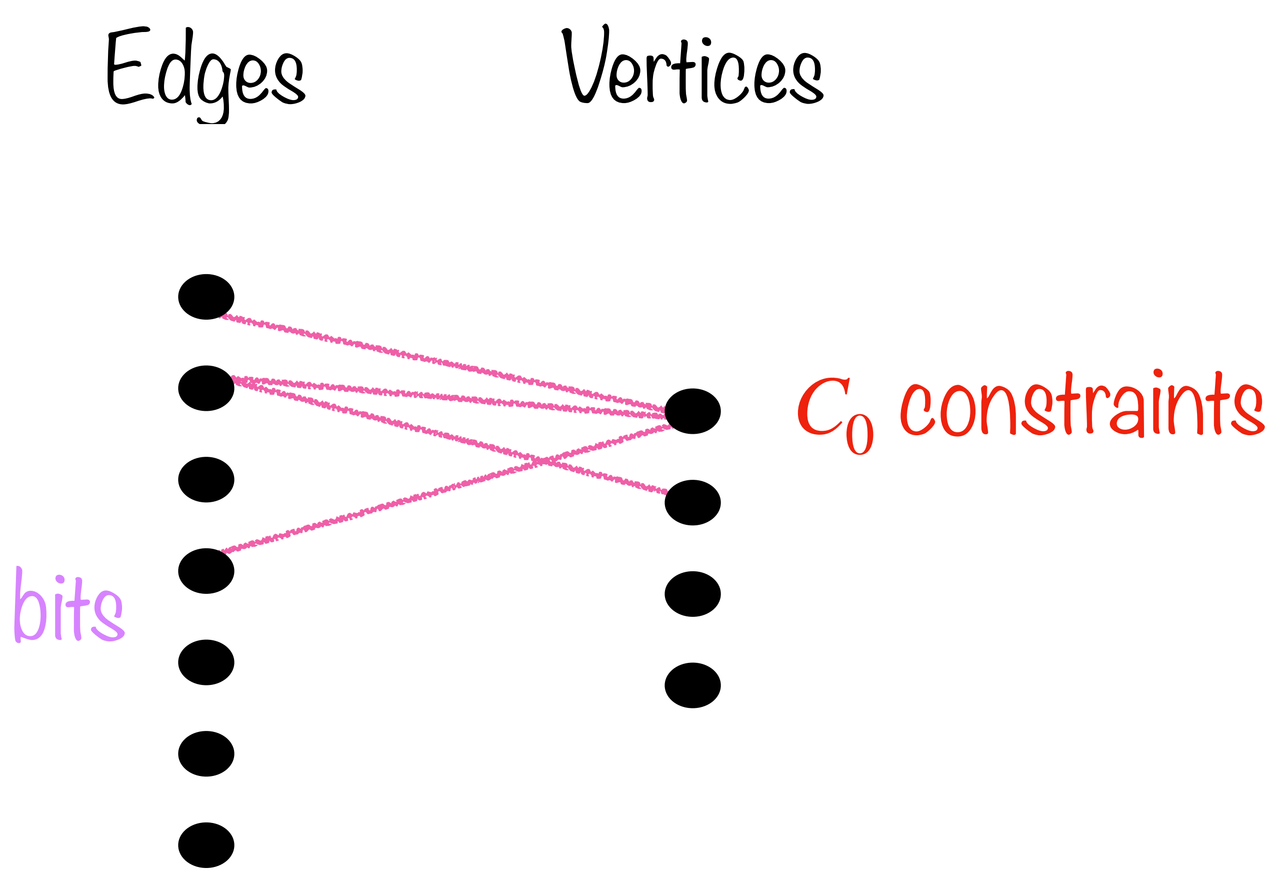
Given

1. A d-regular $\lambda-$expander graph G on n vertices
2. A base code $C_0 \subseteq \{0,1\}^d$ with rate r_0, distance δ_0

Let $C[G, C_0] = \{f : E \to \{0,1\} : \forall v, f|_{edges(v)} \in C_0\}$



- Dim(C) $\geq$ #bits - #constraints = $|E| - |V| \cdot (1 - r_0)d = |E|(2r_0 - 1)$ rate positive if $r_0 > 1/2$
- Distance $\geq \delta_0(\delta_0 - \lambda)$
- Linear time decoding !
- Locally testable?

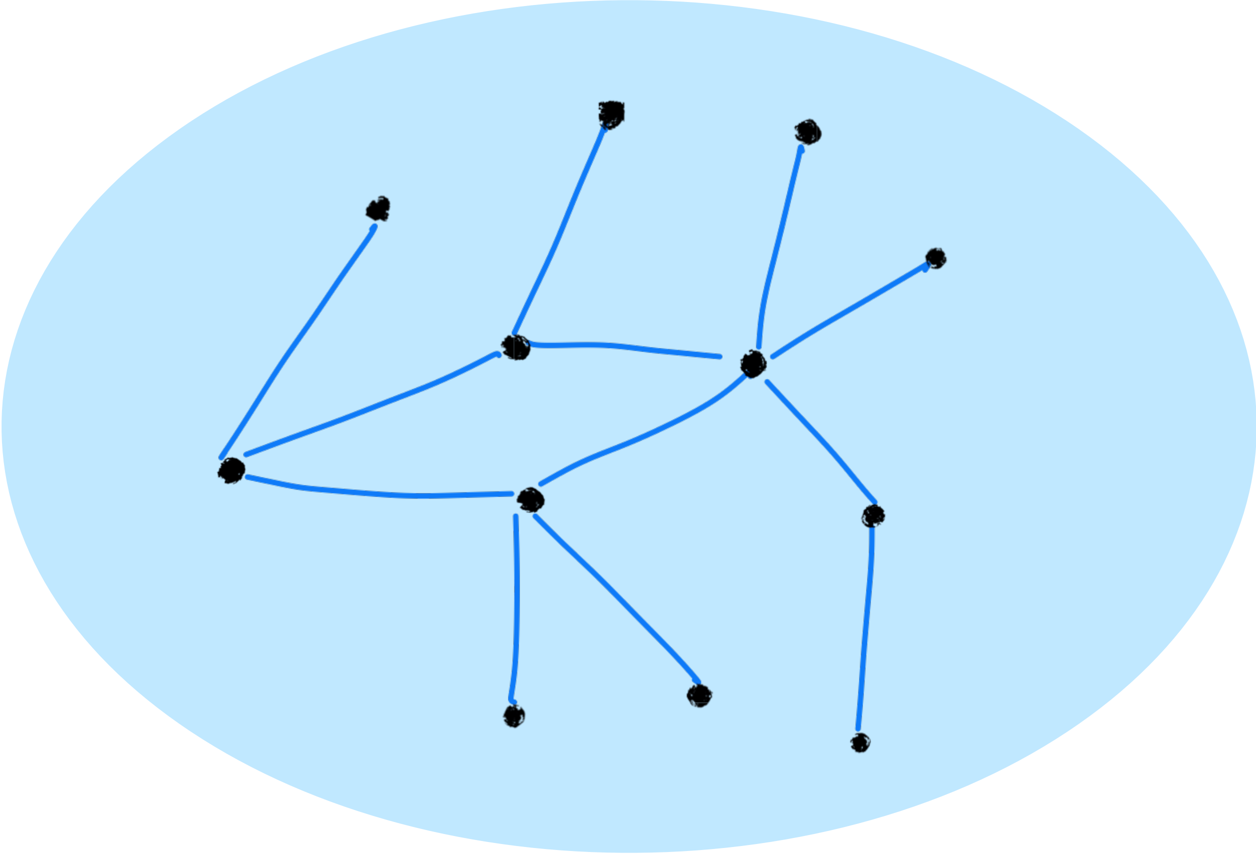


Expander Codes [BenSasson-H.-Raskhodnikova '03]

are typically not locally testable

Expander codes often have a word $w \notin C$ that is both

- Far from the code: $dist(f, C) > const$
- Rejected by only 1 constraint $\rho(f) = 1/|V|$



Proof:

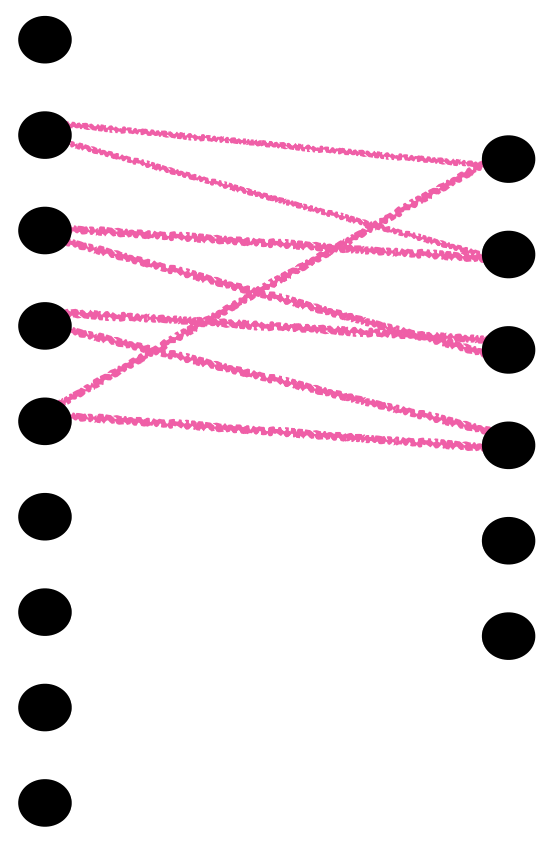
Choose v_0 and remove one constraint from the base-code of v_0

New codewords are far from old code, but violate only one constraint

Other LTCs

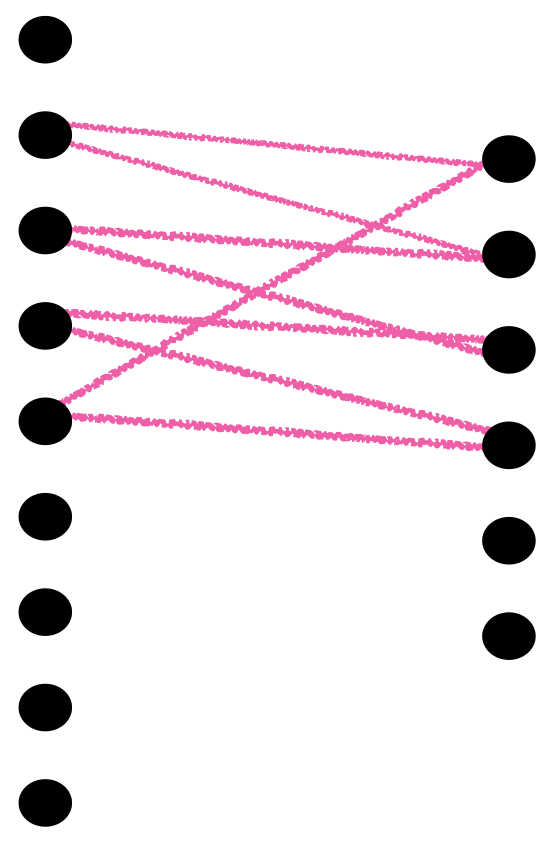
- Hadamard Codes [Blum-Luby-Rubinfeld 1990,…]
- Reed-Muller Codes
  - Large fields [Rubinfeld-Sudan 1992,…]
  - Small fields [Alon-Kaufman-Krivilevich-Litsyn-Ron 2003]

Hadamard Code as Tanner Code



bits C_0 constraints

factor graph

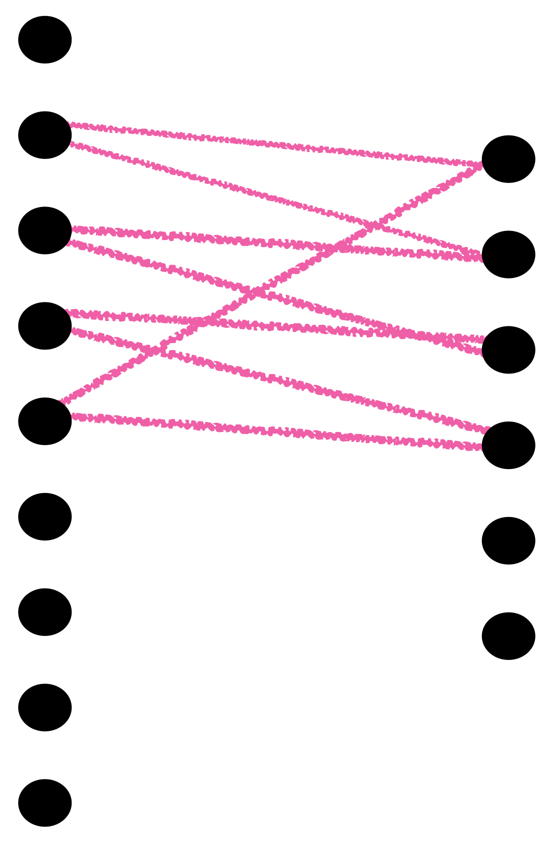


$\{0,1\}^n$
Codeword bits

Triples $(x, y, x+y)$
Constraints

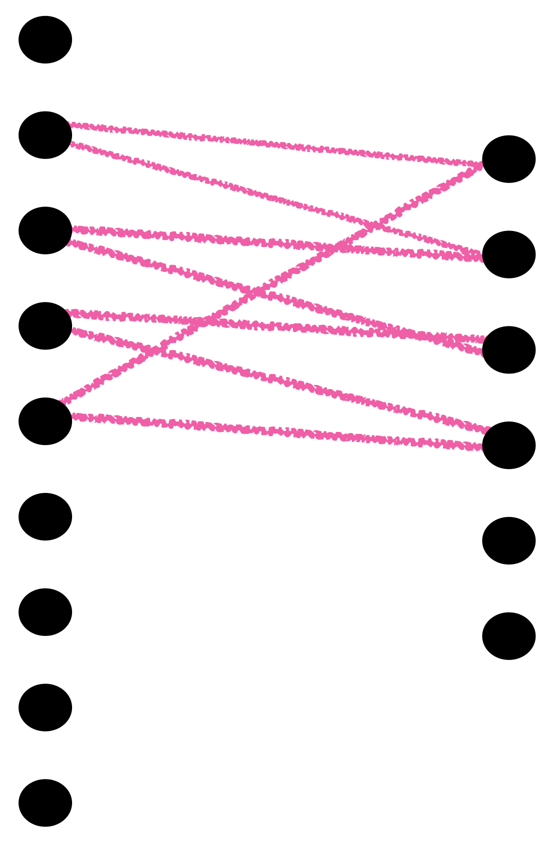
factor graph

Reed-Muller Code as Tanner Code



bits C_0 constraints

factor graph



$\mathbb{F}^m$ Affine Lines

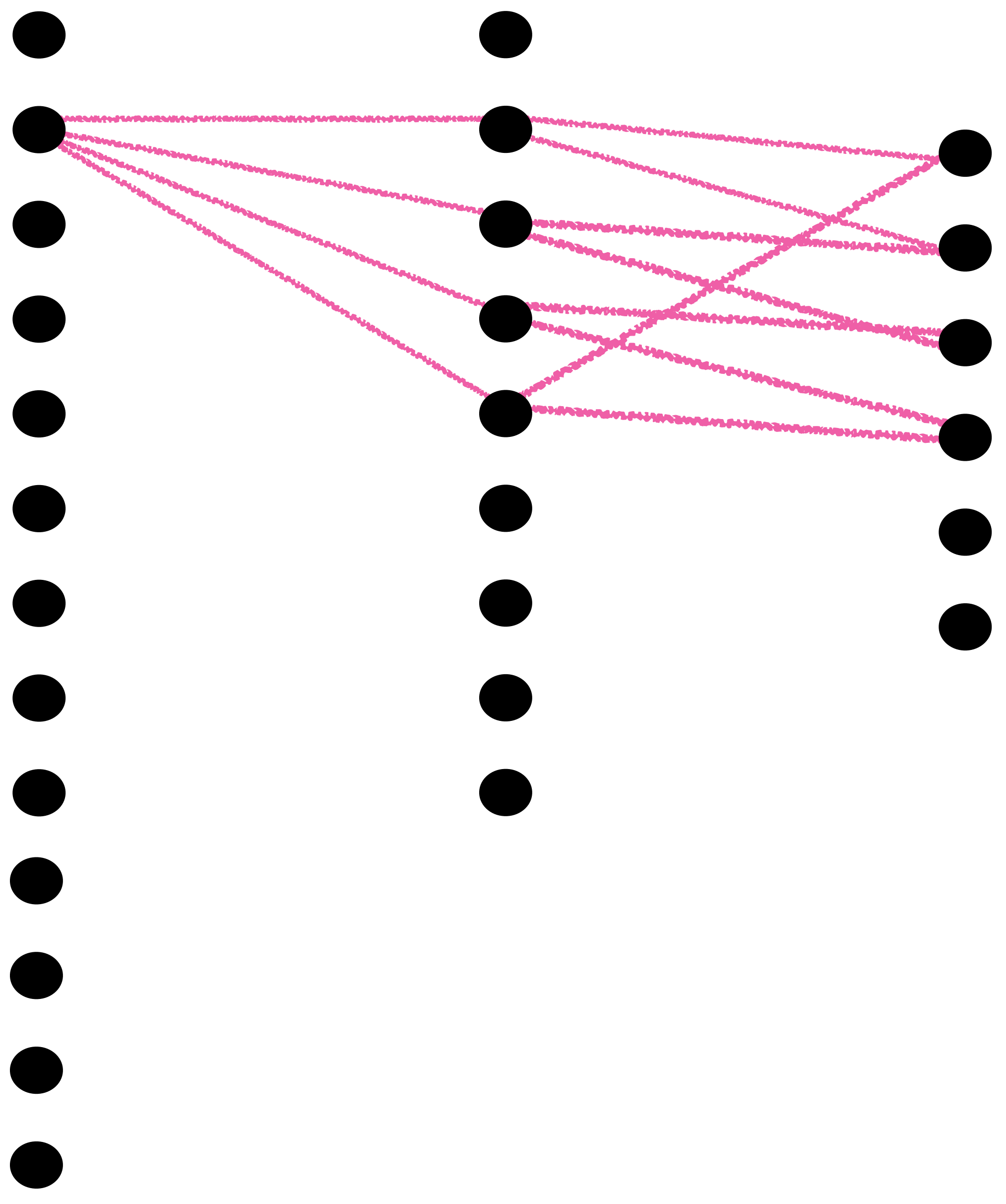
Codeword bits Constraints

factor graph

What makes Hadamard and RM codes testable?

- Hadamard Codes [Blum-Luby-Rubinfeld 1990,…]
- Reed-Muller Codes
  - Large fields [Rubinfeld-Sudan 1992,…]
  - Small fields [Alon-Kaufman-Krivilevich-Litsyn-Ron 2003]

Testability of Hadamard Code

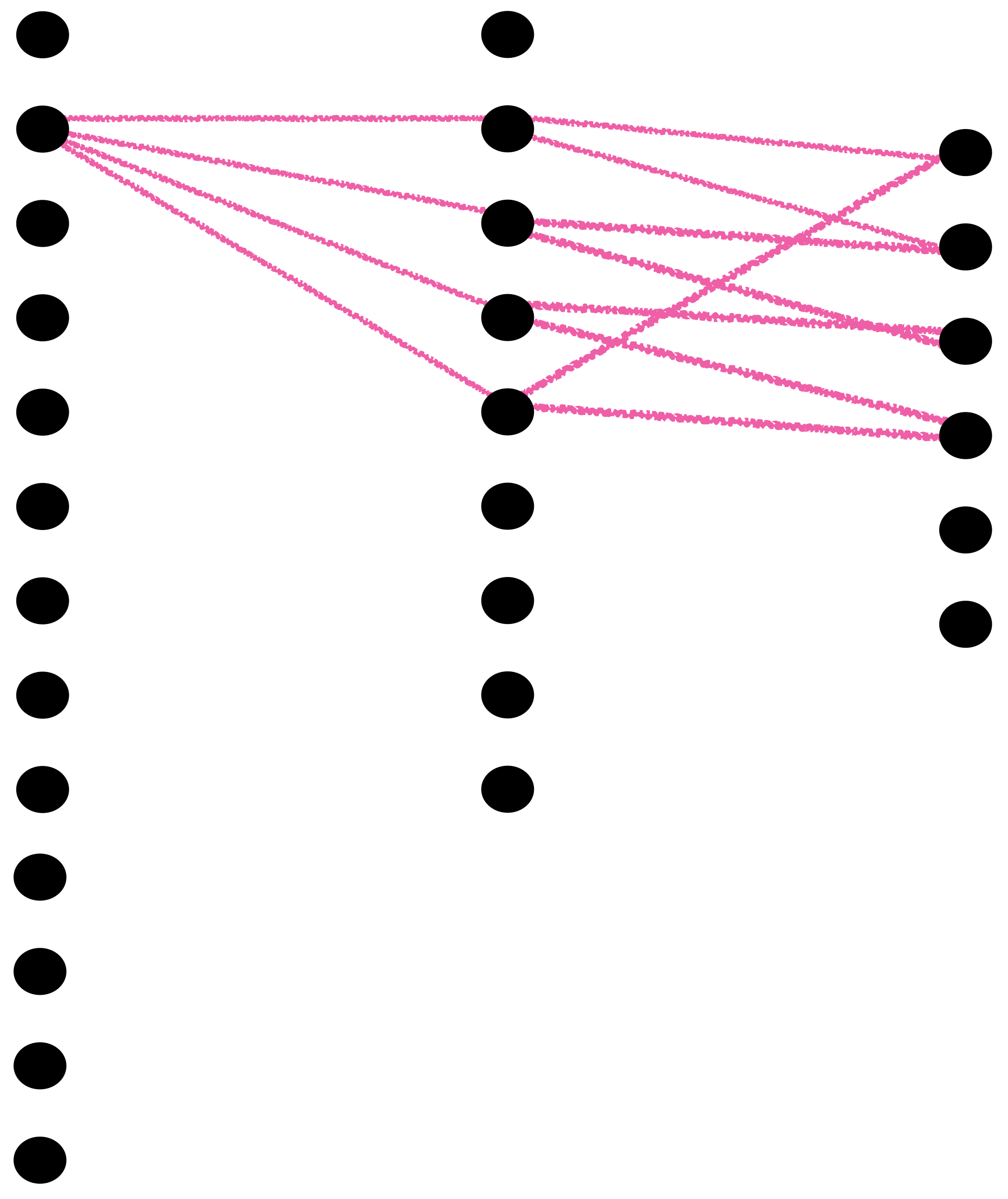


bits

C_0 constraints

dependencies

3-layered factor graph



$\{0,1\}^n$

Codeword bits

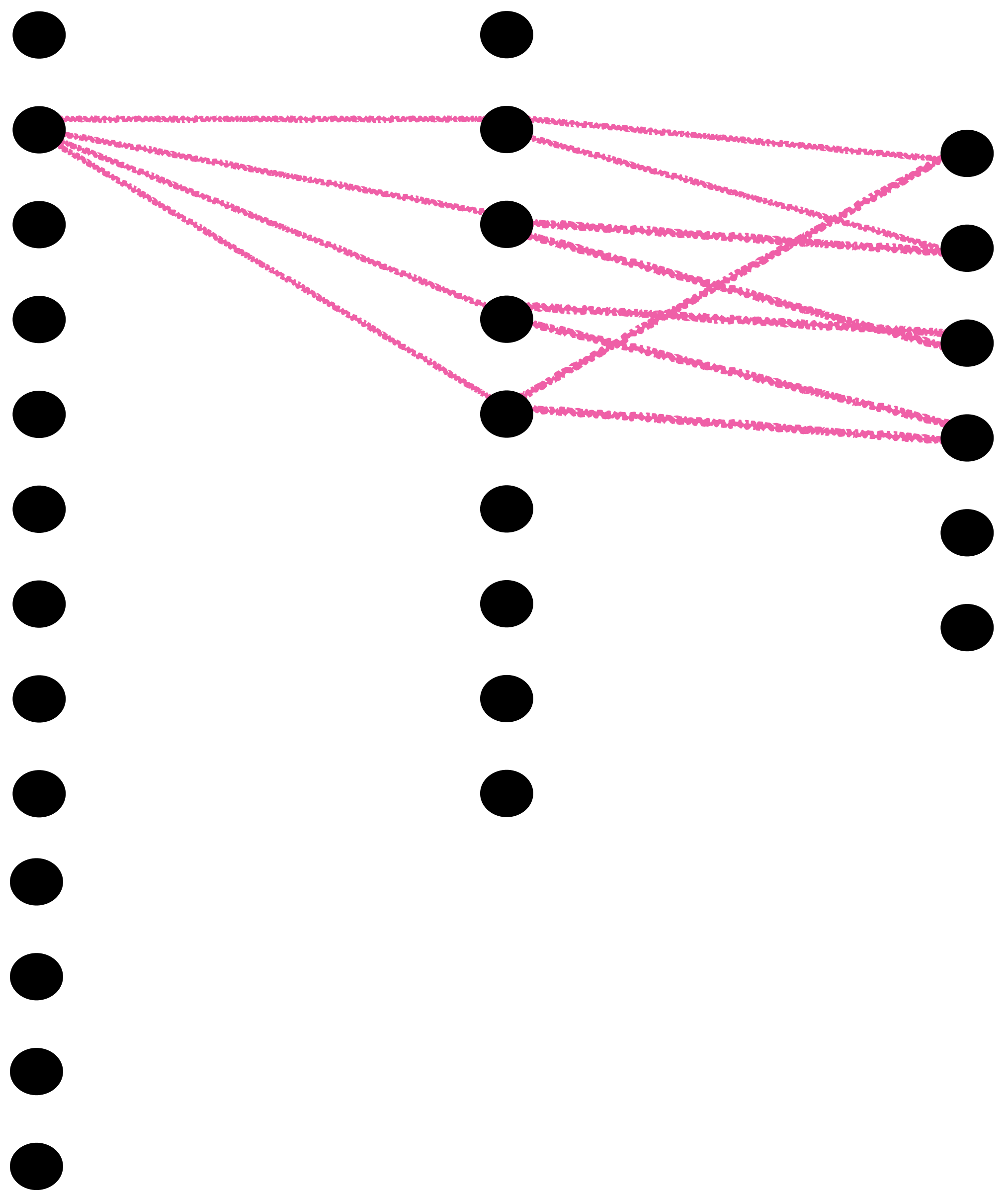
Triples $(x, y, x + y)$

Constraints

Magic Squares

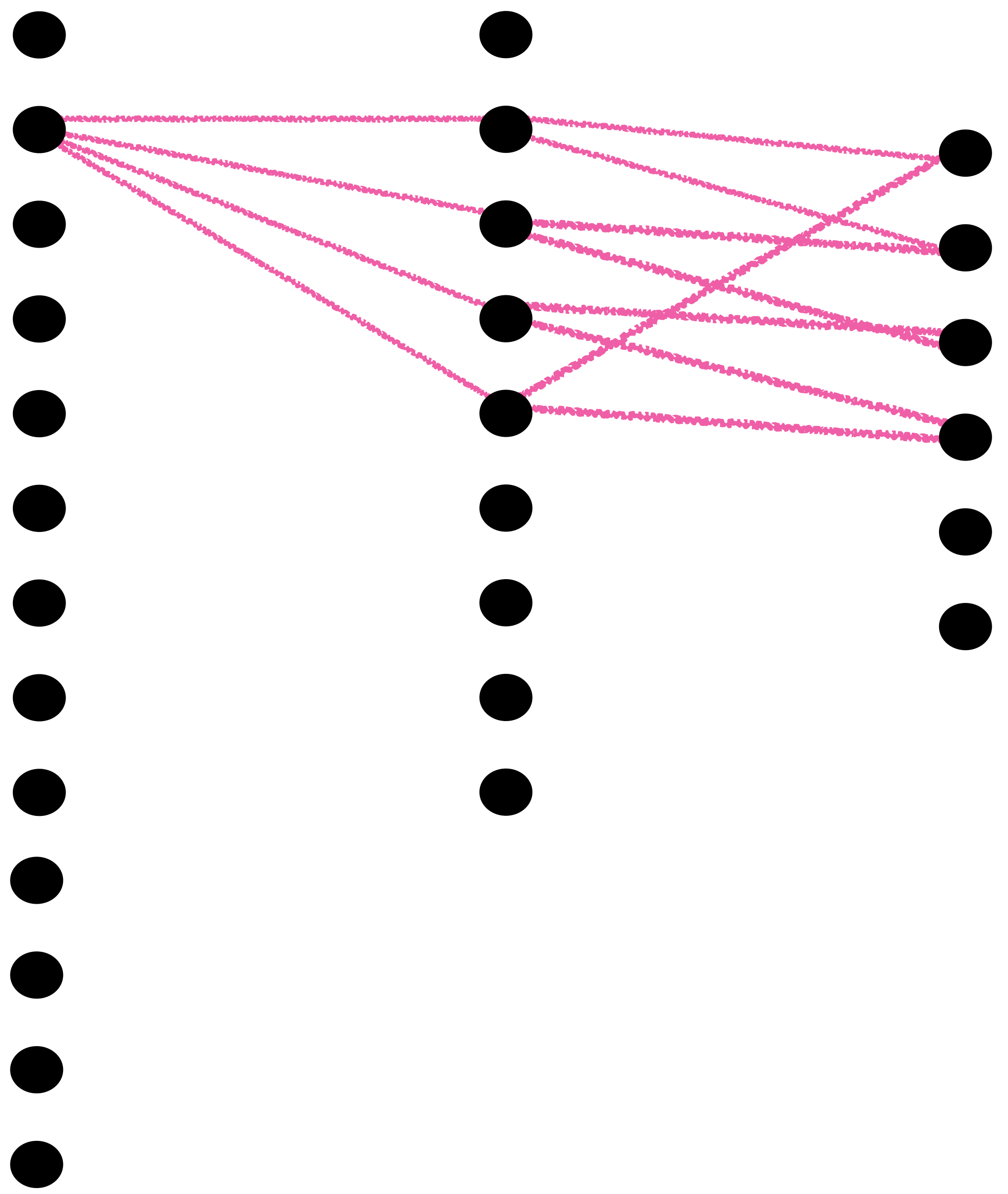
3-layered factor graph

Testability of Reed-Muller Codes



bits C_0 constraints dependencies

3-layered factor graph



$\mathbb{F}^m$ Codeword bits

Affine Lines Constraints

Affine Planes Dependences

3-layered factor graph

High dimensional expansion

The idea of using a higher-dimensional complex instead of a graph for LTCs has been circulating a number of years.

HDXs exhibit local-to-global features: prove something locally and then use expansion to globablize

[Garland 1973, Kaufman-Kazhdan-Lubotzky2014, Evra-Kaufman2016, Oppenheim2017, D.-Kaufman2017, Dinur-H.-Kaufman-LivniNavon-TaShma2018, Anari-Liu-OveisGharan-Vinzant2019]

Dikstein-Dinur-H.-RonZewi2019 proved that if one defines a code on a HDX using a base code that itself is an LTC, (and if there is an agreement-test), then the entire code is an LTC.

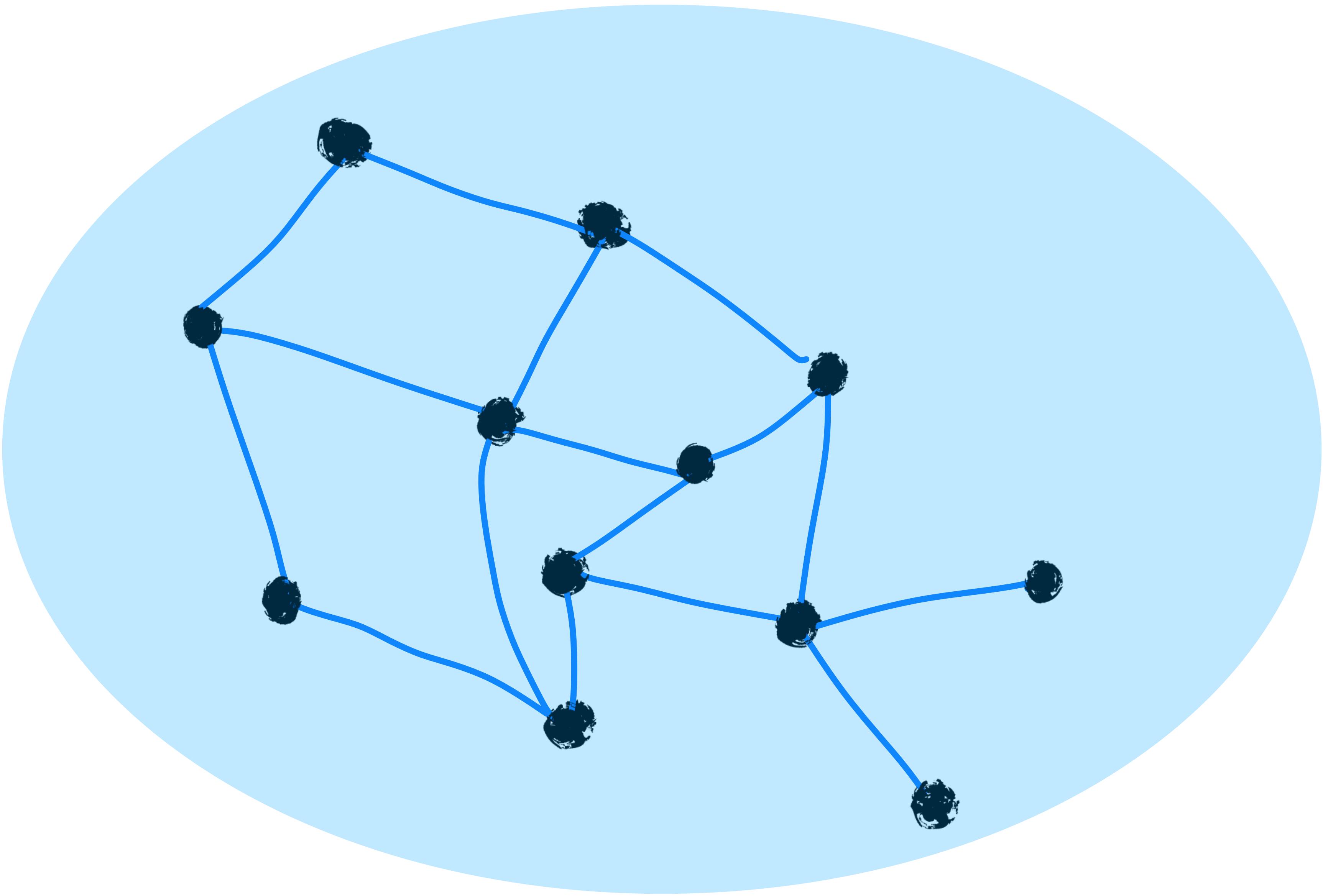
Recently also Kaufman-Oppenheim 2021 proved a similar "schema".

How to"instantiate" this? …we worked on the Lubotzky-Samuels-Vishne complexes (quotients of BT buildings), and have conjectured base codes, but no proof of local LTCness

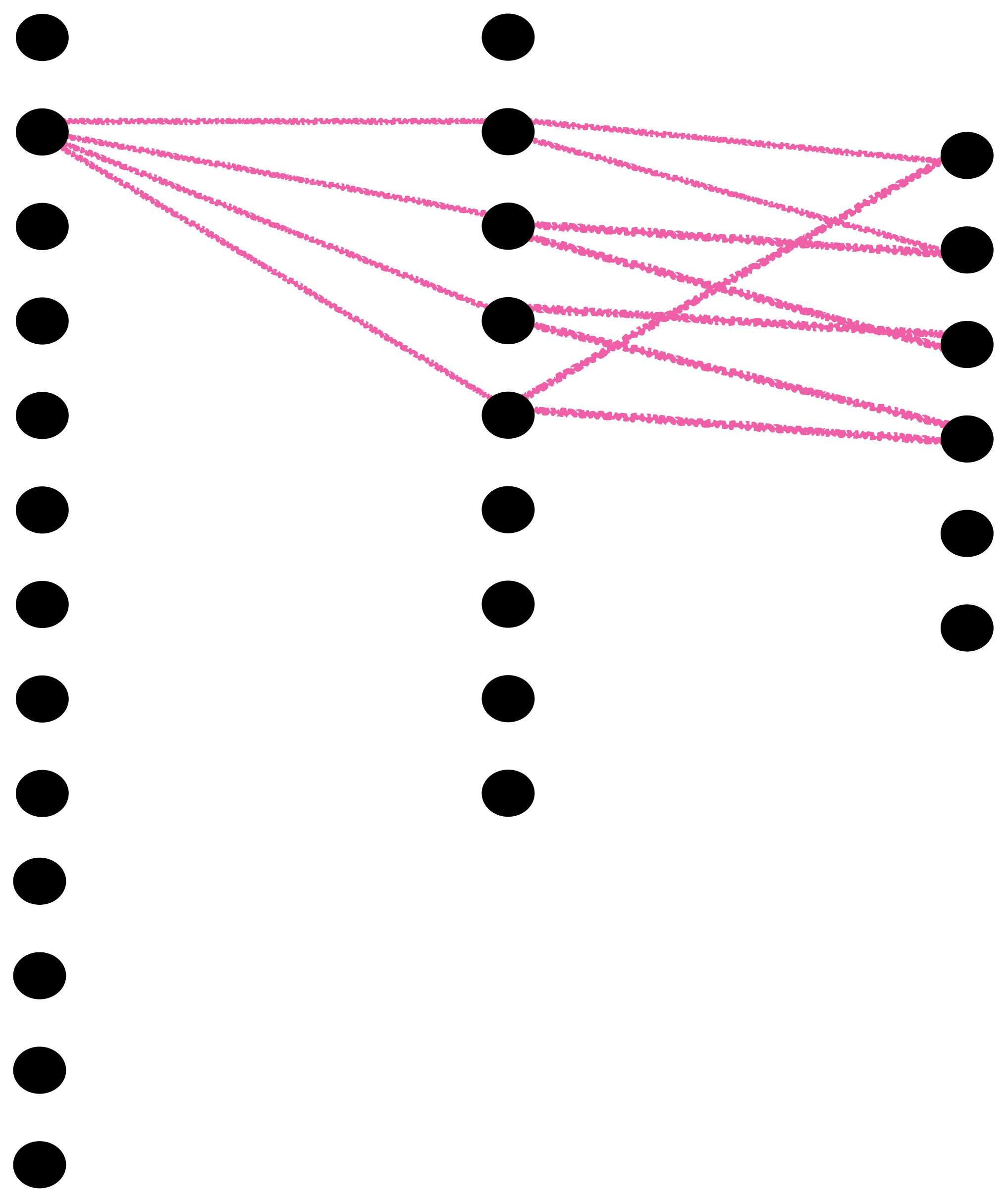
Dinur-Evra-Livne-Lubotzky-Mozes Approach

- High-dimensional expansion not required
- A square complex suffices

Expander Codes, one level up



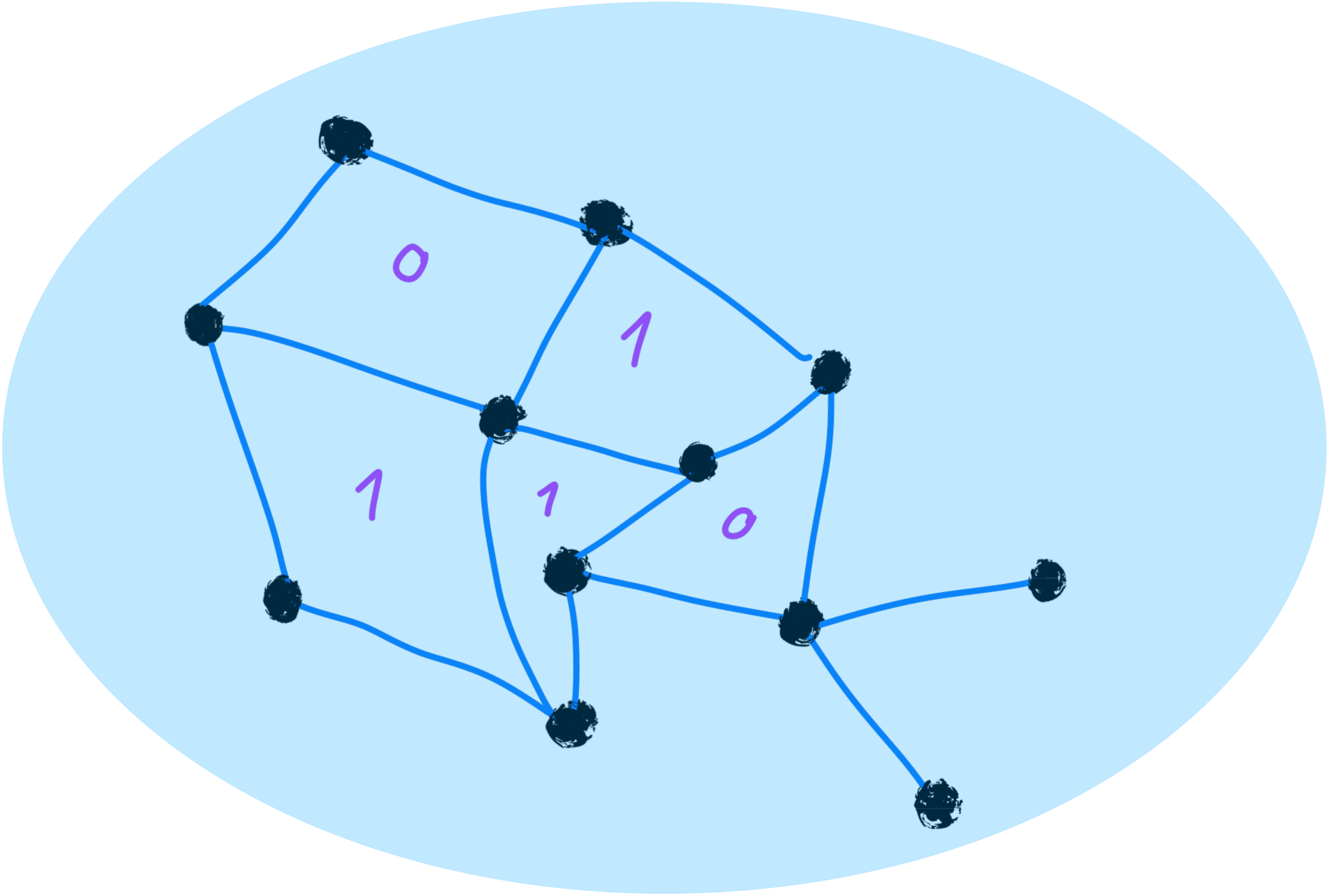
Squares Edges Vertices

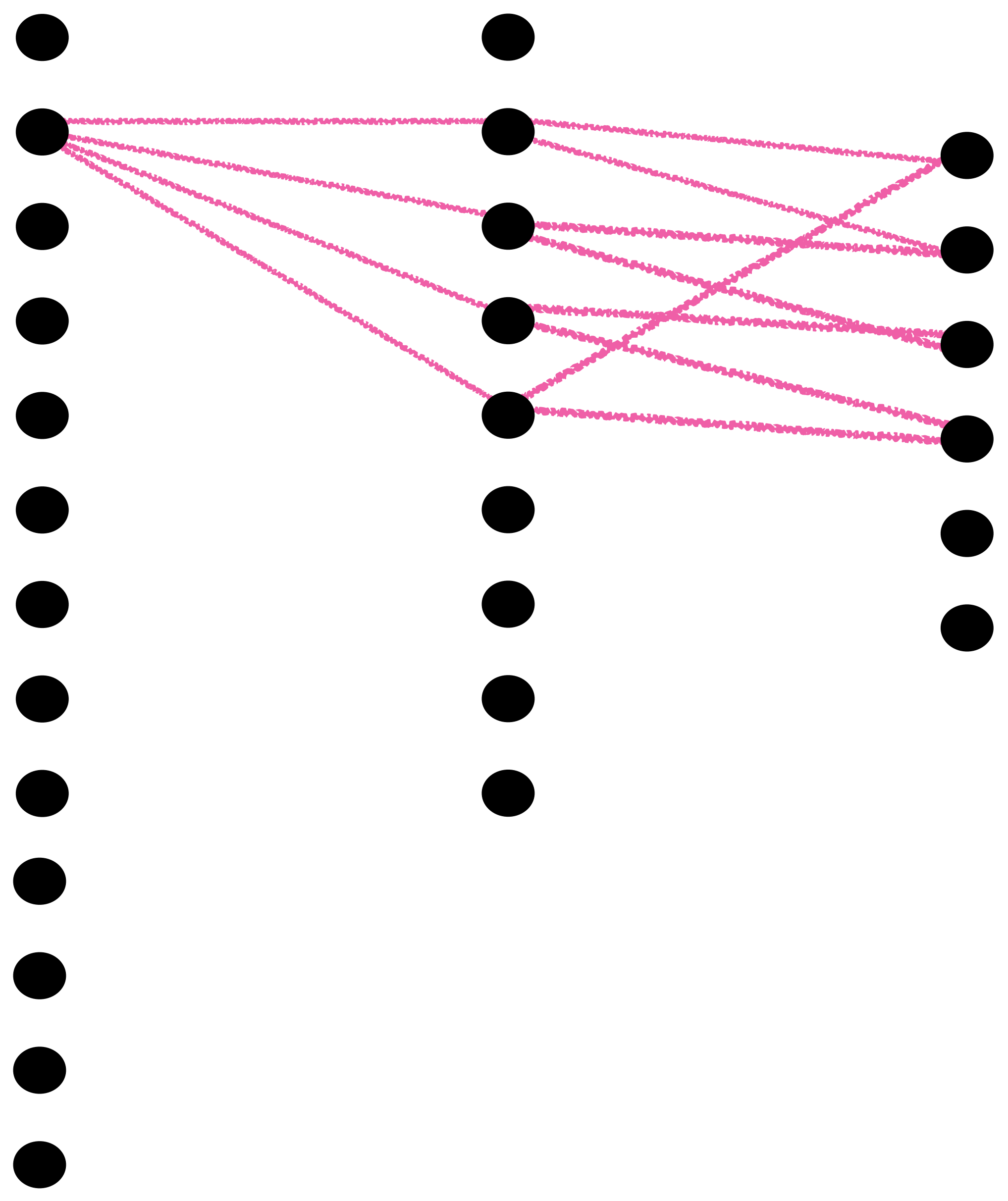


bits C_0 constraints dependencies

factor graph

Expander Codes, one level up





Squares Edges Vertices

bits C_0 constraints dependencies

factor graph

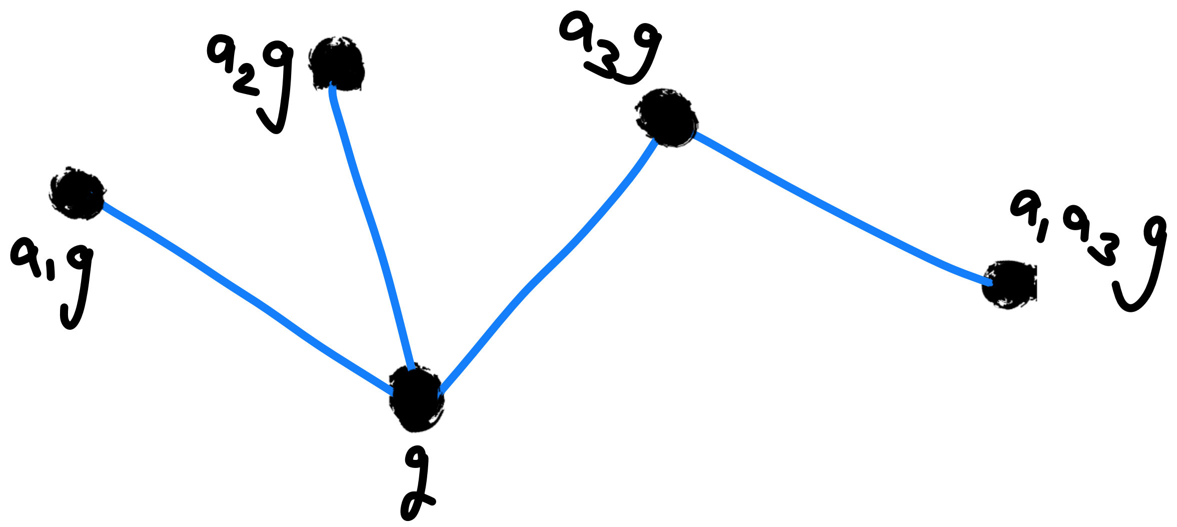
Left-right Cayley Complex

"a graph with squares"

Let G be a finite group,

Let $A \subset G$ be closed under taking inverses, i.e. such that $a \in A \ \rightarrow \ a^{-1} \in A$

Cay(G,A) is a graph with vertices G, and edges $E_A = \{\{g, ag\} : g \in G, a \in A\}$

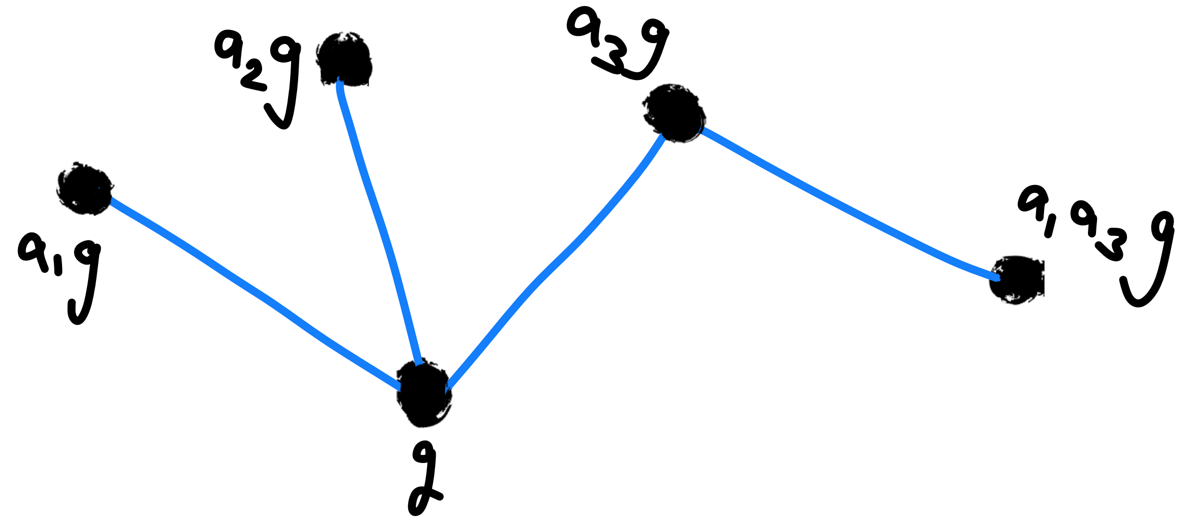


Left-right Cayley Complex

"a graph with squares"

Let G be a finite group,

Let $A, B \subset G$ be closed under taking inverses



Left-right Cayley Complex

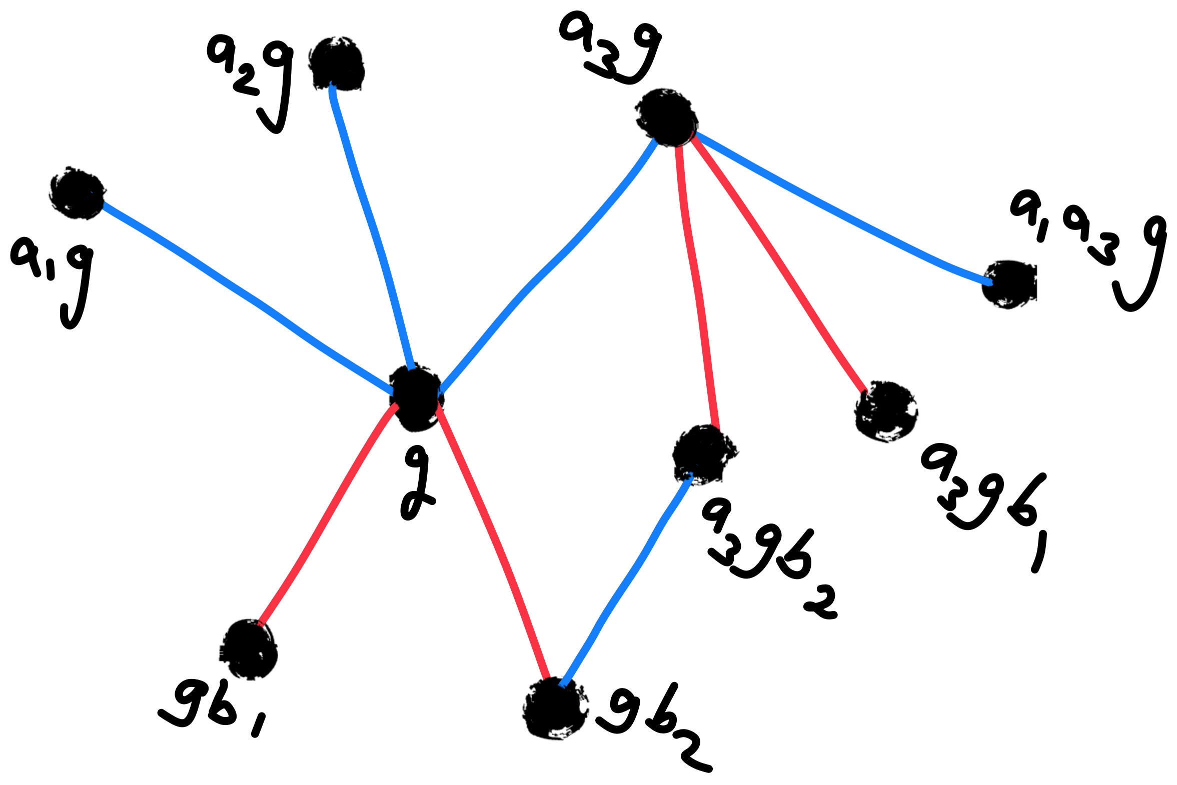
"a graph with squares"

Let G be a finite group,

Let $A, B \subset G$ be closed under taking inverses

Cay(G,A) is a graph with vertices G, and edges $E_A = \{\{g, ag\} : g \in G, a \in A\}$ (left *)

Cay(G,B) is a graph with vertices G, and edges $E_B = \{\{g, gb\} : g \in G, b \in B\}$ (right *)



Left-right Cayley Complex

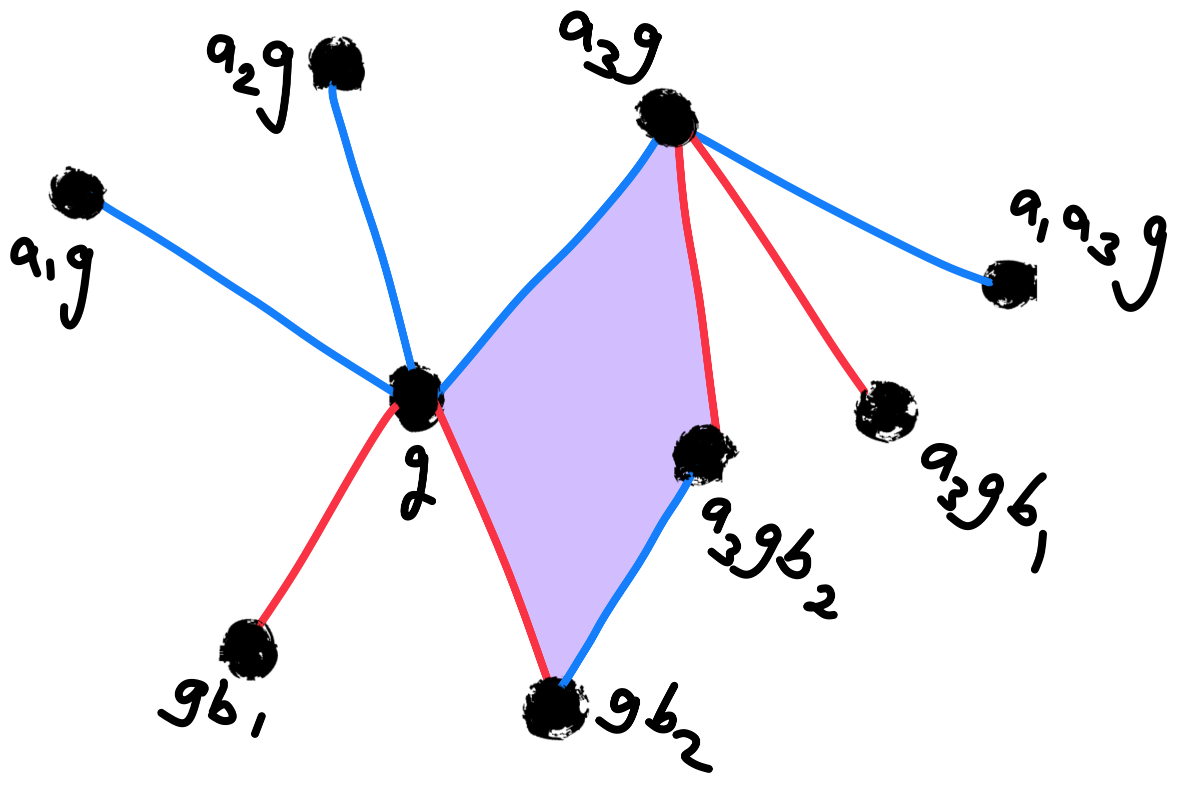
"a graph with squares"

Let G be a finite group,

Let $A, B \subset G$ be closed under taking inverses

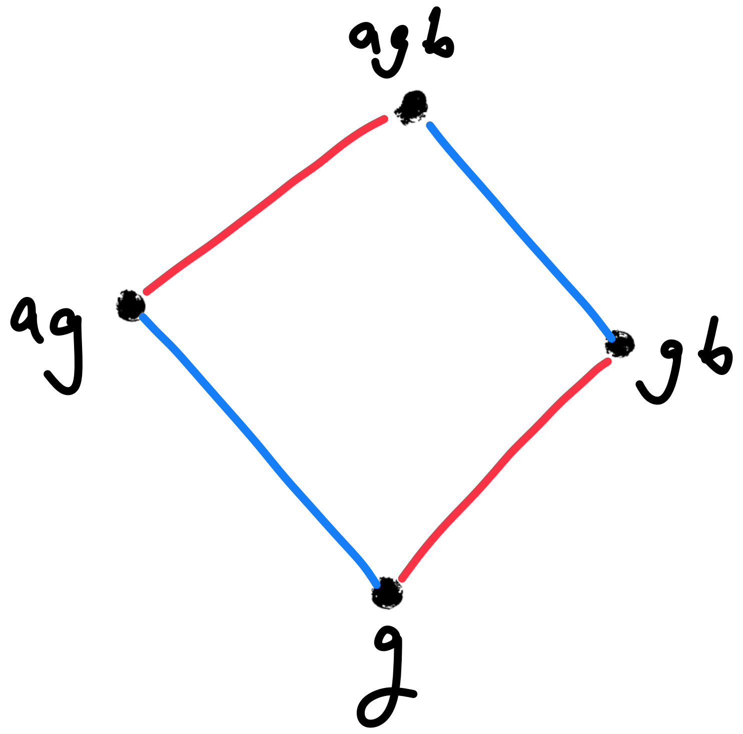
Cay(G,A) is a graph with vertices G, and edges $E_A = \{\{g, ag\} : g \in G, a \in A\}$ (left *)

Cay(G,B) is a graph with vertices G, and edges $E_B = \{\{g, gb\} : g \in G, b \in B\}$ (right *)



Left-right Cayley Complex

“a graph with squares”



(a^{-1}, agb, b^{-1})

(a^{-1}, ag, b)

(a, gb, b^{-1})

(a, g, b)

Each triple $a \in A, g \in G, b \in B$ define a <u>rooted square</u> (a, g, b)

Each square can have 4 roots,

$[a, g, b] = \{ \ (a, g, b), \quad (a^{-1}, ag, b), \quad (a^{-1}, agb, b^{-1}), \quad (a, gb, b^{-1}) \ \}$

This square naturally contains

- The edges {g,ag}, {g,gb}, {gb,agb}, {ag,agb},
- The vertices g,ag,gb,agb

The set of squares is $X(2) = \{[a, g, b] : g \in G, a \in A, b \in B\} \quad = \quad A \times G \times B \,/ \, \sim$

Left-right Cayley Complex Cay^2(A,G,B)

Let G be a finite group, and let $A, B \subset G$ be closed under taking inverses.

The left-right Cayley complex Cay^2(A,G,B) has

- Vertices G
- Edges $E_A \cup E_B$

$$E_A = \{\{g, ag\} : g \in G, a \in A\}, \quad E_B = \{\{g, gb\} : g \in G, b \in B\}$$

- Squares A x G x B / ~

We say that Cay^2(A,G,B) is a λ-expander if Cay(G,A) and Cay(G,B) are λ-expanders.

Lemma: For every $\lambda > 0$ there are explicit infinite families of bounded-degree left-right Cayley complexes that are λ-expanders.

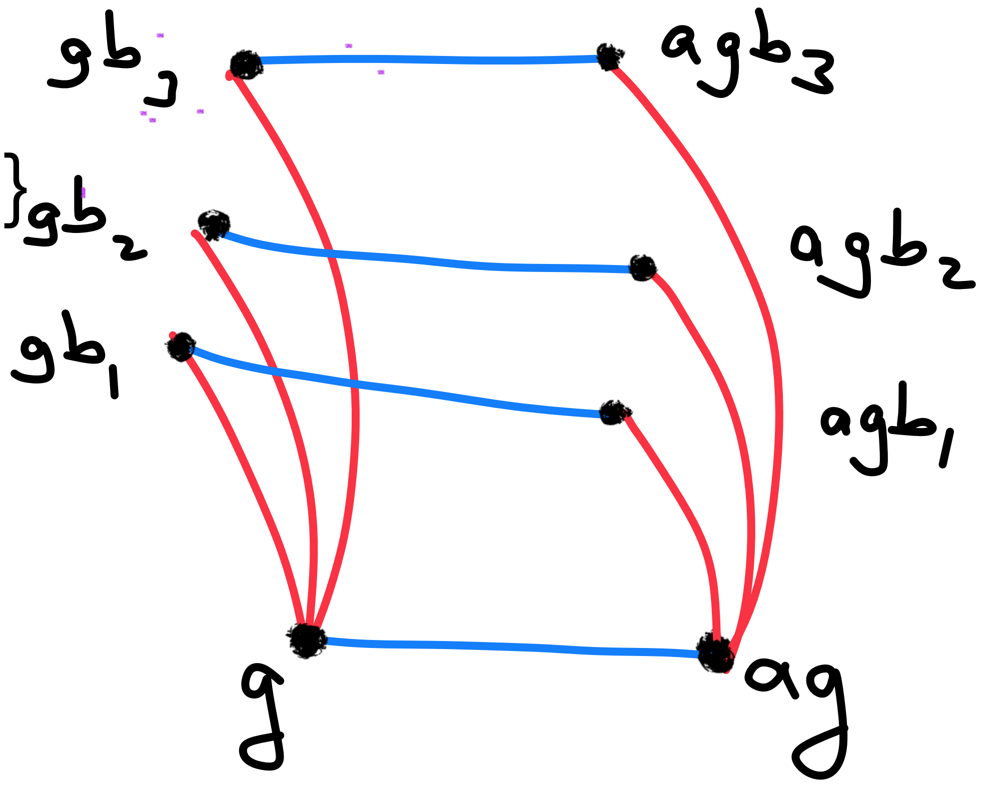
Left-right Cayley Complex

"a graph with squares"

Squares touching the edge {g,ag}

are naturally identified with B

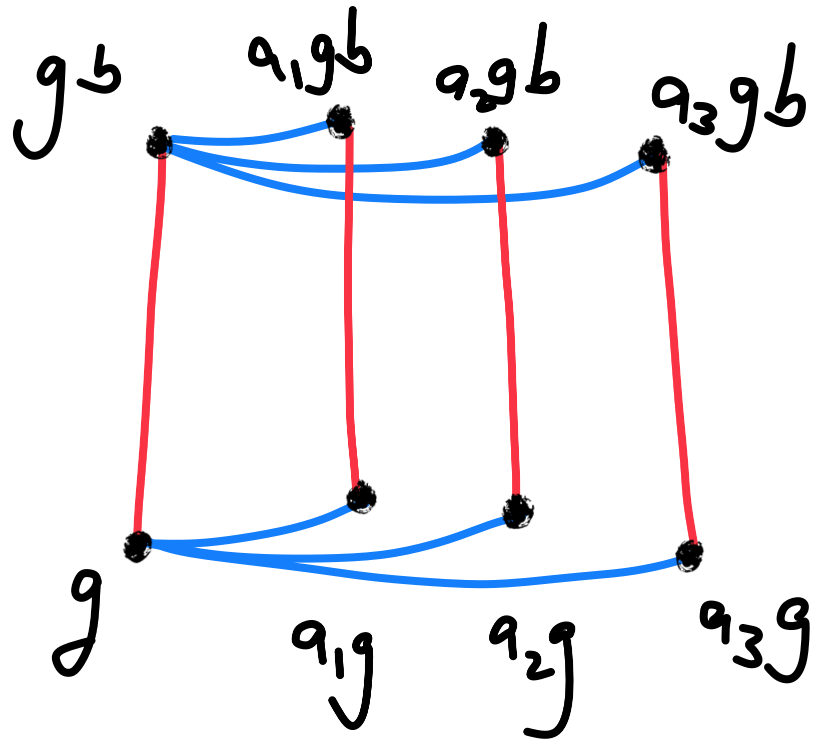
$$b \mapsto [a, g, b]$$



Squares touching the edge {g,gb}

are naturally identified with A

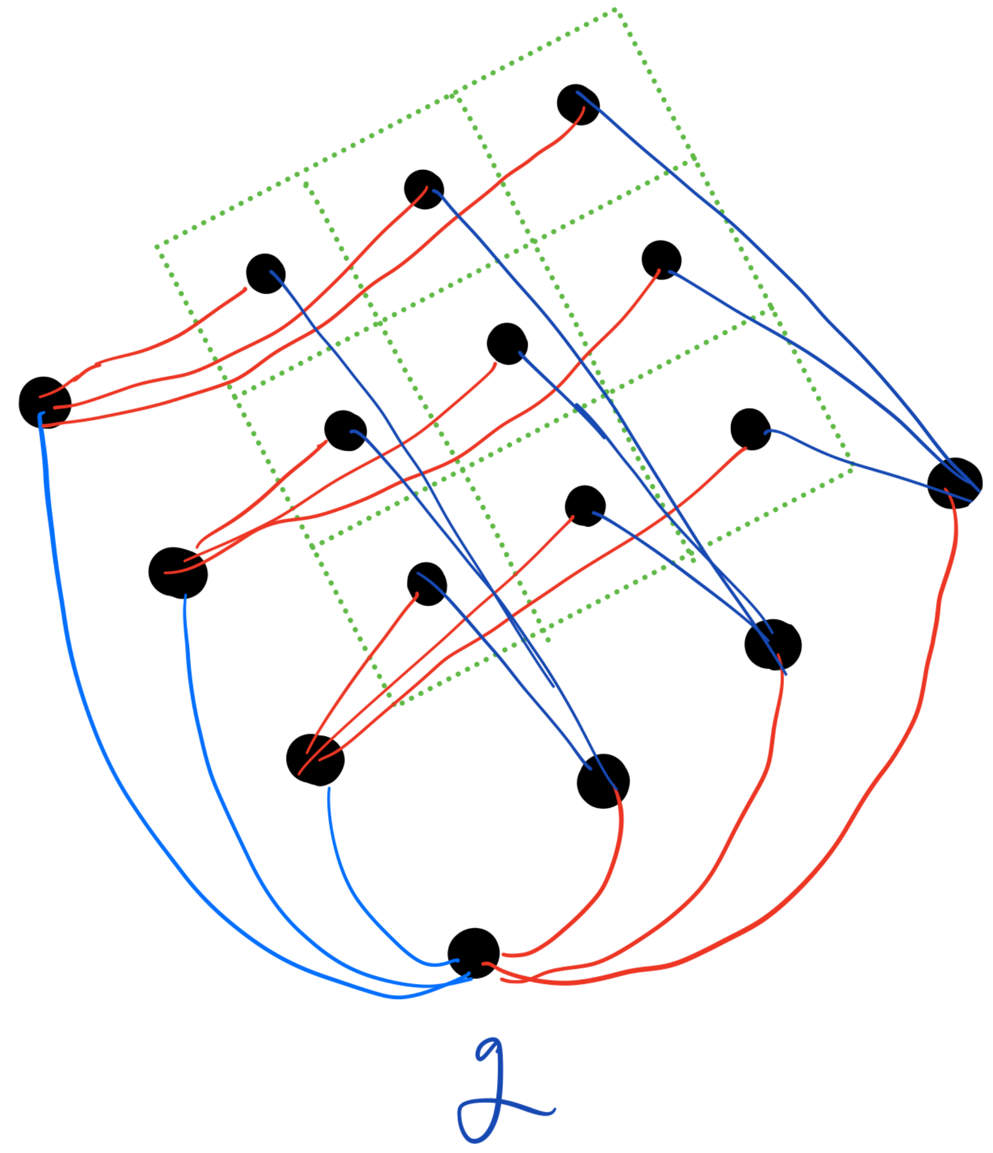
$$a \mapsto [a, g, b]$$



A vertex g has |A| + |B| neighbors

For each $a \in A, b \in B$ there is one square touching g,

so there is a natural bijection* $(a, b) \mapsto [a, g, b]$



* it is a bijection assuming $\forall a, b, g, \quad g^{-1}ag \neq b$

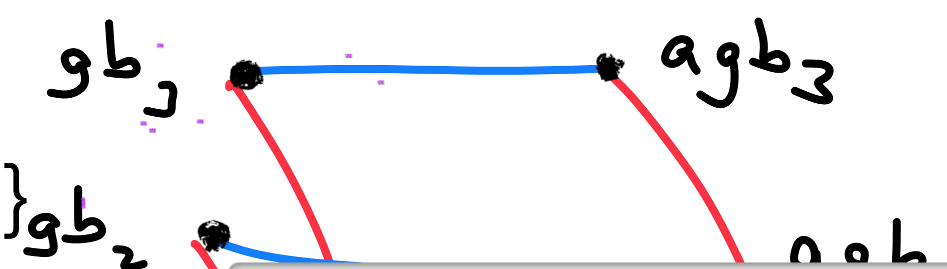
Left-right Cayley Complex

"a graph with squares"

Squares touching the edge {g,ag}

are naturally identified with B

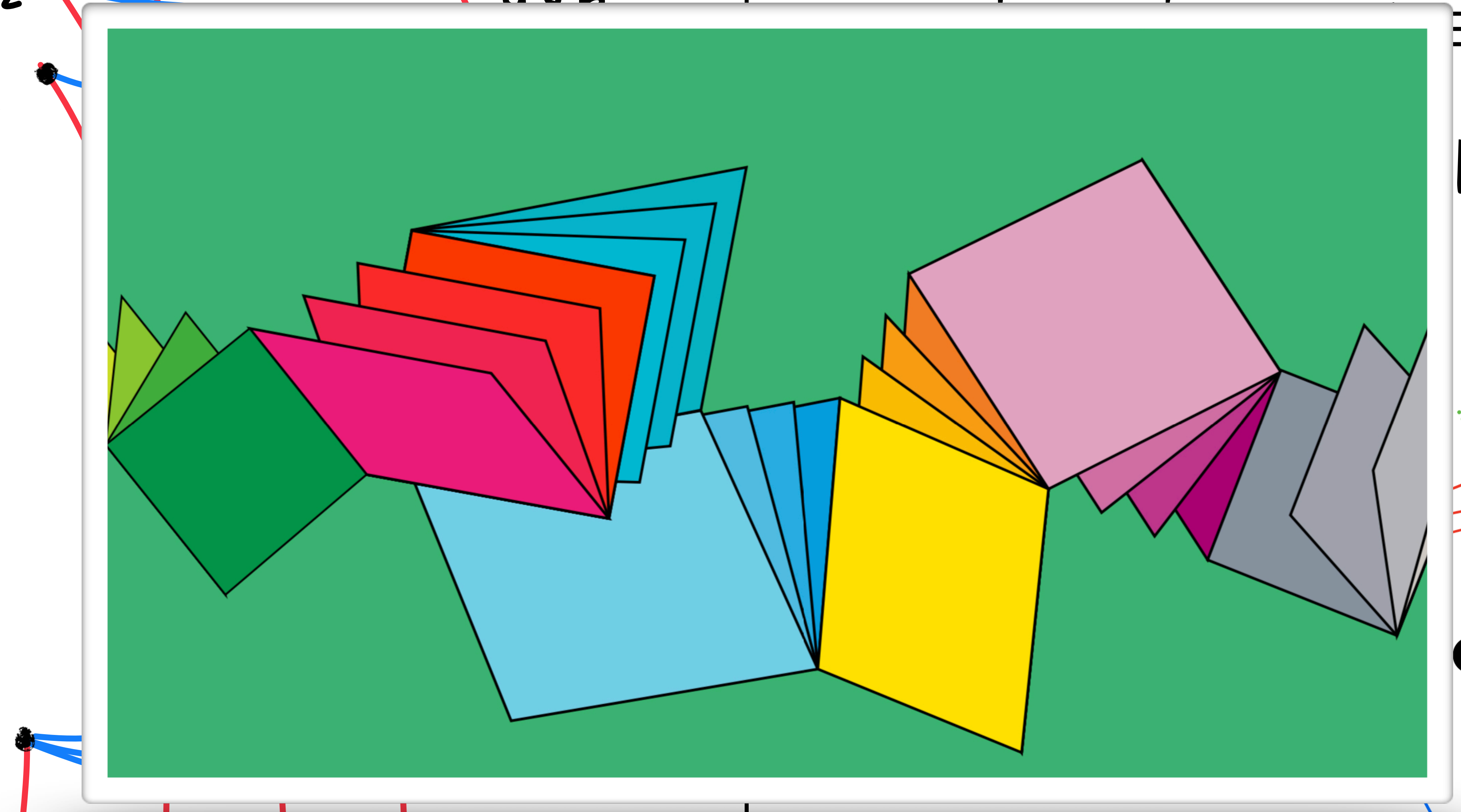
$b \mapsto [a,g,b]$

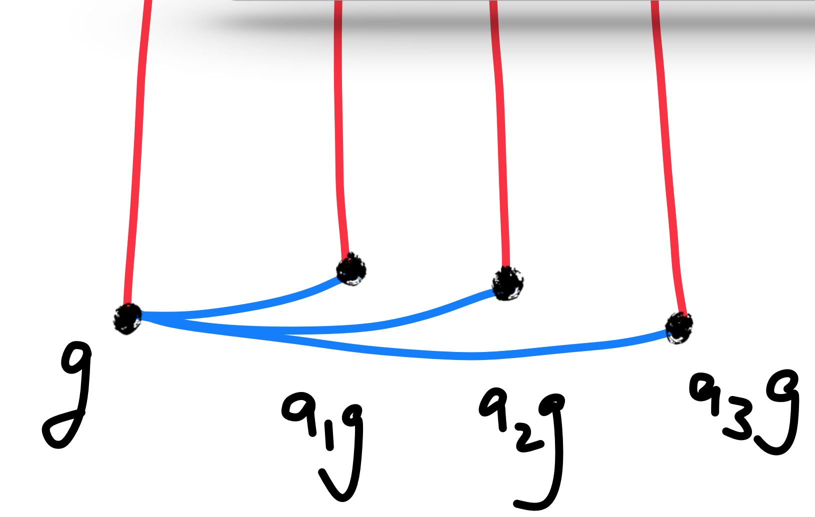


Squares touching the edge {g,gb}

are naturally identified with A

$a \mapsto [a,g,b]$

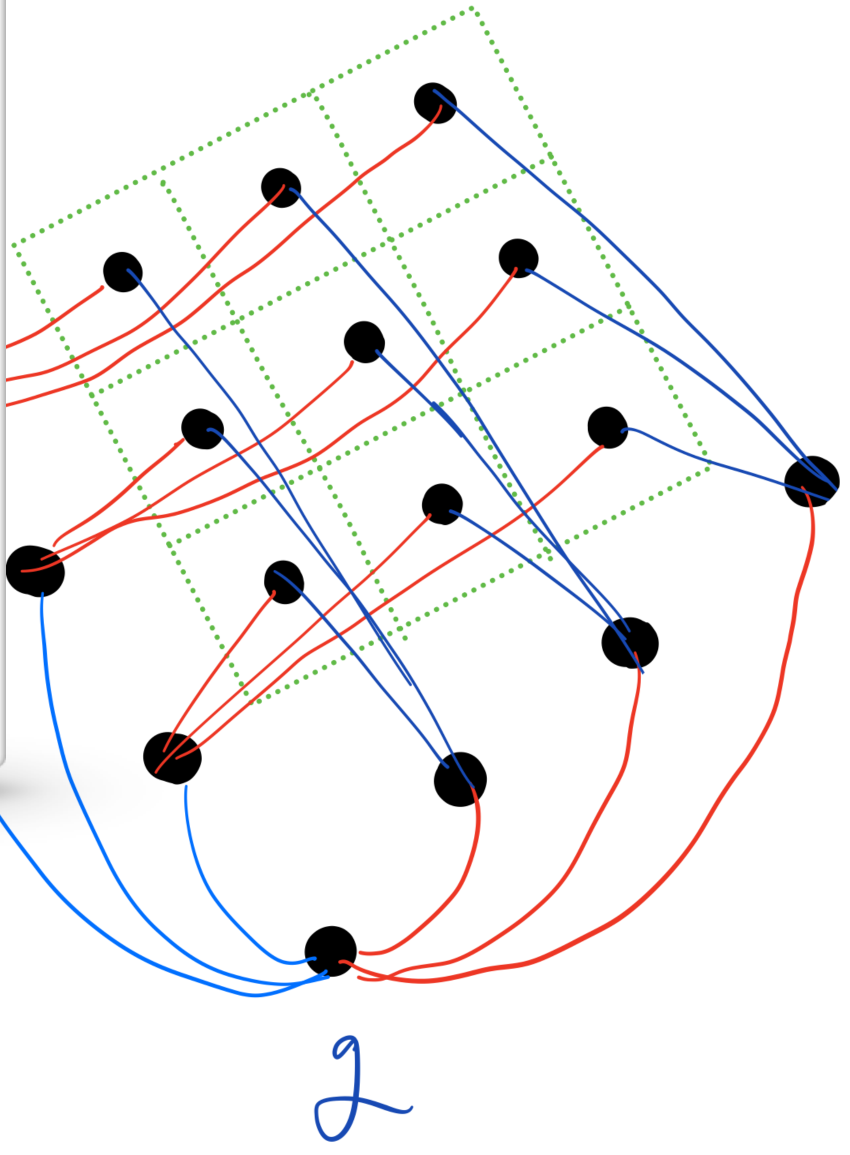




A vertex g has |A| + |B| neighbors

F $\in B$ there is one square touching g,

bijection* $(a,b) \mapsto [a,g,b]$

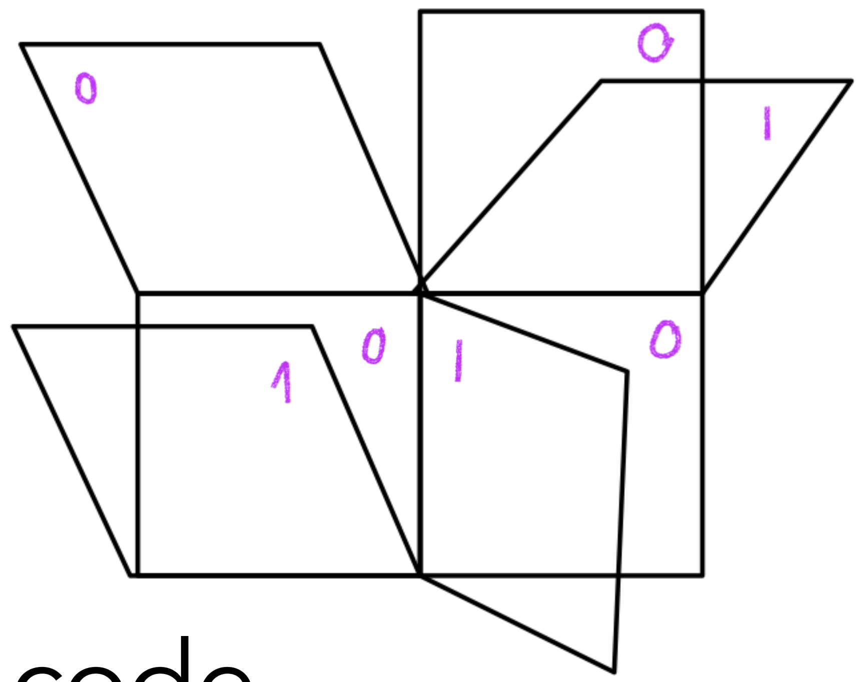


* it is a bijection assuming $\forall a,b,g,\quad g^{-1}ag \neq b$

The Code

Let Cay²(A,G,B) be a left-right Cayley complex.

Fix base codes $C_A \subseteq \{0,1\}^A, C_B \subseteq \{0,1\}^B$ (assuming |A| = |B| = d we can take one base code $C_0 \subseteq \{0,1\}^d$ and let $C_A, C_B \simeq C_0$)



Define a code CODE = $C[G, A, B, C_A, C_B]$:

- The codeword bits are placed on the squares
- Each edge requires that the bits on the squares around it are in the base code

$$\text{CODE} = \{f : Squares \to \{0,1\} : \forall a, g, b, \quad f([\,\cdot\,, g, b]) \in C_A, f([a, g, \cdot\,]) \in C_B\}$$

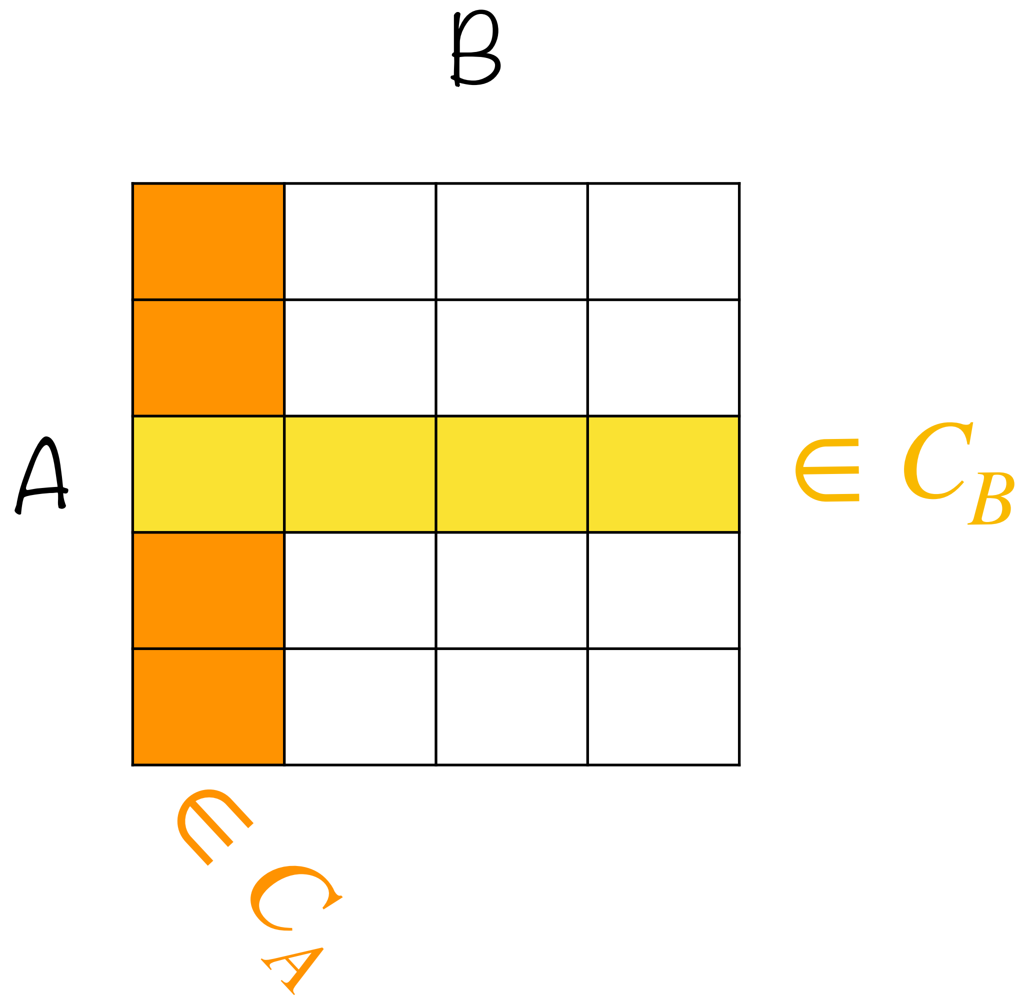
Rate: $\geq 4r_0 - 3$ [calc: #squares - #constraints]

Distance: $\geq \delta_0^2(\delta_0 - \lambda)$ [easy propagation argument]

Local views are tensor codes

Claim: Fix f∈CODE. For each $g \in G$, $f([\cdot, g, \cdot]) \in C_A \otimes C_B$

Theorem: Assume Cay^2(A,G,B) is a λ-expander, and $C_A \otimes C_B$ is ρ-robustly testable. If $\lambda < \delta_0\rho/5$, then $C[G, A, B, C_A, C_B]$ is locally testable.



The tester is as follows:

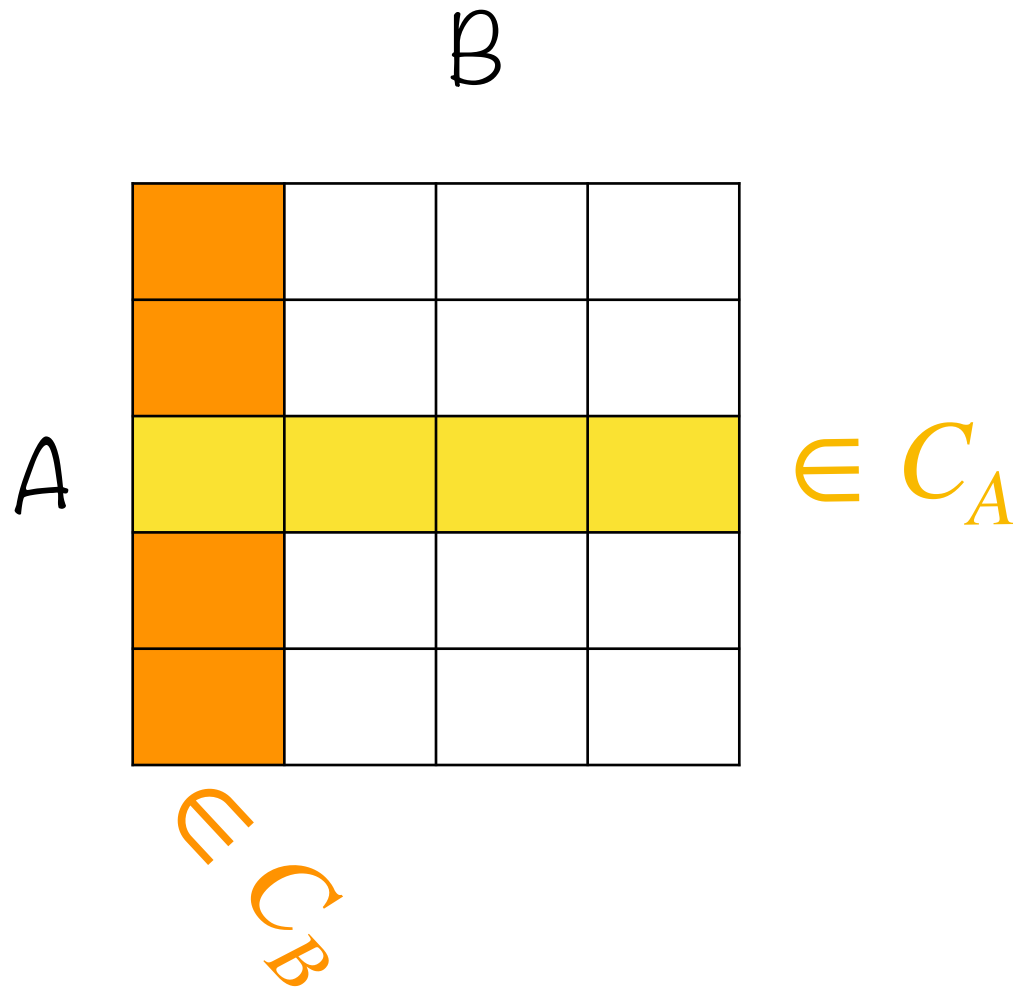
```
1. Select a vertex g uniformly,
2. Read f on all $|A| \cdot |B|$ squares touching g, namely $f([\cdot, g, \cdot])$.
3. Accept iff this belongs to $C_A \otimes C_B$
```

Then $\Pr_{g \in G}[f([\cdot, g, \cdot]) \notin C_A \otimes C_B) \geq const \cdot dist(f, C[G, A, B, C_A, C_B])$

CODE = $\{f : Squares \to \{0,1\} : \forall a, g, b, \quad f([\cdot, g, b]) \in C_A, f([a, g, \cdot]) \in C_B\}$

Robustly-testable tensor codes

Definition [Ben-Sasson-Sudan'05]: $C_A \otimes C_B$ is ρ-robustly testable if for all $w : A \times B \to \{0,1\}$, $\rho \cdot dist(w, C_A \otimes C_B) \leq$ row-distance + column-distance



Row-distance : average distance of each row to C_A

Column-distance : average distance of each column to C_B

Lemma [Ben-Sasson-Sudan'05, Dinur-Sudan-Wigderson2006, Ben-Sasson-Viderman2009]:

For every r>0 there exist base codes with rate r and constant distance whose tensors are robustly-testable. (Random LDPC codes, LTCs)

Proof of local-testability

Start with $f : Squares \to \{0,1\}$ and find $f' \in C,\ \ rej(f) \geq dist(f, f') \cdot const$

ALG "self-correct":

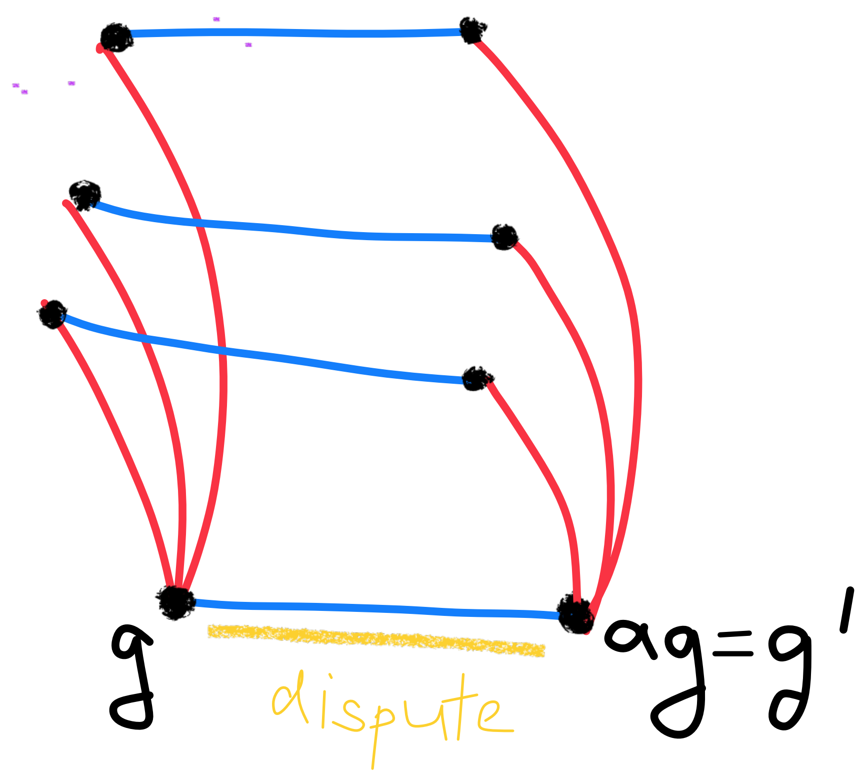
```
1. Init: Each $g \in G$ finds $T_g \in C_A \otimes C_B$ closest to
   $f([\cdot, g, \cdot])$

   [ define a progress measure $\Phi$ = # dispute edges ]

2. Loop: If g can change $T_g$ and reduce $\Phi$ then do it

3. End: If $\Phi = 0$ let $f'([a, g, b]) = T_g(a, b)$ and output $f'$,
   If $\Phi > 0$ quit
```

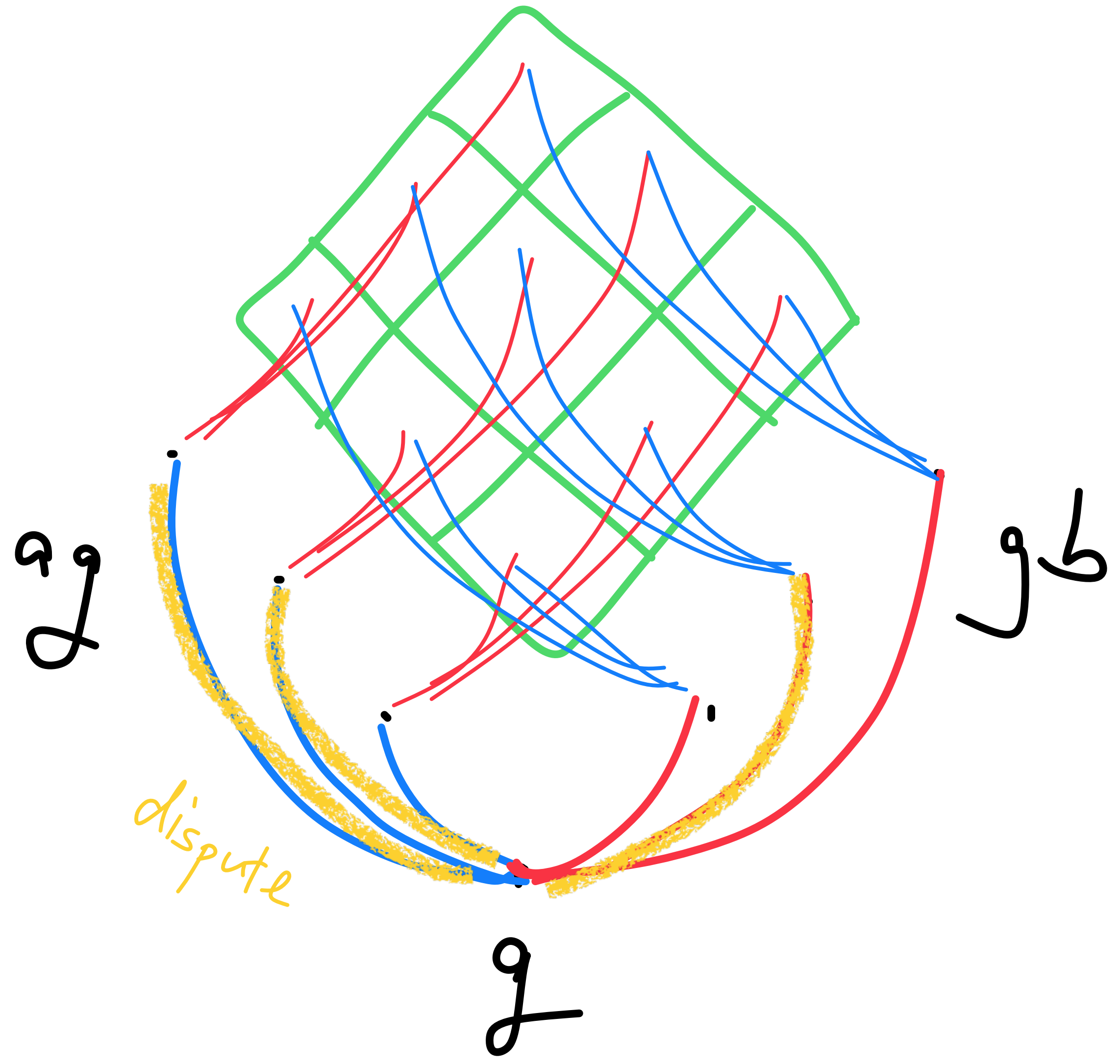
- steps $\leq \Phi \ \approx$ rej(f)
- If $\Phi = 0$ then $rej(f) \geq dist(f, f') \cdot const$
- **If $\Phi > 0$ then $\Phi > 0.1$** so $rej(f) \geq dist(f, f') \cdot 0.1$

Proof of local-testability

If ALG "self-correct" is stuck then rej (f) > 0.1

- If g,g′ are in dispute, there must be many squares on {g,g′} with further dispute edges
- Can try to propagate, but, they all might be clumped around g
- But then g's neighbors all agree, so there must have been a better choice for T_g (using the LTCness of tensor codes)
- Random walk **edge—>square—>edge** + expansion ==> dispute set is large





A concrete choice of group & base codes

Theorem: For all $0 < r < 1$ there exist $\delta > 0$ and $q \in \mathbb{N}$ and an explicit construction of an infinite family of error-correcting codes $\{C_n\}_n$ with rate $\geq r$, distance $\geq \delta$ and locally testable with q queries.

Proof: Take

1. Family of base codes $\{C_d\}_d$ with rate $> \frac{r+3}{4}$ and constant robustness ρ and distance δ
2. Set λ small enough wrt δ and ρ
3. Choose a family $\{Cay^2(A_n, G_n, B_n)\}_n$ of λ expanding left-right Cayley complexes, with $d = |A_n| = |B_n| = O(1/\lambda^2)$
4. Output $\{C[G_n, A_n, B_n, C_d, C_d]\}_n$

…questions?

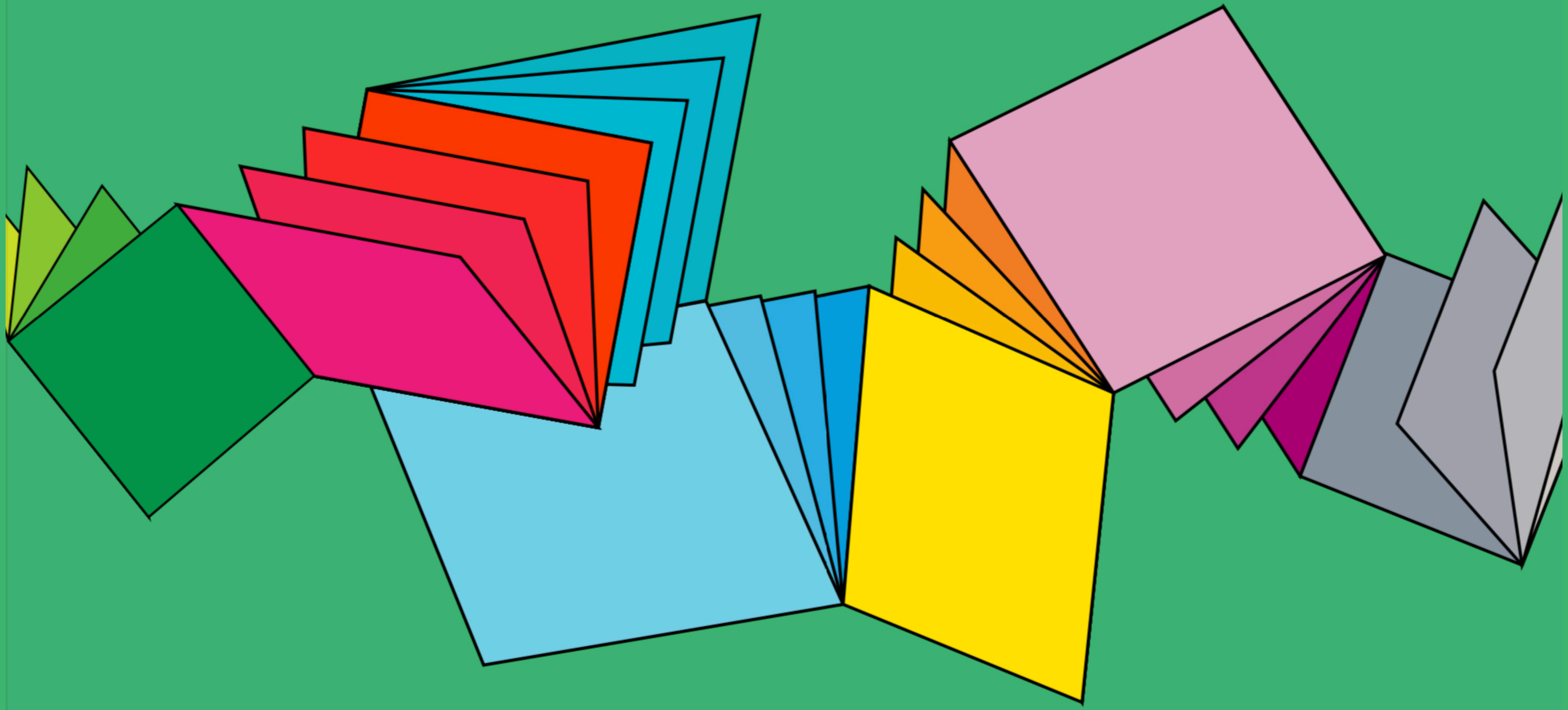
- Can such ideas be used for constructing PCPs?
- Can these codes be made practical?
- Can one consrtuct LTCs on other HDX's such as LSV simplical complexes? It all boils down to building one finite code in the links
- Can one construct higher dimensional (e.g. cubical) complexes similarly?

References

- Irit Dinur, Shai Evra, Ron Livne, Alexander Lubotzky, and Shahar Mozes. *Locally testable codes with constant rate, distance, and locality*. In Stefano Leonardi and Anupam Gupta, eds., Proc. 54th ACM Symp. on Theory of Computing (STOC), pages 357–374. 2022. arXiv:2111.04808

- Prahladh Harsha, *The Blooming of the c3 LTC Flowers*, A blogpost on the constant-query locally testable code construction due to Dinur, Evra, Livne, Lubotzky and Mozes, In Calvin Café: The Simons Institute Blog, September 2022.

c^3-LTC constructions



Thank You