Probabilistic and Combinatorial Methods - Part 2

Error-Correcting Codes: Theory and Practice Boot Camp

Jonathan Mosheiff
Ben-Gurion University

Let’s recall

Definition: A **random code ensemble** $C \subseteq \mathbb{F}_q^n$ is **k-locally-similar** to an **RLC of rate R** if

$$\Pr\left[\{v_1, \ldots, v_k\} \subseteq C\right] \lesssim 2^{-(1-R)\cdot n \cdot \dim\{v_1,\ldots,v_k\}}$$

for all $v_1, \ldots, v_k \in \mathbb{F}_q^n$.

Let's recall

Definition: A **random code ensemble** $C \subseteq \mathbb{F}_q^n$ is k**-locally-similar** to an **RLC of rate** R if

$$\Pr\left[\{v_1, \ldots, v_k\} \subseteq C\right] \lesssim 2^{-(1-R)\cdot n \cdot \dim\{v_1,\ldots,v_k\}}$$

for all $v_1, \ldots, v_k \in \mathbb{F}_q^n$.

Theorem: If a **random code ensemble** C is k**-locally-similar** to an **RLC of rate** R then it **achieves the list-decoding GV-bound** and any other **monotone,** k**-local and symmetric property** of **RLC codes** with high probability.

Warm up: The generalized Wozencraft Ensemble

Definition: Let $n = 2m$. Let $\varphi : \mathbb{F}_{2^m} \to \mathbb{F}_2^m$ be the natural binary encoding.

Sample α uniformly at random from $\mathbb{F}_q$.

The **Wozencraft ensemble** is

$$C_\alpha = \left\{ (\varphi(x), \varphi(\alpha x)) \mid x \in \mathbb{F}_q) \right\} \subseteq \mathbb{F}_2^n$$

Warm up: The generalized Wozencraft Ensemble

Claim: C_α is 1**-locally-similar** to an **RLC of rate** $\frac{1}{2}$**.**

Warm up: The generalized Wozencraft Ensemble

Claim: C_α is **1-locally-similar** to an **RLC of rate $\frac{1}{2}$.**

Proof: Let $y \in \mathbb{F}_2^n \backslash \{0\}$.

There is a unique way to write $y = (\varphi(x), \varphi(\beta x))$ for some $\beta \in \mathbb{F}_2^n$.

So $y \in \mathbb{C}_\alpha$ only if $\alpha = \beta$, which happens with probability $2^{-m} = 2^{-\frac{n}{2}}$.

Warm up: The generalized Wozencraft Ensemble

Claim: C_α is 1**-locally-similar** to an **RLC of rate** $\frac{1}{2}$**.**

Corollary: The **Wozencraft ensemble** achieves the same 1**-local properties** as an **RLC of rate** $\frac{1}{2}$**.**

In particular, it **achieves the GV-bound for minimal distance**.

Warm up: The generalized Wozencraft Ensemble

Definition: Let $n = 2m$. Let $\varphi : \mathbb{F}_{2^m} \to \mathbb{F}_2^m$ be the natural binary encoding.

Sample a uniformly random polynomial p of degree $< k$ from $\mathbb{F}_q[x]$.

The k**-generalized Wozencraft ensemble** is

$$C_p = \left\{ (\varphi(x), \varphi(p(x))) \mid x \in \mathbb{F}_q \right\} \subseteq \mathbb{F}_2^n$$

Claim: C_p is k**-locally-similar** to an **RLC of rate** $\frac{1}{2}$**.**

Warm up: The generalized Wozencraft Ensemble

Claim: C_p is **k-locally-similar** to an **RLC of rate $\frac{1}{2}$.**

Warm up: The generalized Wozencraft Ensemble

Claim: C_p is k**-locally-similar** to an **RLC of rate $\frac{1}{2}$.**

Corollary: The **Wozencraft ensemble** achieves the same k**-local properties** as an **RLC of rate $\frac{1}{2}$.**

In particular, it **achieves the list-decoding GV-bound** for list size up to k.

Randomly Punctured Low-Bias codes

Theorem [Guruswami-M]: Let D be a **low-bias code** and let C be a **random puncturing of D.** Then C is **k-locally similar** to an **RLC** of similar rate.

Randomly Punctured Low-Bias codes

Theorem [Guruswami-M]: Let D be a **low-bias code** and let C be a **random puncturing of D.** Then C is **k-locally similar** to an **RLC** of similar rate.

Caveat: This theorem requires **constant alphabet size**.

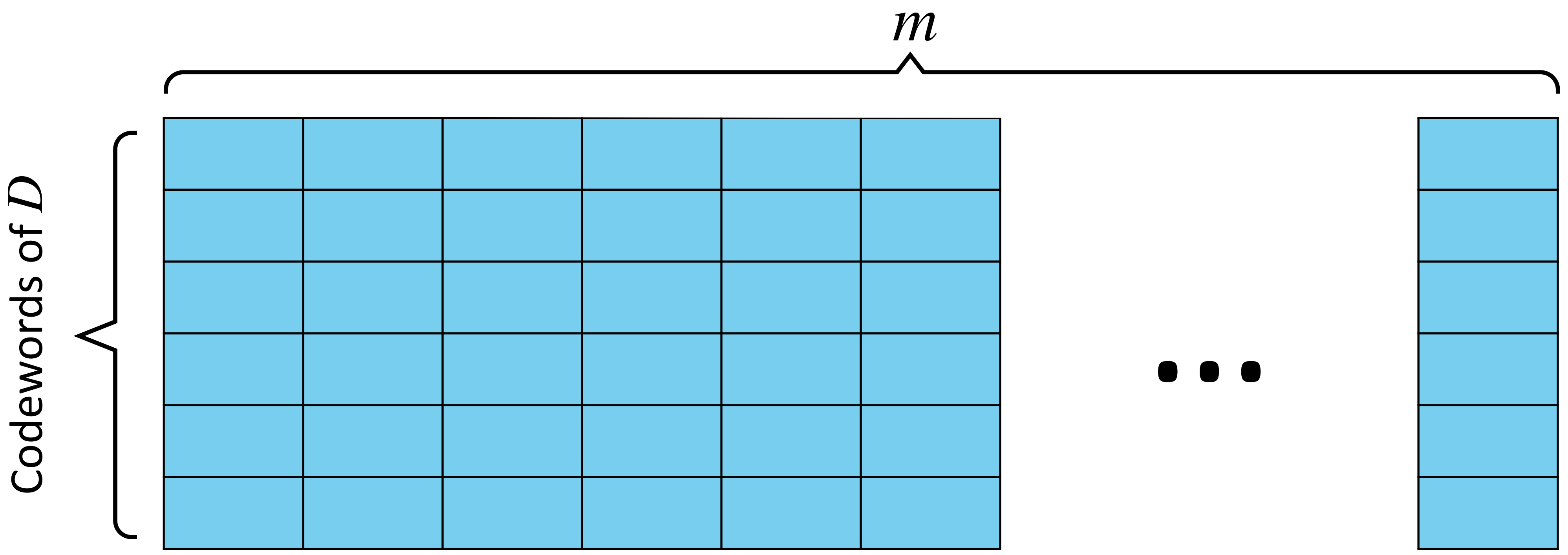
Randomly Punctured Low-Bias codes

Partially derandomized by **[Putterman-Pyne]**

Theorem [Guruswami-M]: Let D be a **low-bias code** and let C be a **random puncturing of D.** Then C is **k-locally similar** to an **RLC** of similar rate.

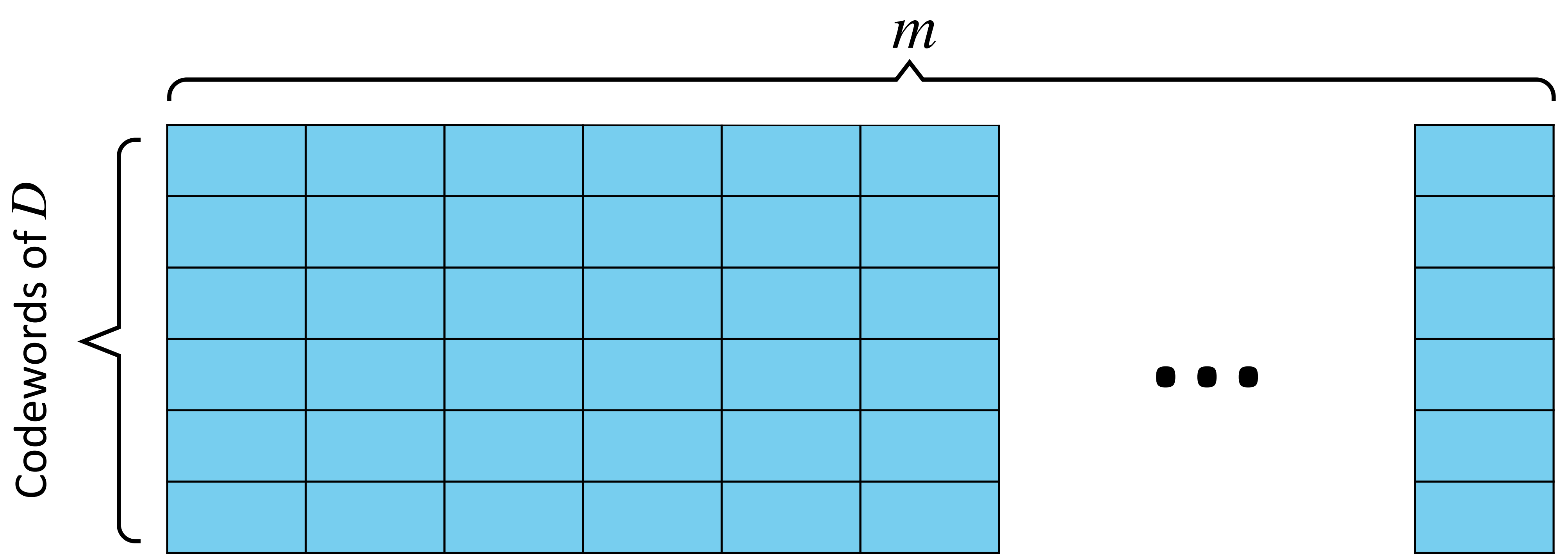
Caveat: This theorem requires **constant alphabet size**.

Puncturing of Codes



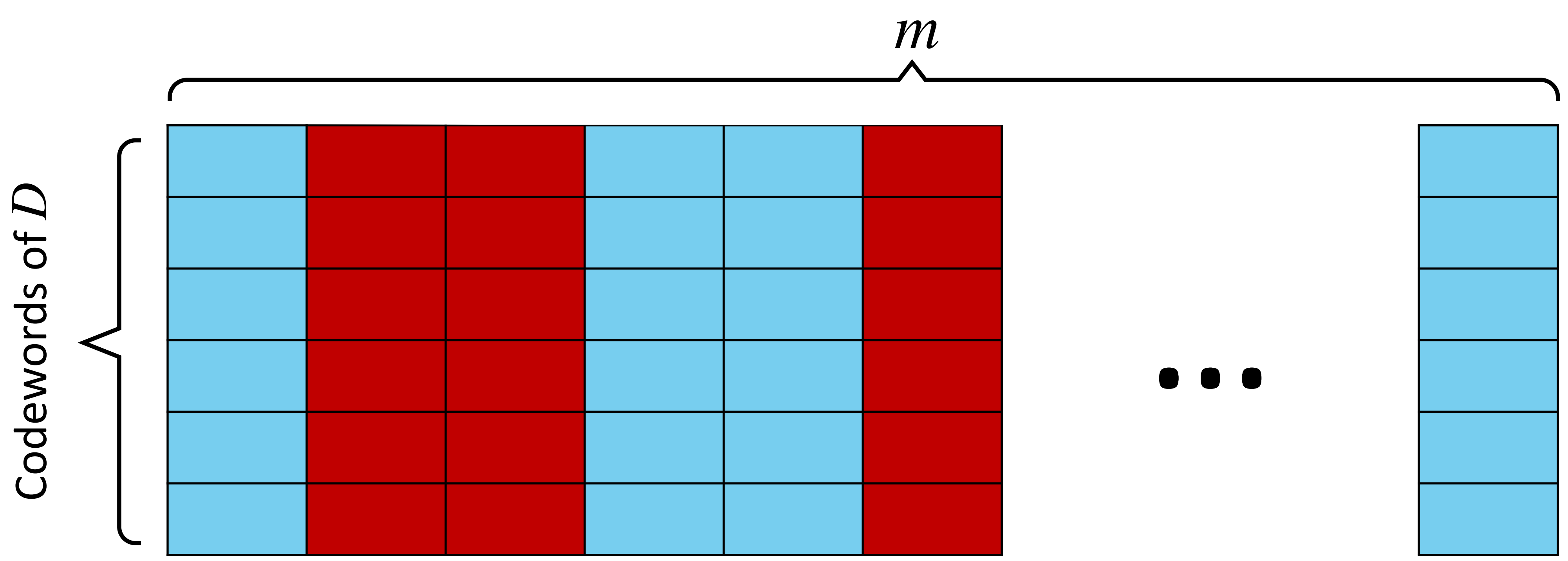
Puncturing of Codes

- From a code $D \subseteq \mathbb{F}_q^m$ create a new code $C \subseteq \mathbb{F}_q^n$. Usually $n \ll m$.



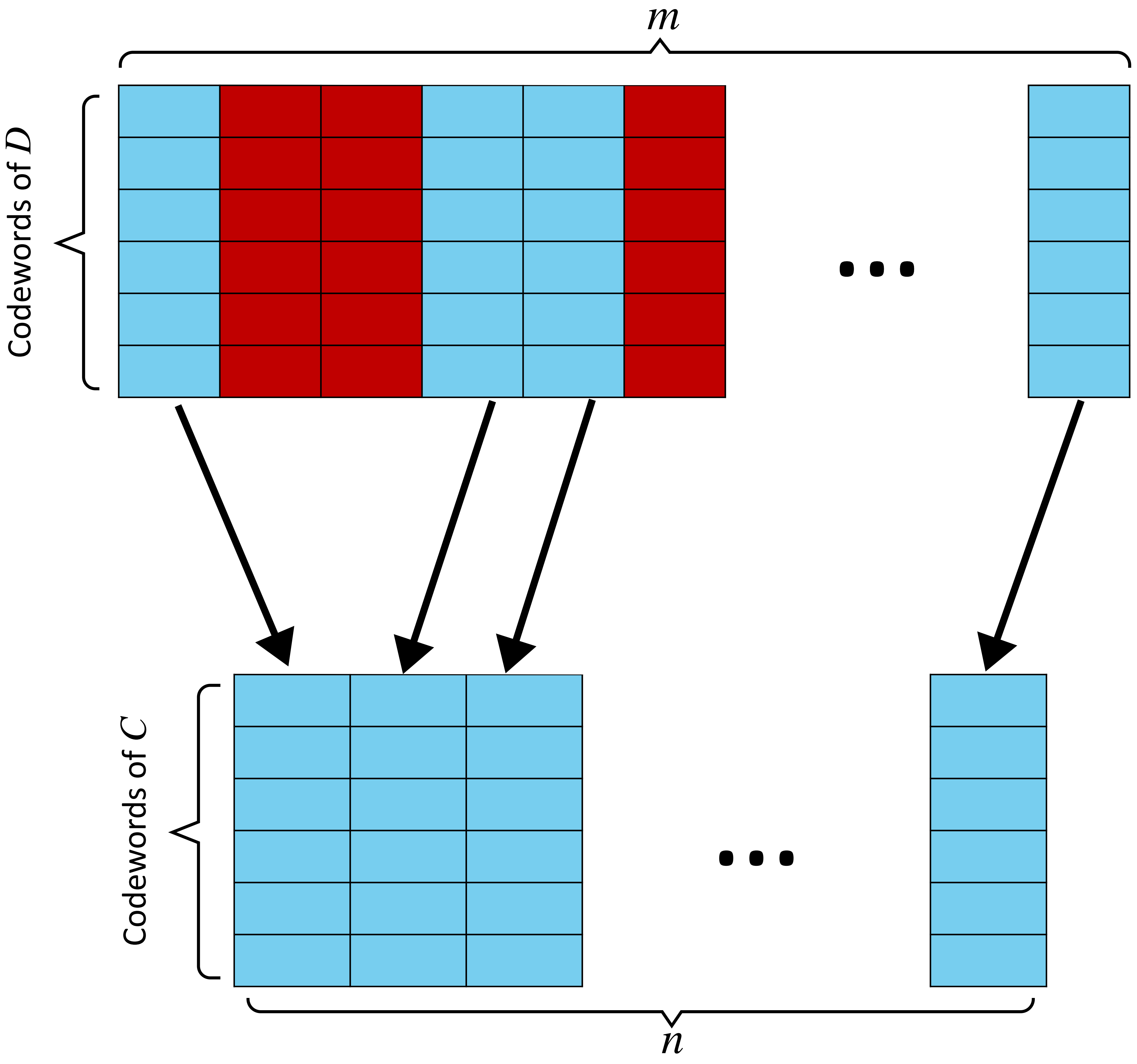
Puncturing of Codes

- From a code $D \subseteq \mathbb{F}_q^m$ create a new code $C \subseteq \mathbb{F}_q^n$. Usually $n \ll m$.



Puncturing of Codes

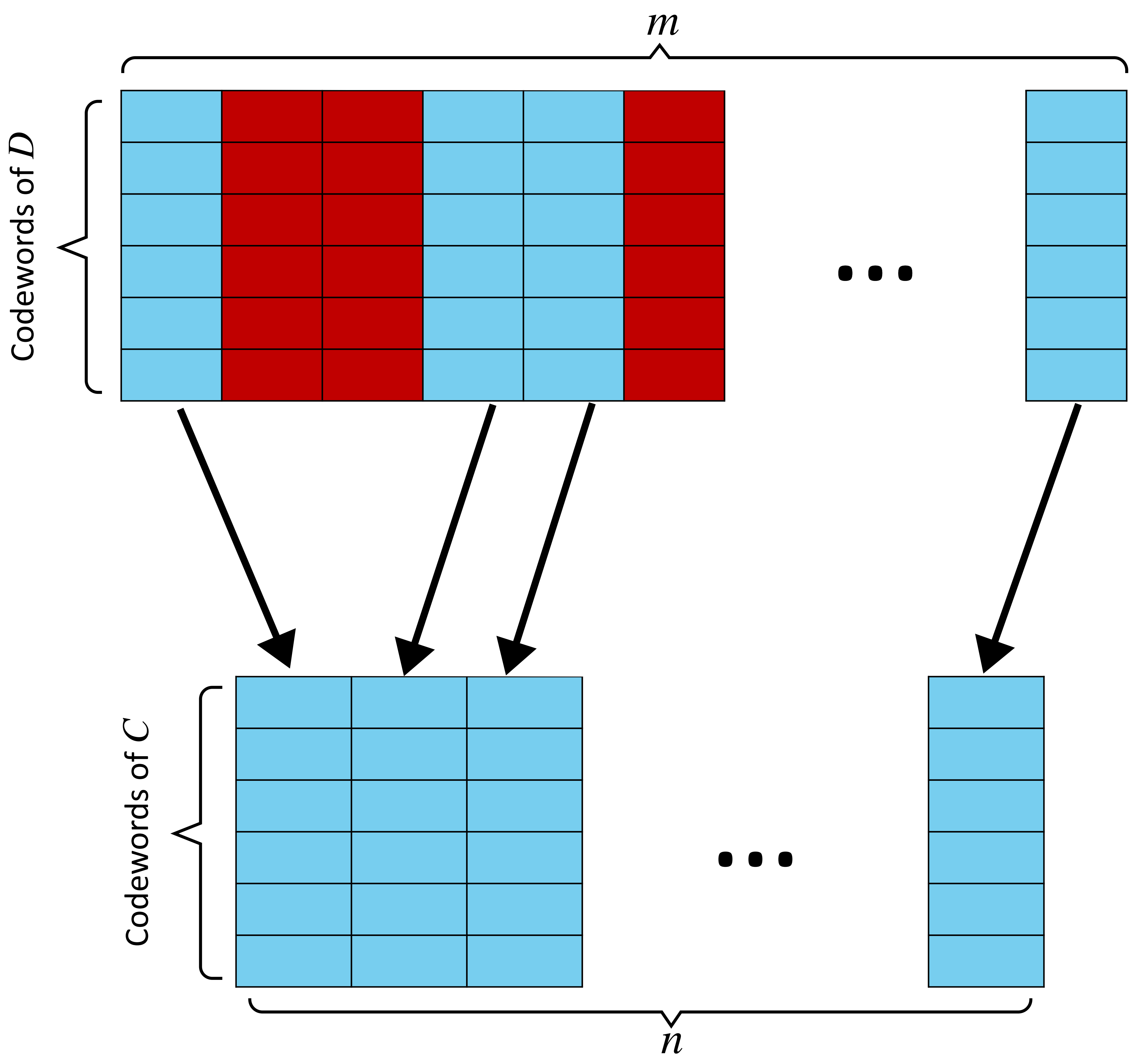
- From a code $D \subseteq \mathbb{F}_q^m$ create a new code $C \subseteq \mathbb{F}_q^n$. Usually $n \ll m$.



Puncturing of Codes

- From a code $D \subseteq \mathbb{F}_q^m$ create a new code $C \subseteq \mathbb{F}_q^n$. Usually $n \ll m$.

- If the punctured columns are chosen at random, C is said to be a **random n-puncturing of D**.

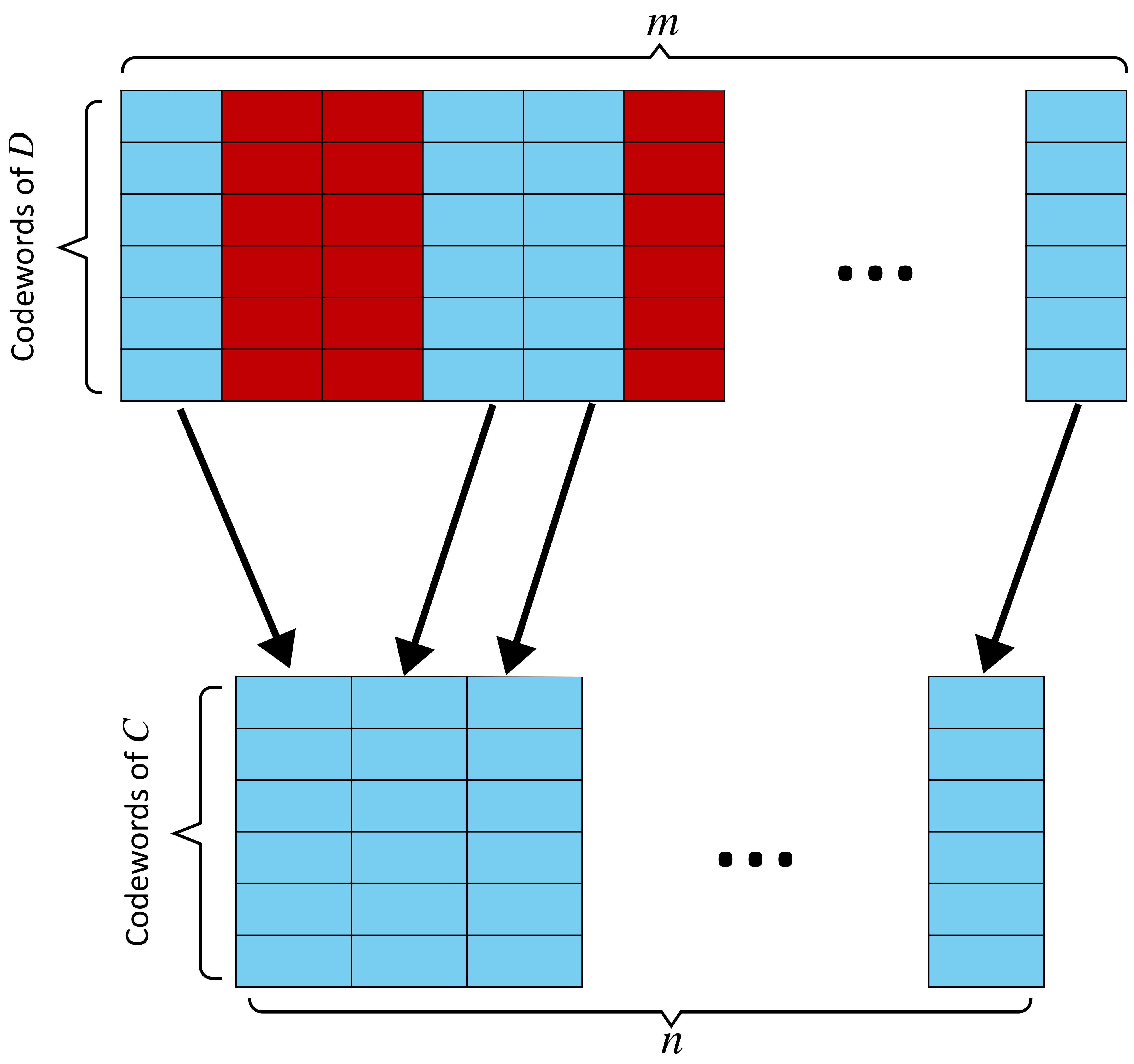


Puncturing of Codes

- From a code $D \subseteq \mathbb{F}_q^m$ create a new code $C \subseteq \mathbb{F}_q^n$. Usually $n \ll m$.

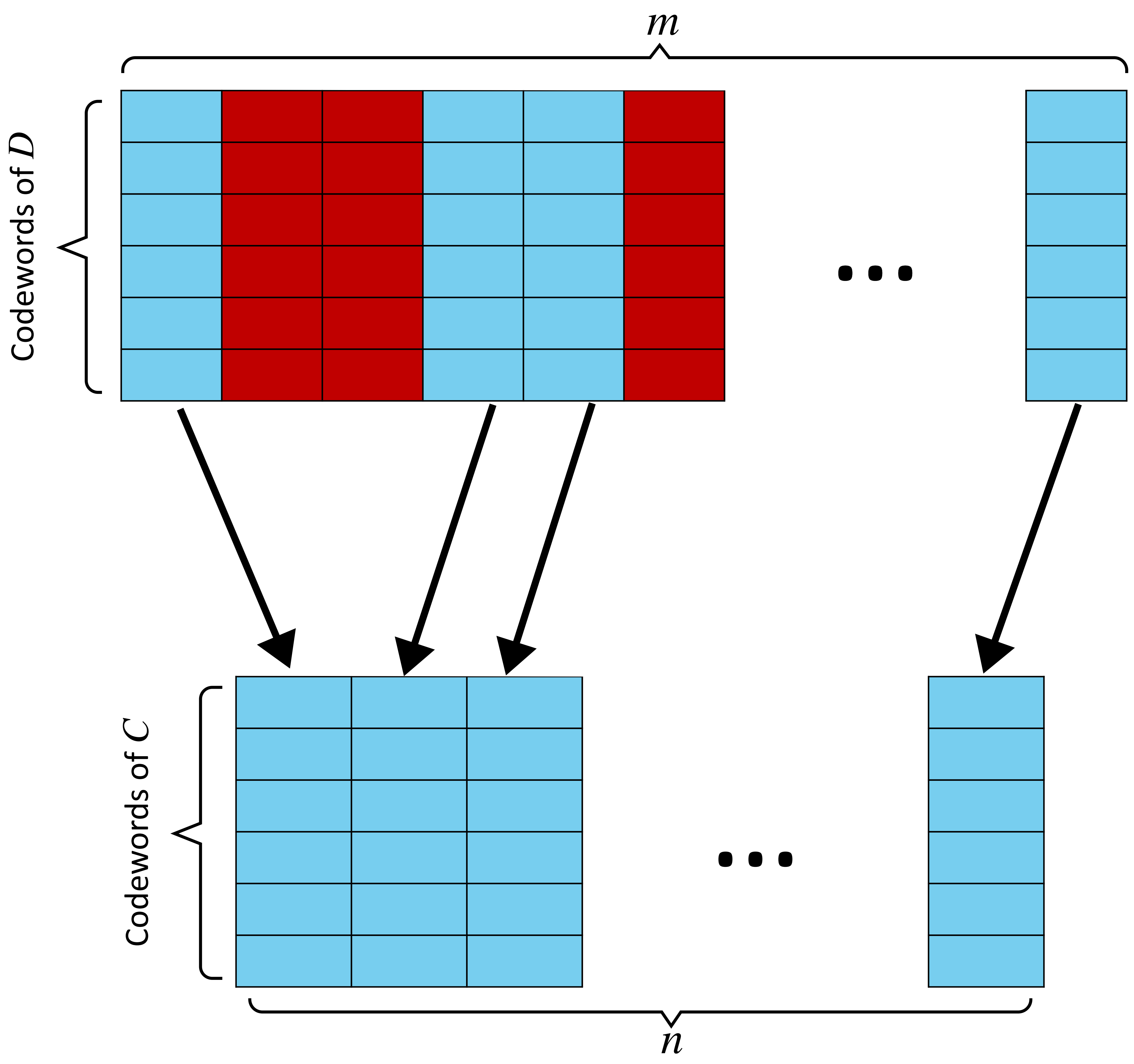
- If the punctured columns are chosen at random, C is said to be a **random n-puncturing of D**.

- **Example:** An **RLC** of rate R in $\mathbb{F}_q^n$ is a **random puncturing** of the **Hadamard code** $H \subseteq \mathbb{F}_q^{q^{Rn}}$.

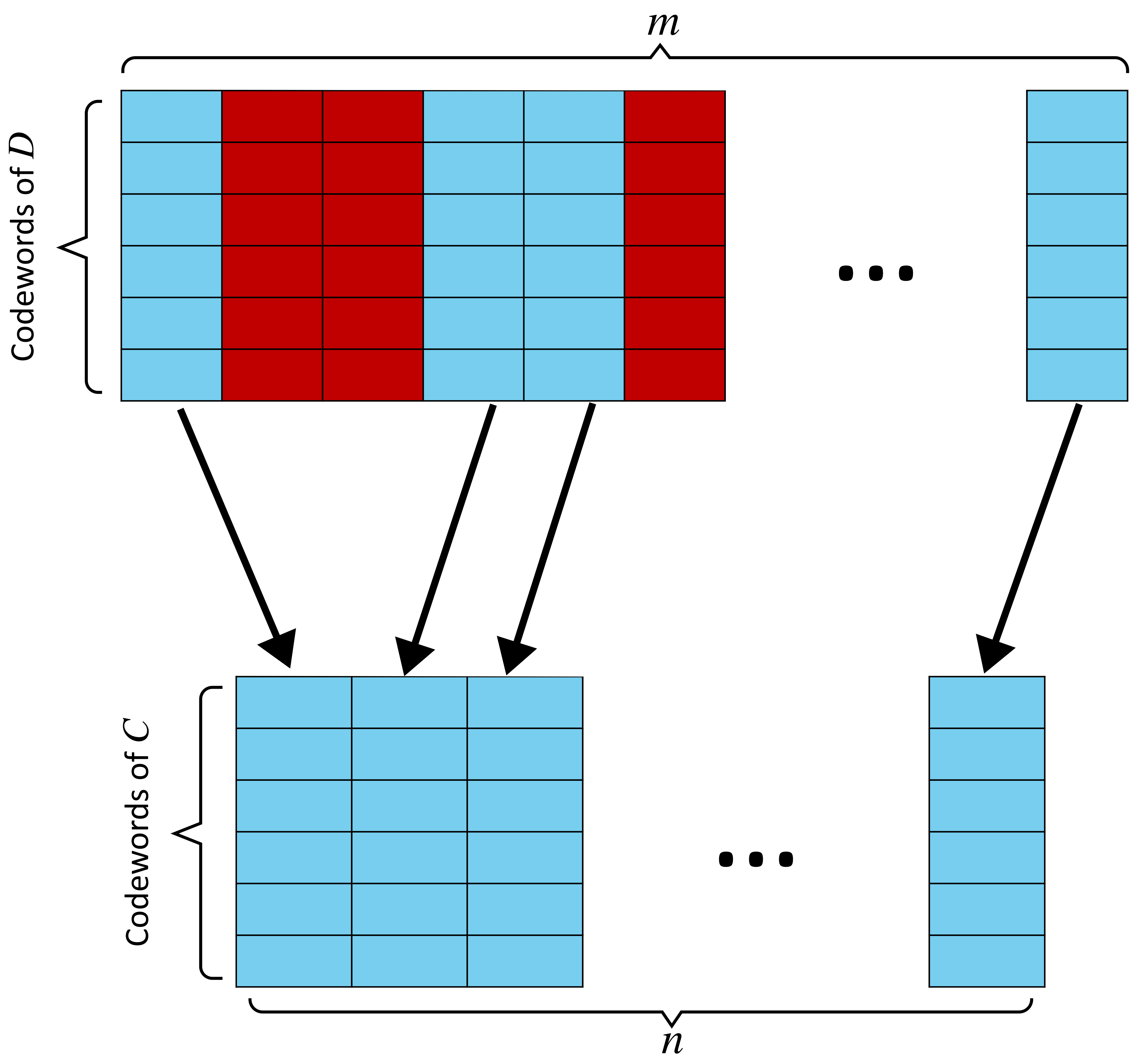


Puncturing of Codes

- From a code $D \subseteq \mathbb{F}_q^m$ create a new code $C \subseteq \mathbb{F}_q^n$. Usually $n \ll m$.
- If the punctured columns are chosen at random, C is said to be a **random n-puncturing of D**.
- **Example:** An **RLC** of rate R in $\mathbb{F}_q^n$ is a **random puncturing** of the **Hadamard code** $H \subseteq \mathbb{F}_q^{q^{Rn}}$.
- A **Reed-Solomon code over a random evaluation set** is a **random puncturing** of the **full Reed-Solomon code.**

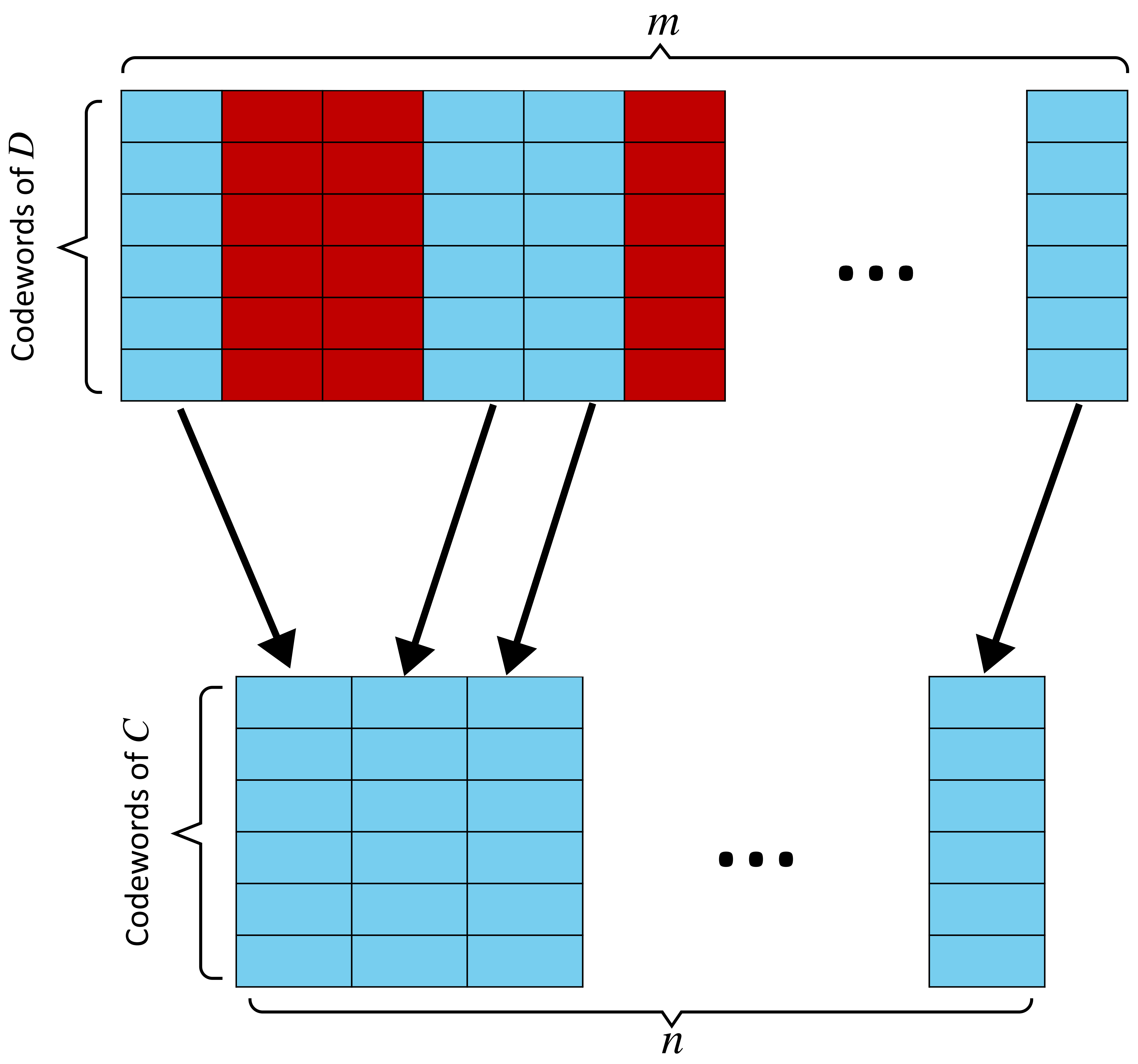


Puncturing of low-bias codes



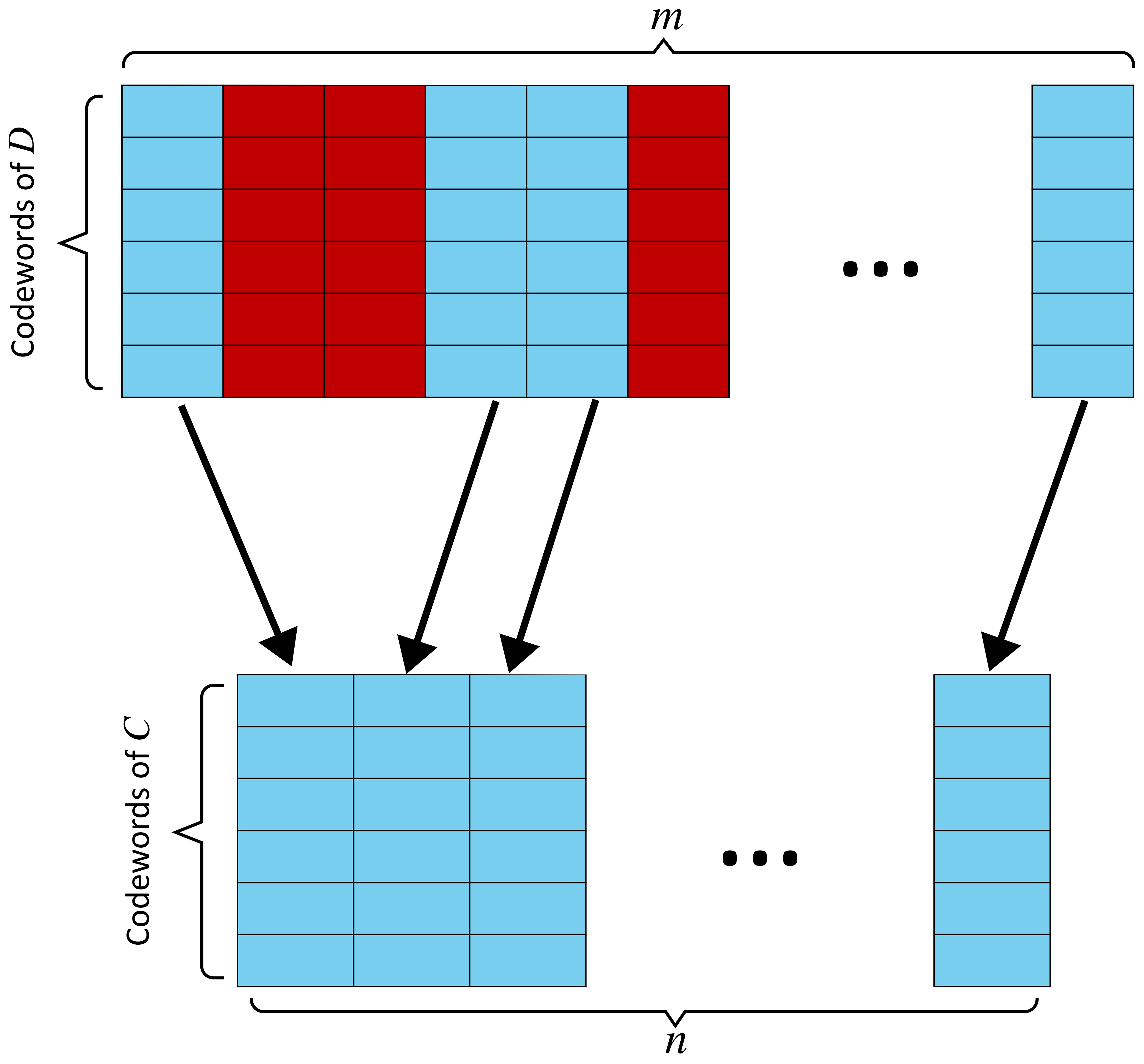
Puncturing of low-bias codes

- Let's focus on $q = 2$



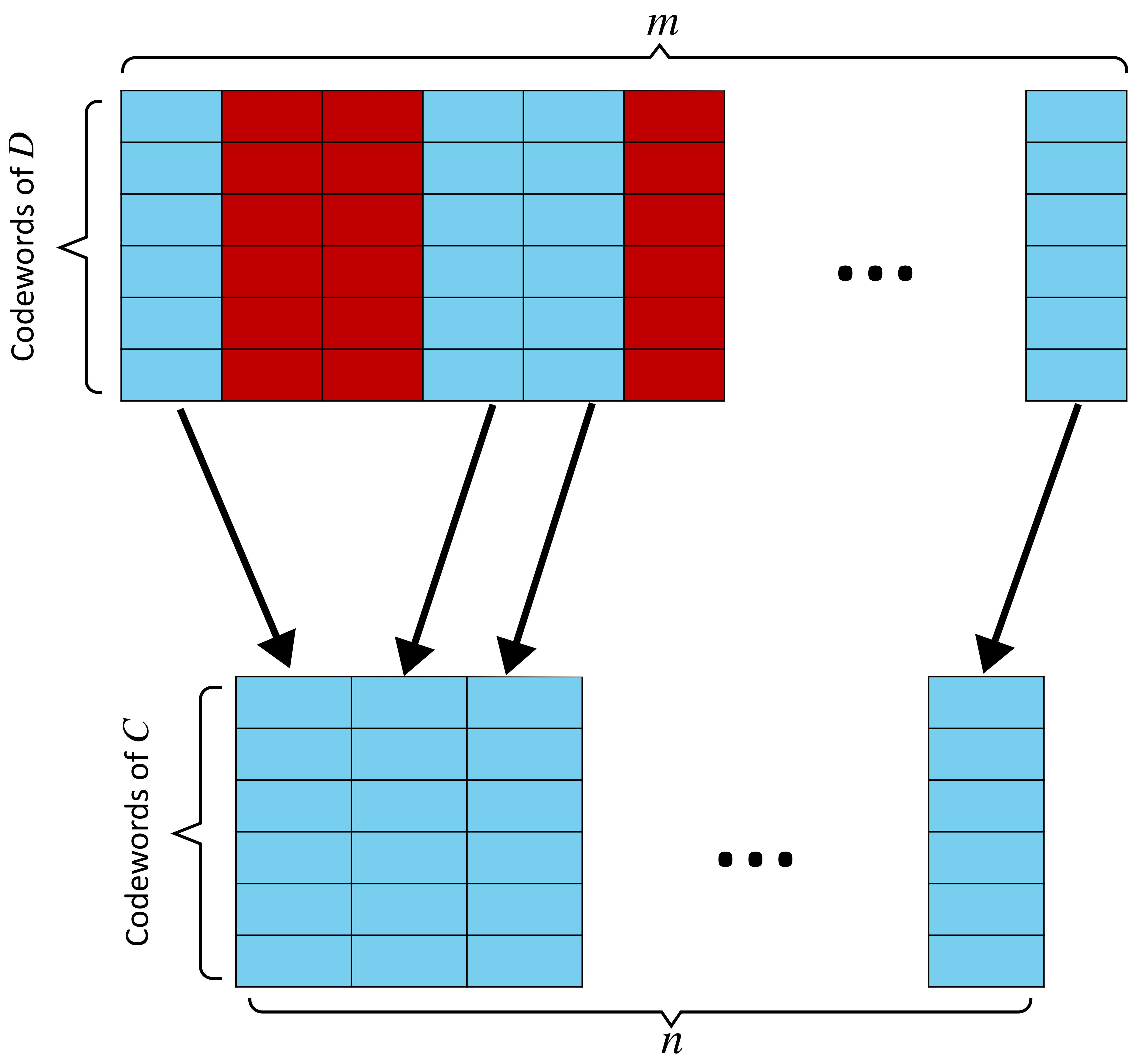
Puncturing of low-bias codes

- Let's focus on $q = 2$
- Suppose every $u \in D$ has weight close to $\dfrac{m}{2}$ (low-bias).



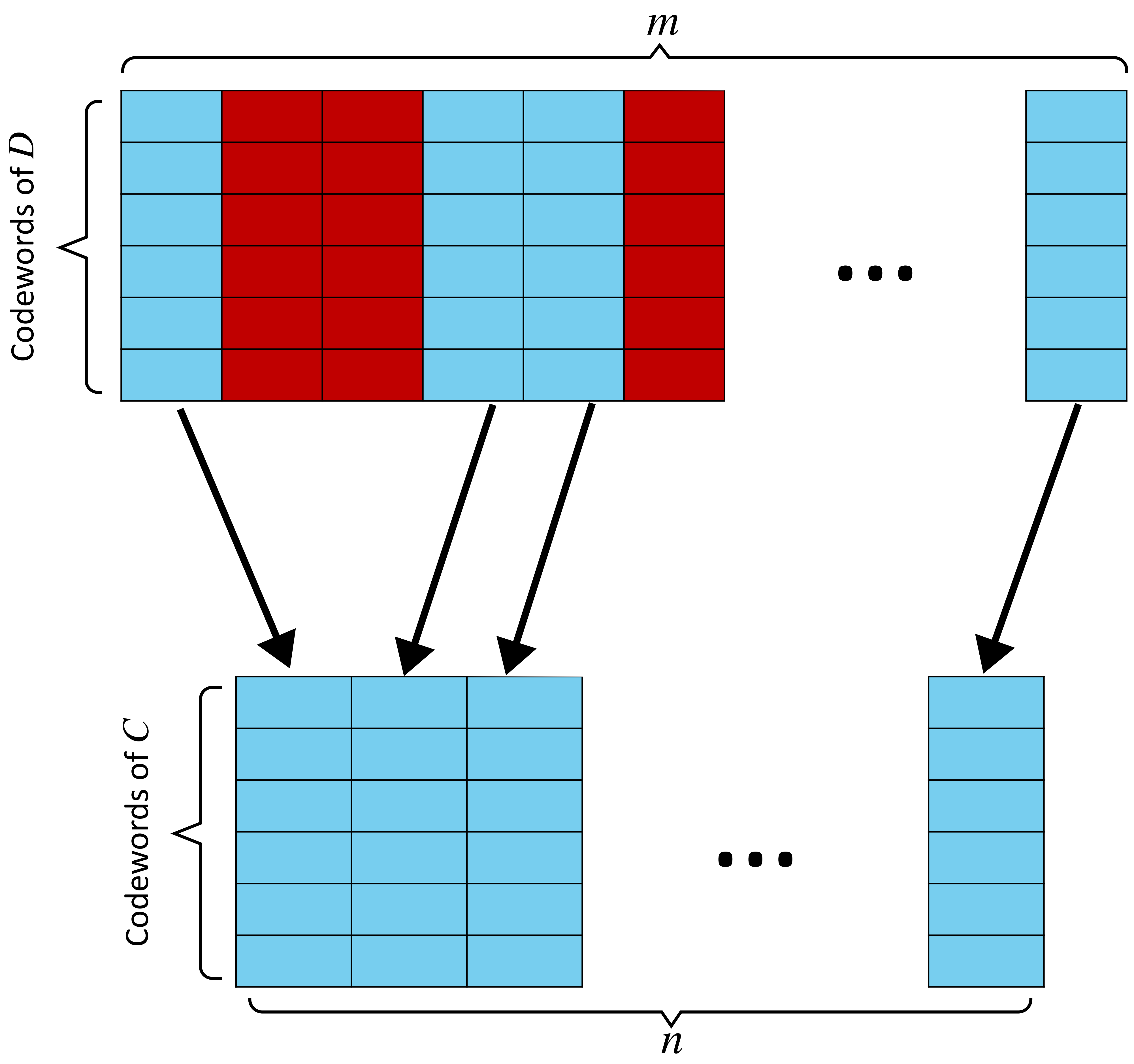
Puncturing of low-bias codes

- Let's focus on $q = 2$
- Suppose every $u \in D$ has weight close to $\dfrac{m}{2}$ (low-bias).
- D can be, e.g., a **dual-BCH** code.



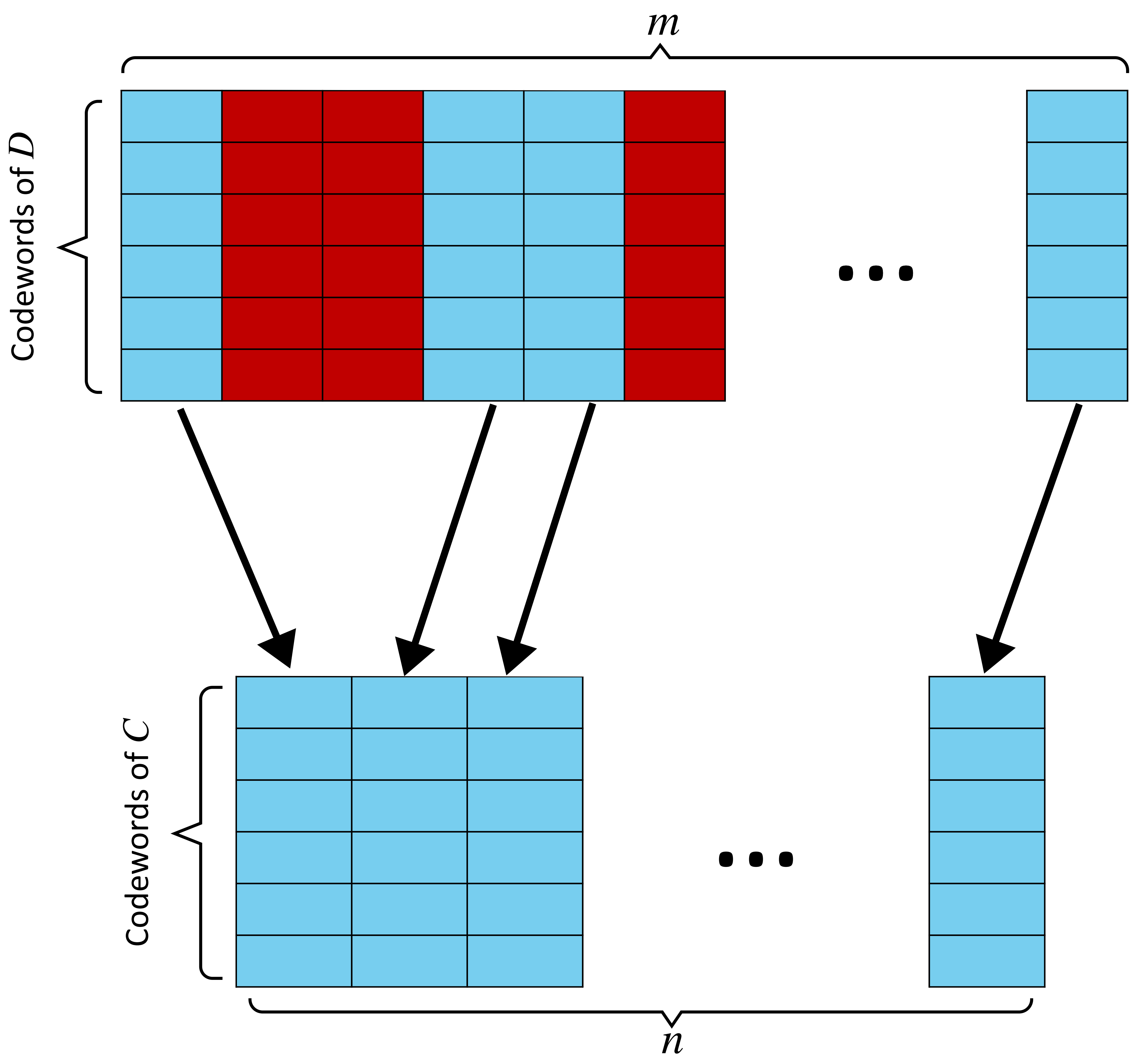
Puncturing of low-bias codes

- Let's focus on $q = 2$
- Suppose every $u \in D$ has weight close to $\dfrac{m}{2}$ (low-bias).
- D can be, e.g., a **dual-BCH** code.
- **Theorem:** C is **locally-similar** to an **RLC**.



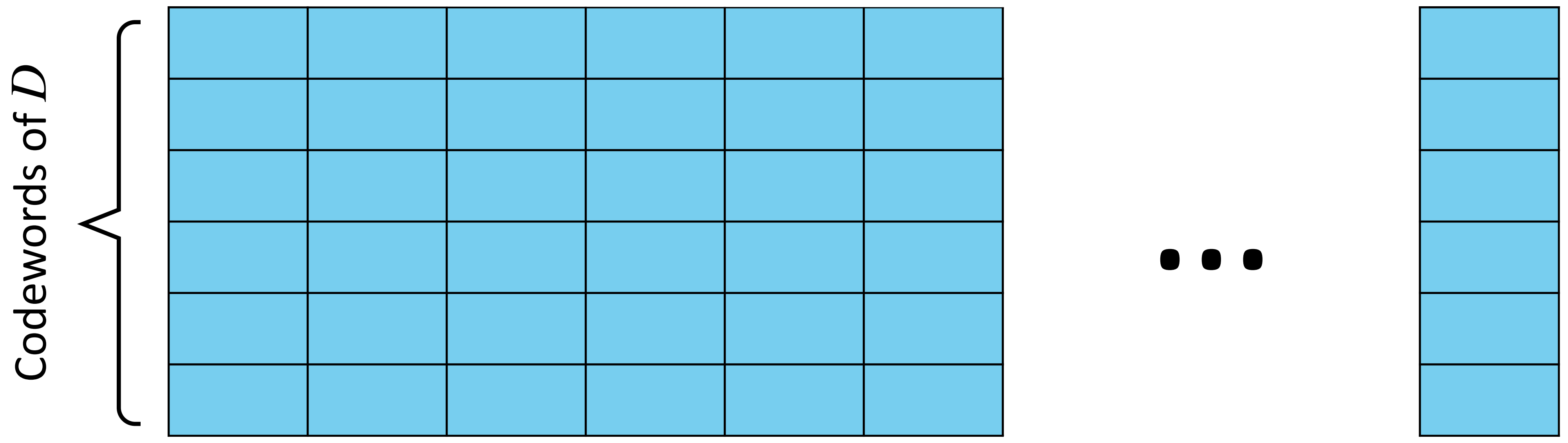
Puncturing of low-bias codes

- Let's focus on $q = 2$
- Suppose every $u \in D$ has weight close to $\dfrac{m}{2}$ (low-bias).
- D can be, e.g., a **dual-BCH** code.
- **Theorem:** C is **locally-similar** to an **RLC**.
- **Corollary:** C is as **list-decodable** and **list-recoverable** as an **RLC**.

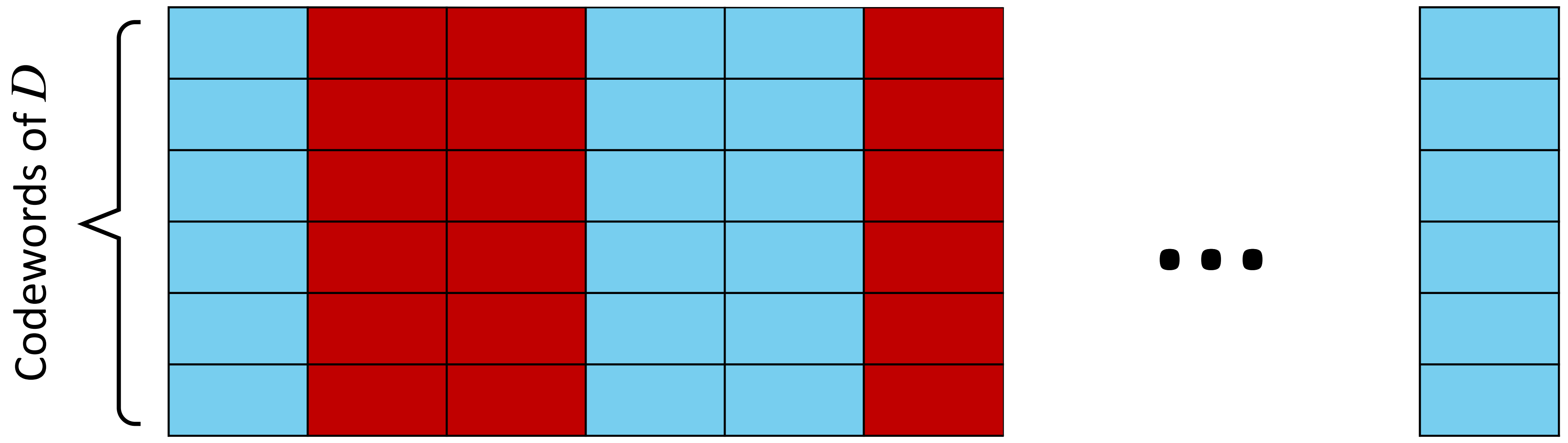


Proof sketch: C is locally-similar to an RLC.

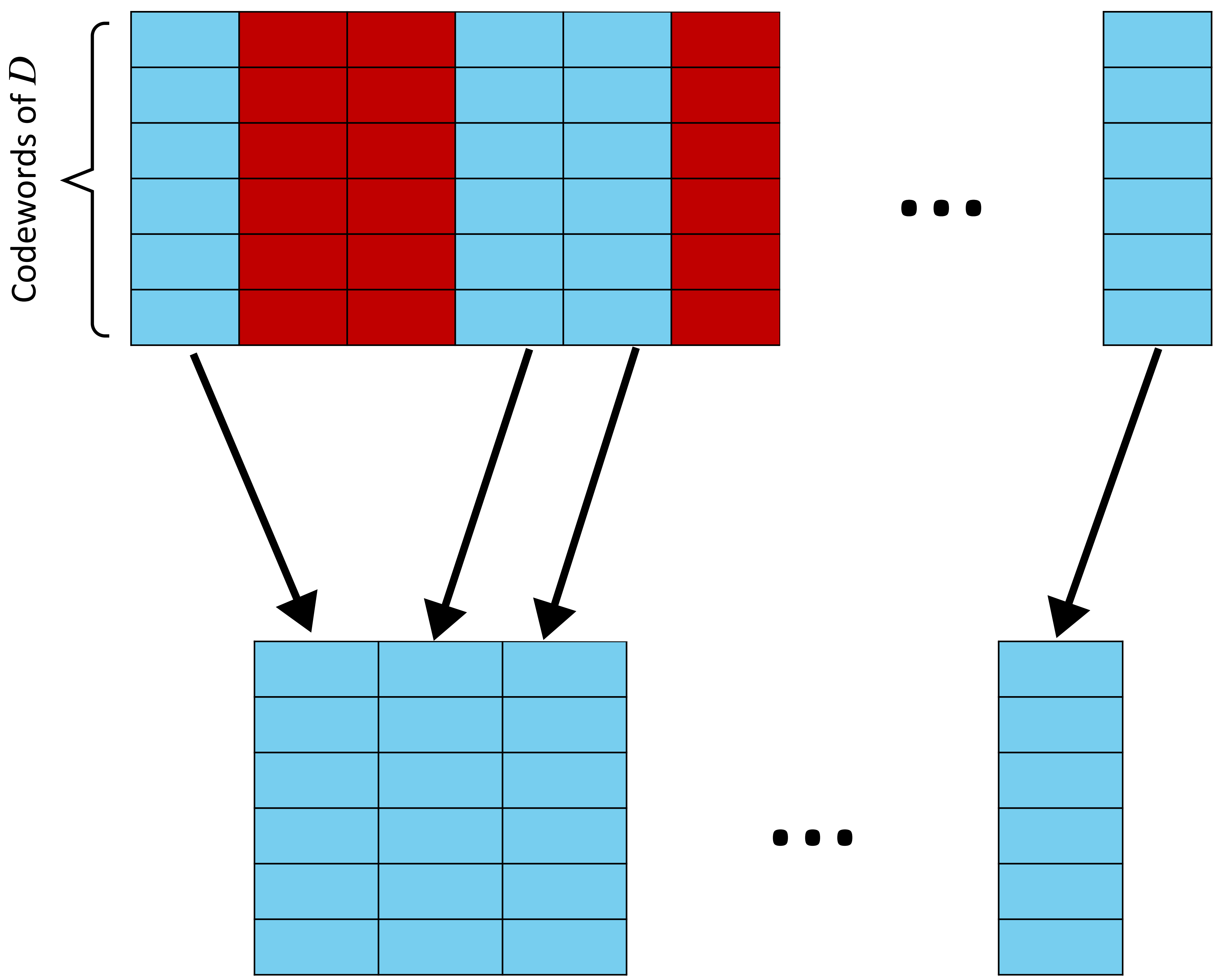
Proof sketch: C is locally-similar to an RLC.



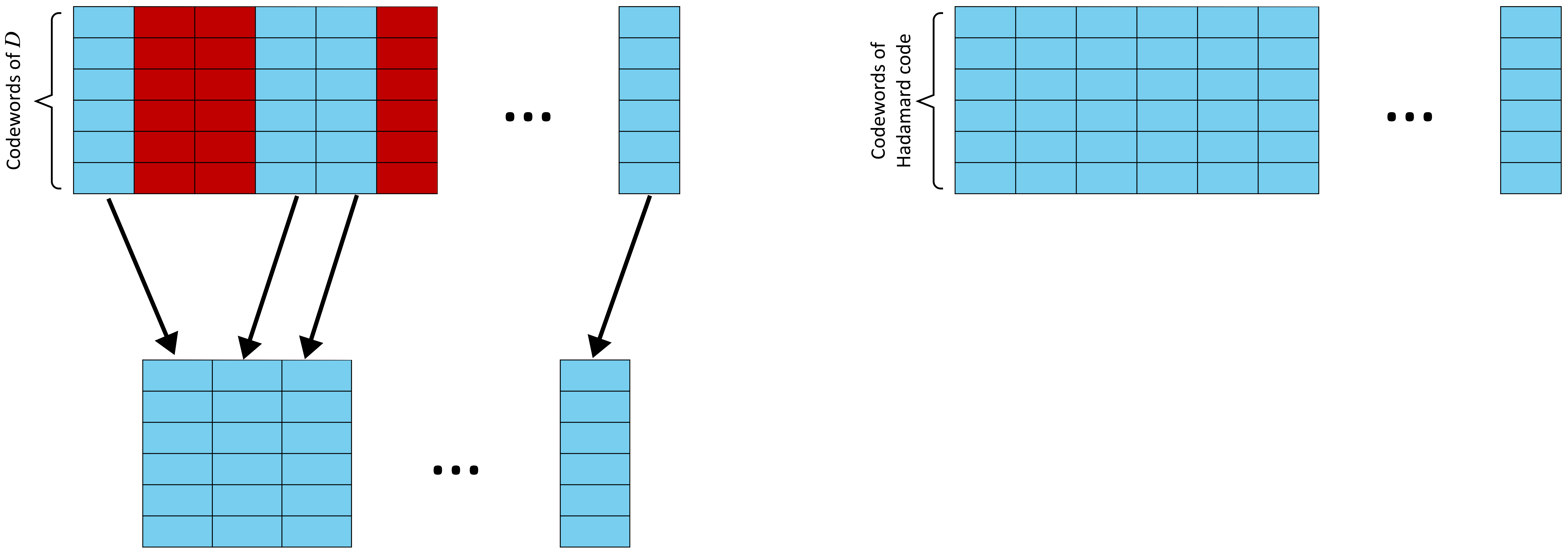
Proof sketch: C is locally-similar to an RLC.



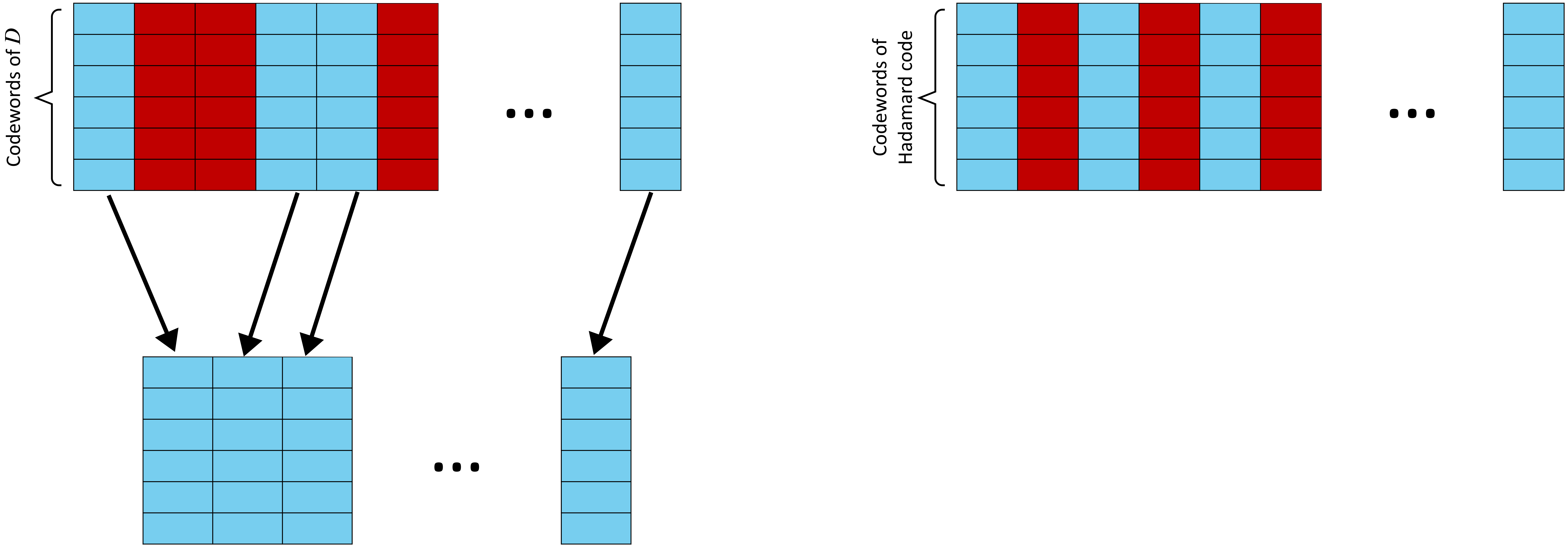
Proof sketch: C is locally-similar to an RLC.



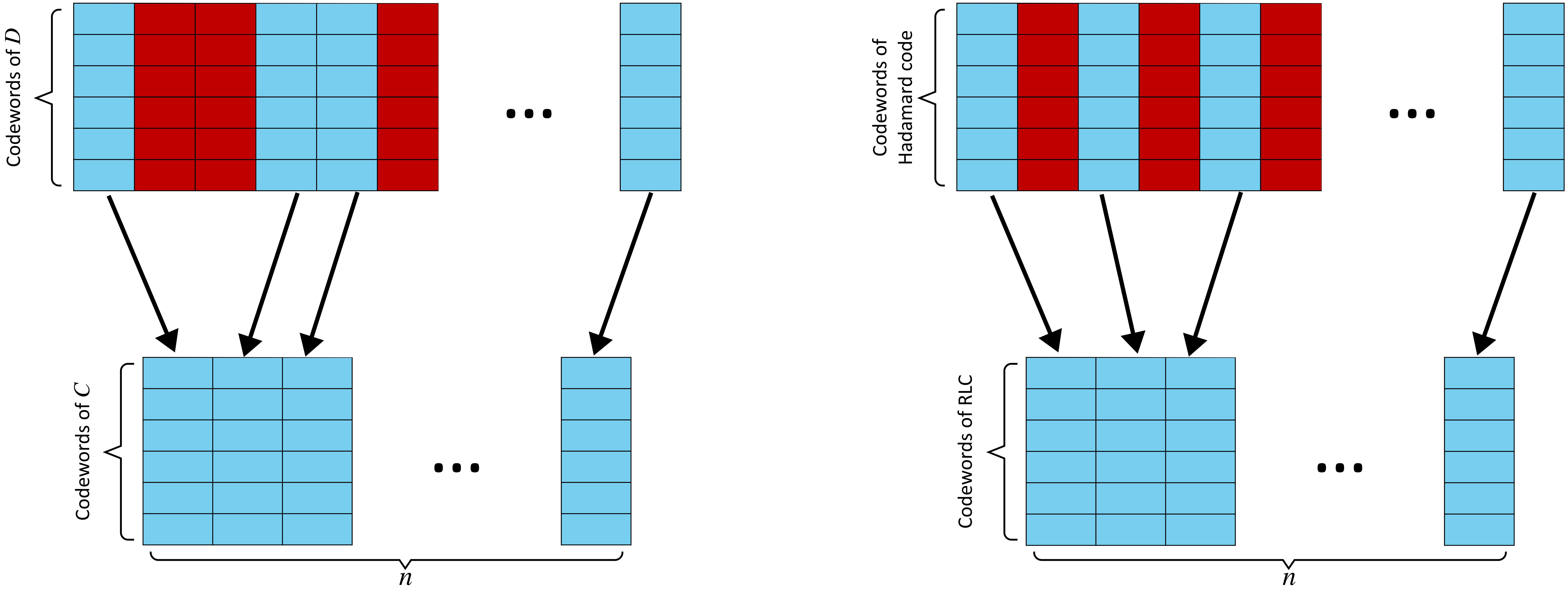
Proof sketch: C is locally-similar to an RLC.



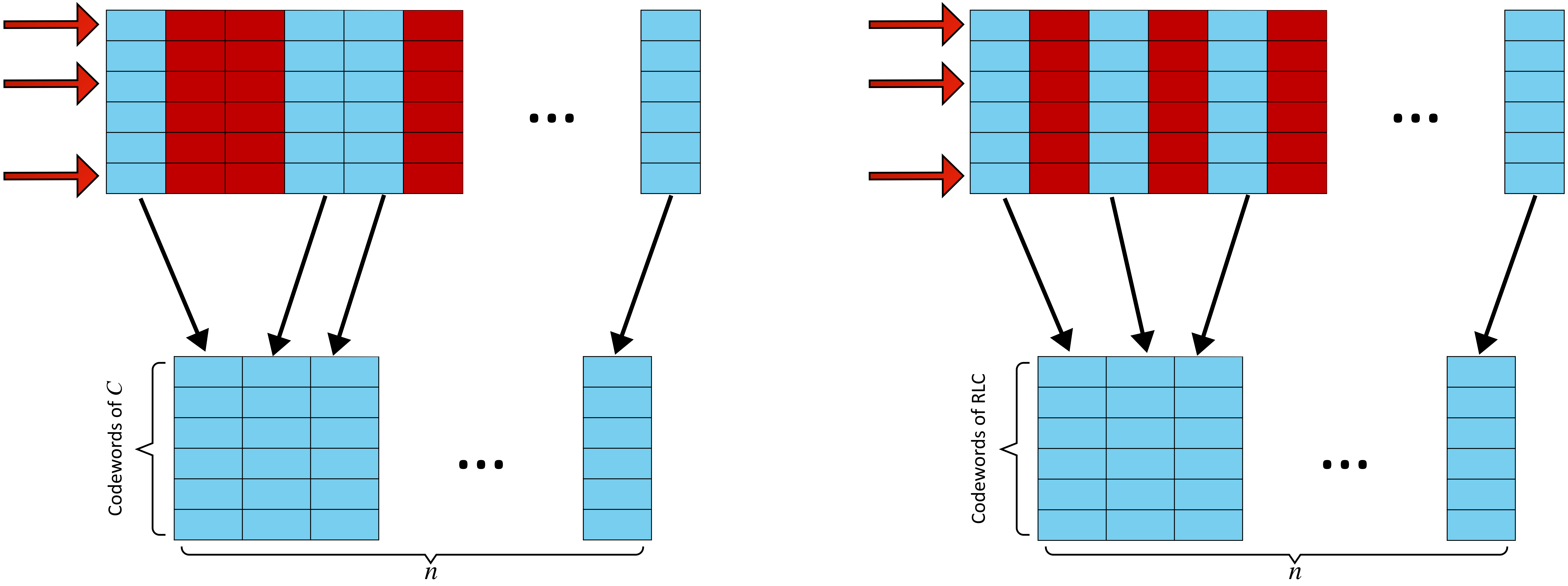
Proof sketch: C is locally-similar to an RLC.



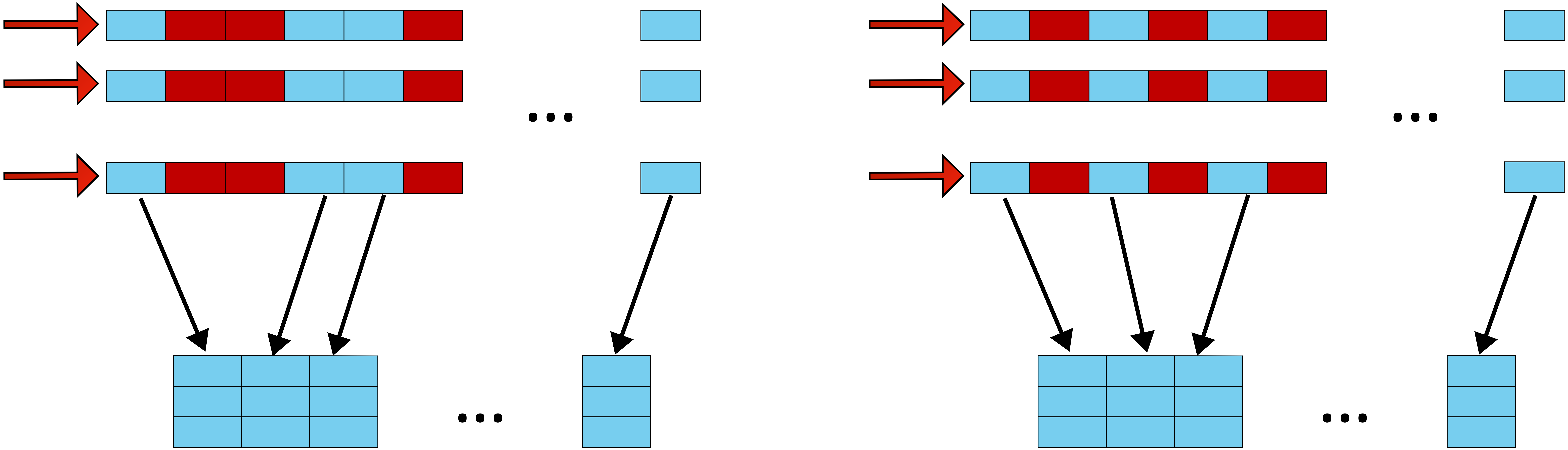
Proof sketch: C is locally-similar to an RLC.



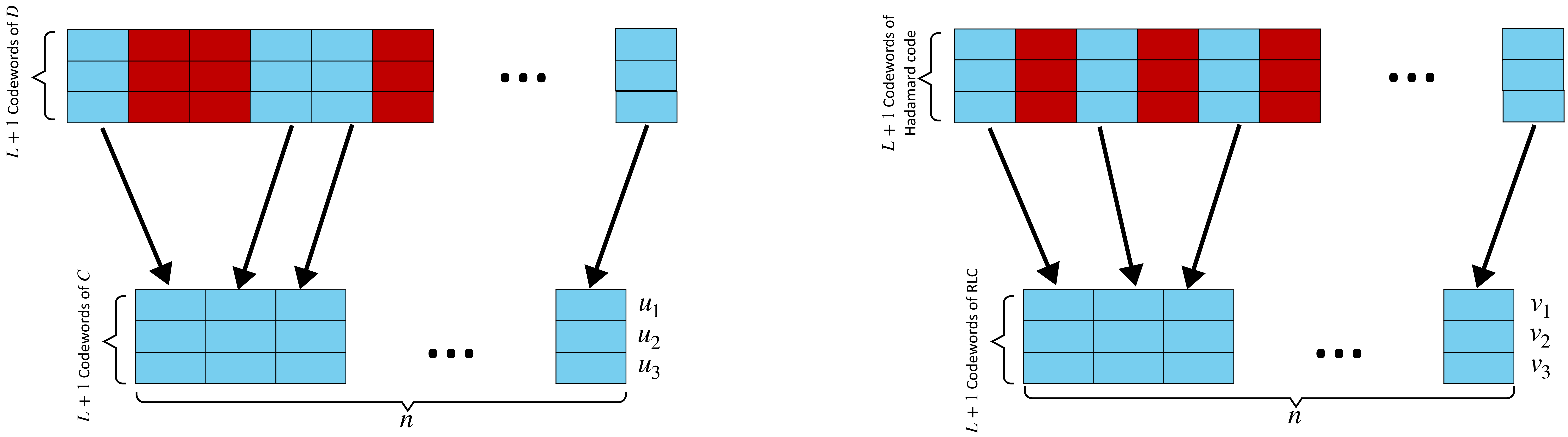
Proof sketch: C is locally-similar to an RLC.



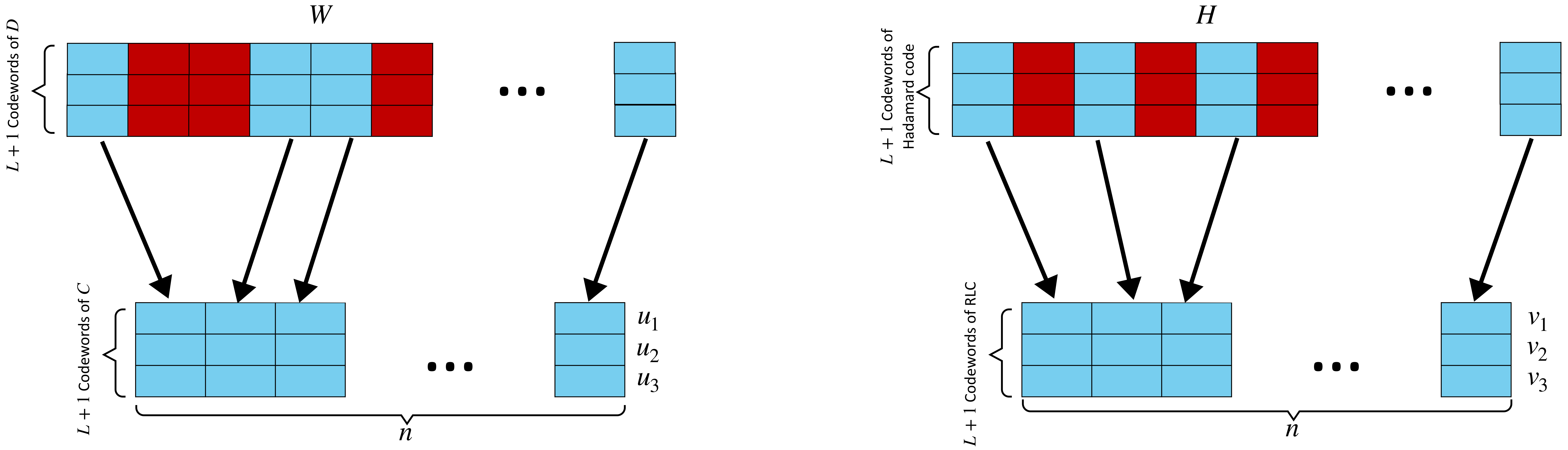
Proof sketch: C is locally-similar to an RLC.



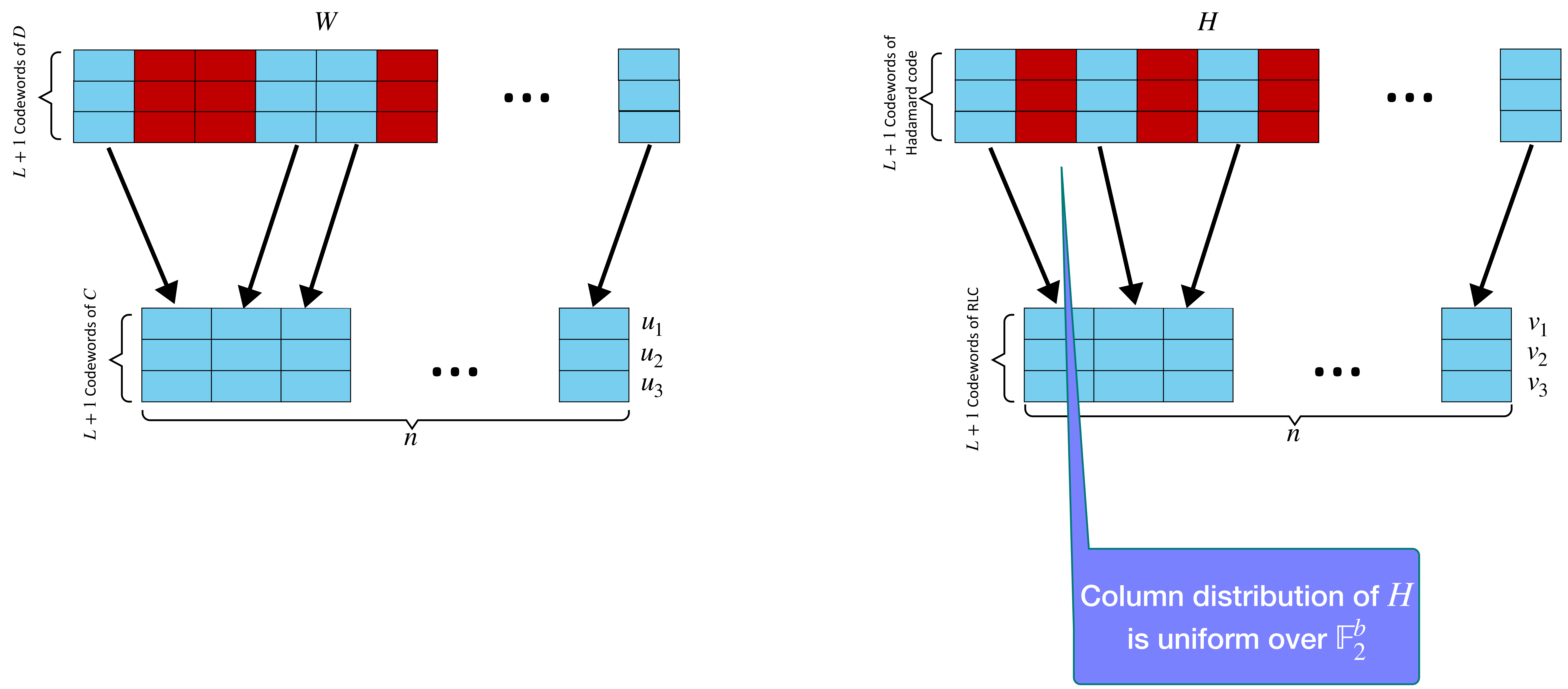
Proof sketch: C is locally-similar to an RLC.



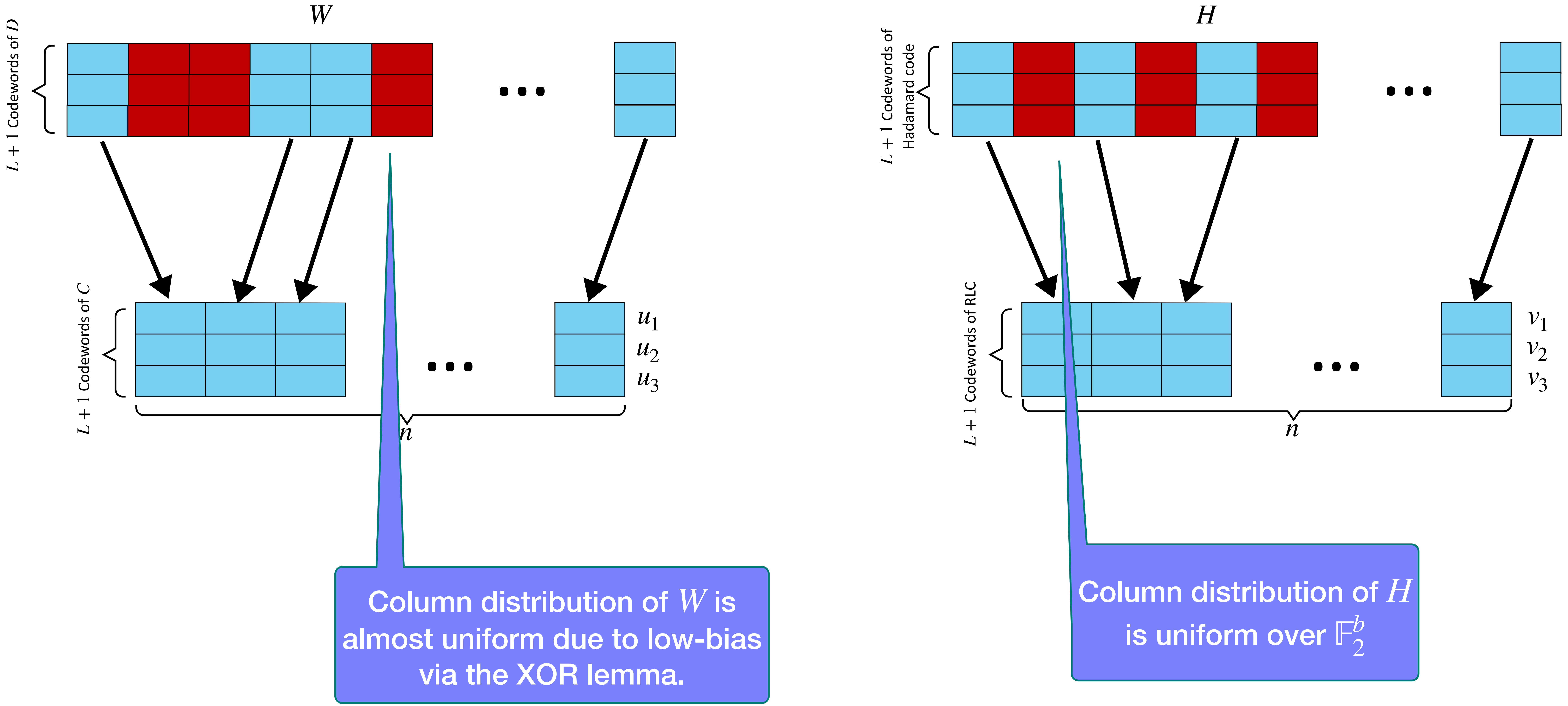
Proof sketch: C is locally-similar to an RLC.



Proof sketch: C is locally-similar to an RLC.



Proof sketch: C is locally-similar to an RLC.



Drawbacks of the method

Drawbacks of the method

Locality is necessary

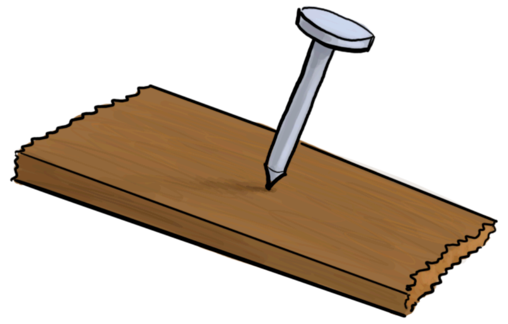
Drawbacks of the method

Locality is necessary

Open problem:

Let $C \subseteq \mathbb{F}_q^n$ be an **RLC** and fix $\epsilon > 0$.

Prove that C is $\left(\frac{q}{2},\ q^{Rn} \cdot 2^{-n} \cdot (1+\epsilon)\right)$**-list-recoverable** with high probability.



Drawbacks of the method

Can only deal with “Σ_1” properties.

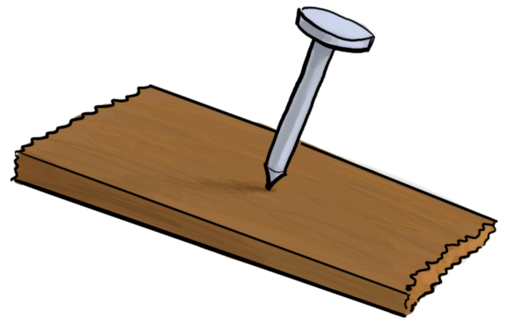
Drawbacks of the method

Can only deal with "Σ_1" properties.

Open problem:

Say that a code C is (ρ, L)-**covering** if every $x \in \mathbb{F}_2^n$ is ρ-**close** to at least L codewords of C.

Find the rate threshold for (ρ, L)-**covering** with regard to **RLCs**.



Drawbacks of the method

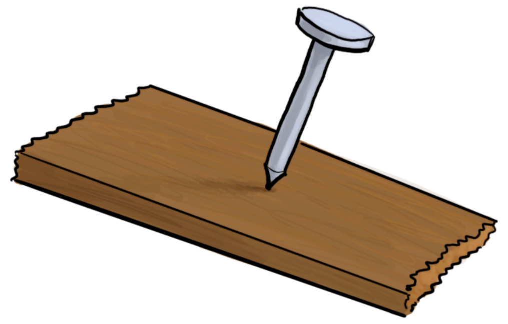
Local-similarity to RLC requires $\Omega(n)$ random bits

Drawbacks of the method

Local-similarity to RLC requires $\Omega(n)$ random bits

Open problem:

Construct a code **achieving the GV bound** with $o(n)$ random bits.



Drawbacks of the method

Alphabet cannot be large.

Drawbacks of the method

Alphabet cannot be large.

- Recall that the number of possible row distributions for a matrix in $\mathbb{F}_2^L$ is roughly n^{2^L}. We need to union bound over this.

Drawbacks of the method

Alphabet cannot be large.

- Recall that the number of possible row distributions for a matrix in $\mathbb{F}_2^L$ is roughly n^{2^L}. We need to union bound over this.
- For general q this is n^{q^L}.

Drawbacks of the method

Alphabet cannot be large.

- Recall that the number of possible row distributions for a matrix in $\mathbb{F}_2^L$ is roughly n^{2^L}. We need to union bound over this.
- For general q this is n^{q^L}.
- Suppose $q = n$, then **there are n^{n^L} types!** 😱

Drawbacks of the method

Alphabet cannot be large.

Drawbacks of the method

Alphabet cannot be large.

- Is there any hope for reasoning about **Reed-Solomon** codes with this method?

Reed-Solomon codes

Reed-Solomon codes

- A **Reed-Solomon (RS) code** over $\mathbb{F}_q$ is defined by:
  - A **rank** $1 \leq k \leq q$
  - An **evaluation set** $S \subseteq \mathbb{F}_q$.

Reed-Solomon codes

- A **Reed-Solomon (RS) code** over $\mathbb{F}_q$ is defined by:
    - A **rank** $1 \leq k \leq q$
    - An **evaluation set** $S \subseteq \mathbb{F}_q$.
- The codewords are $\big(p(x)\big)_{x \in S}$ where $p \in \mathbb{F}_q[x]$ has degree $< k$.

Reed-Solomon codes

- A **Reed-Solomon (RS) code** over $\mathbb{F}_q$ is defined by:
  - A **rank** $1 \leq k \leq q$
  - An **evaluation set** $S \subseteq \mathbb{F}_q$.
- The codewords are $\big(p(x)\big)_{x \in S}$ where $p \in \mathbb{F}_q[x]$ has degree $< k$.
- We denote $\mathrm{RS}[S, k]$.

Reed-Solomon codes

- A **Reed-Solomon (RS) code** over $\mathbb{F}_q$ is defined by:
  - A **rank** $1 \leq k \leq q$
  - An **evaluation set** $S \subseteq \mathbb{F}_q$.
- The codewords are $\big(p(x)\big)_{x \in S}$ where $p \in \mathbb{F}_q[x]$ has degree $< k$.
- We denote $\mathrm{RS}[S, k]$.
- The code has dimension k and length $n = |S|$, so $R = \dfrac{k}{n}$.

Reed-Solomon codes

- A **Reed-Solomon (RS) code** over $\mathbb{F}_q$ is defined by:
    - A **rank** $1 \leq k \leq q$
    - An **evaluation set** $S \subseteq \mathbb{F}_q$.
- The codewords are $\big(p(x)\big)_{x \in S}$ where $p \in \mathbb{F}_q[x]$ has degree $< k$.
- We denote $\mathrm{RS}[S, k]$.
- The code has dimension k and length $n = |S|$, so $R = \dfrac{k}{n}$.
- Note that $n \leq q$.

List-Decodability of RS codes

List-Decodability of RS codes

Problem:

Are there **RS codes** that **achieve the list-decoding GV-bound**?

How large does q need to be in terms of n?

List-Decodability of RS codes

List-Decodability of RS codes

- Many works about **list-decodability** of $\mathrm{RS}[S,k]$ where $S \subseteq \mathbb{F}_q$ is random ("**random RS code**")
  - **[Rudra-Wootters], [Shangguan-Tamo], [Goldberg-Shangguan-Tamo][Guo-Li-Shangguan-Tamo-Wootters], [Ferber-Kwan-Sauermann], [Brakensiek-Gopi-Makam], [Guo-Zhang], [Alrabiah-Guruswami-Li].**

List-Decodability of RS codes

- Many works about **list-decodability** of $\mathrm{RS}[S, k]$ where $S \subseteq \mathbb{F}_q$ is random ("**random RS code**")
  - **[Rudra-Wootters], [Shangguan-Tamo], [Goldberg-Shangguan-Tamo][Guo-Li-Shangguan-Tamo-Wootters], [Ferber-Kwan-Sauermann], [Brakensiek-Gopi-Makam], [Guo-Zhang], [Alrabiah-Guruswami-Li].**
- Most recently:
  - **[BGM]** - **List-decoding GV-bound** with $q = 2^{O(n)}$
  - **[GZ]** - **List-decoding GV-bound** with $q = O(n^2)$
  - **[AGL]** - **List-decoding GV-bound** with $q = O(n)$

List-Decodability of RS codes

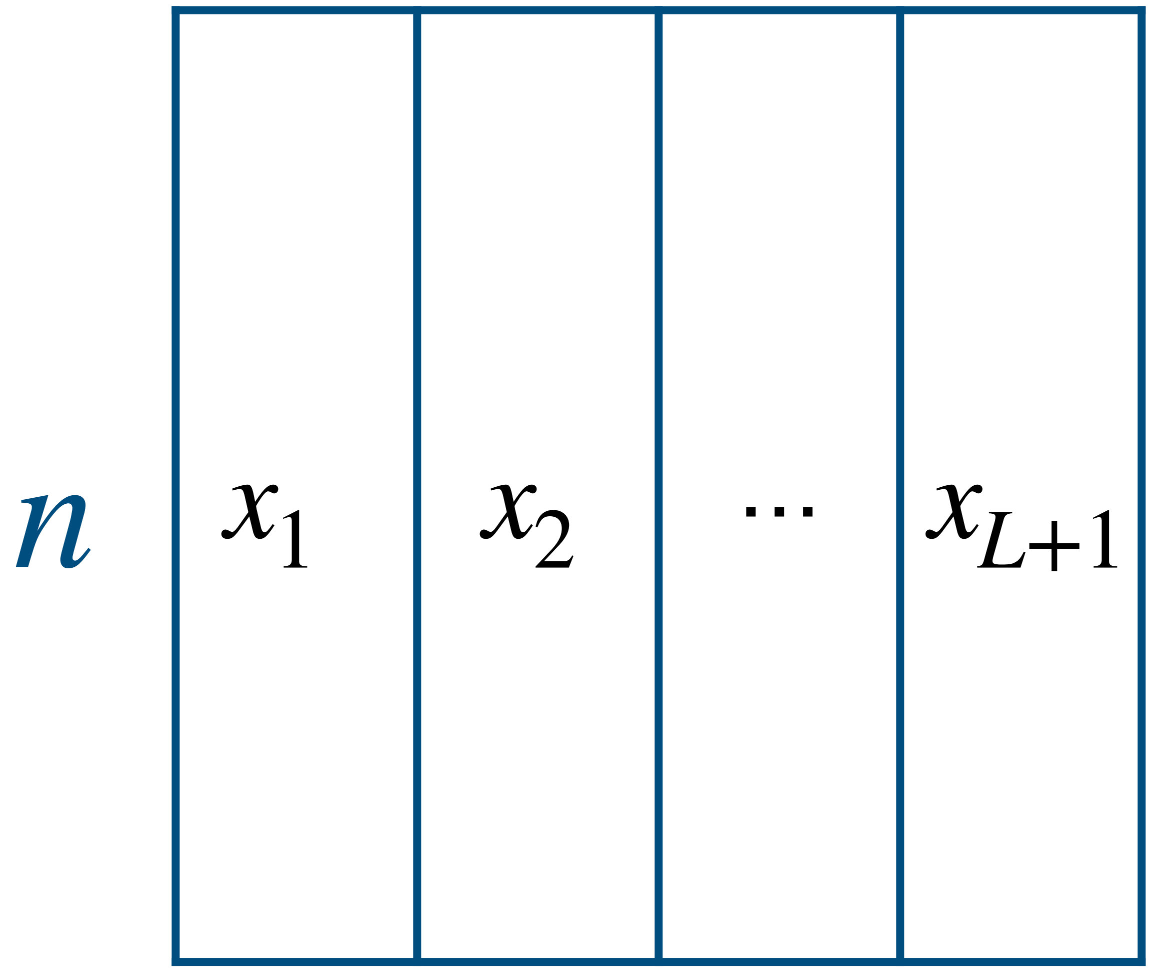
- Many works about **list-decodability** of $\mathrm{RS}[S, k]$ where $S \subseteq \mathbb{F}_q$ is random ("**random RS code**")
  - **[Rudra-Wootters], [Shangguan-Tamo], [Goldberg-Shangguan-Tamo][Guo-Li-Shangguan-Tamo-Wootters], [Ferber-Kwan-Sauermann], [Brakensiek-Gopi-Makam], [Guo-Zhang], [Alrabiah-Guruswami-Li].**
- Most recently:
  - **[BGM]** - **List-decoding GV-bound** with $q = 2^{O(n)}$
  - **[GZ]** - **List-decoding GV-bound** with $q = O(n^2)$
  - **[AGL]** - **List-decoding GV-bound** with $q = O(n)$
- Less is known for **list-recovery**.

List-Decodability of RS codes

- Many works about **list-decodability** of $\mathrm{RS}[S,k]$ where $S \subseteq \mathbb{F}_q$ is random ("**random RS code**")
  - **[Rudra-Wootters], [Shangguan-Tamo], [Goldberg-Shangguan-Tamo][Guo-Li-Shangguan-Tamo-Wootters], [Ferber-Kwan-Sauermann], [Brakensiek-Gopi-Makam], [Guo-Zhang], [Alrabiah-Guruswami-Li].**
- Most recently:
  - **[BGM]** - **List-decoding GV-bound** with $q = 2^{O(n)}$
  - **[GZ]** - **List-decoding GV-bound** with $q = O(n^2)$
  - **[AGL]** - **List-decoding GV-bound** with $q = O(n)$
- Less is known for **list-recovery**.

Types for large alphabet

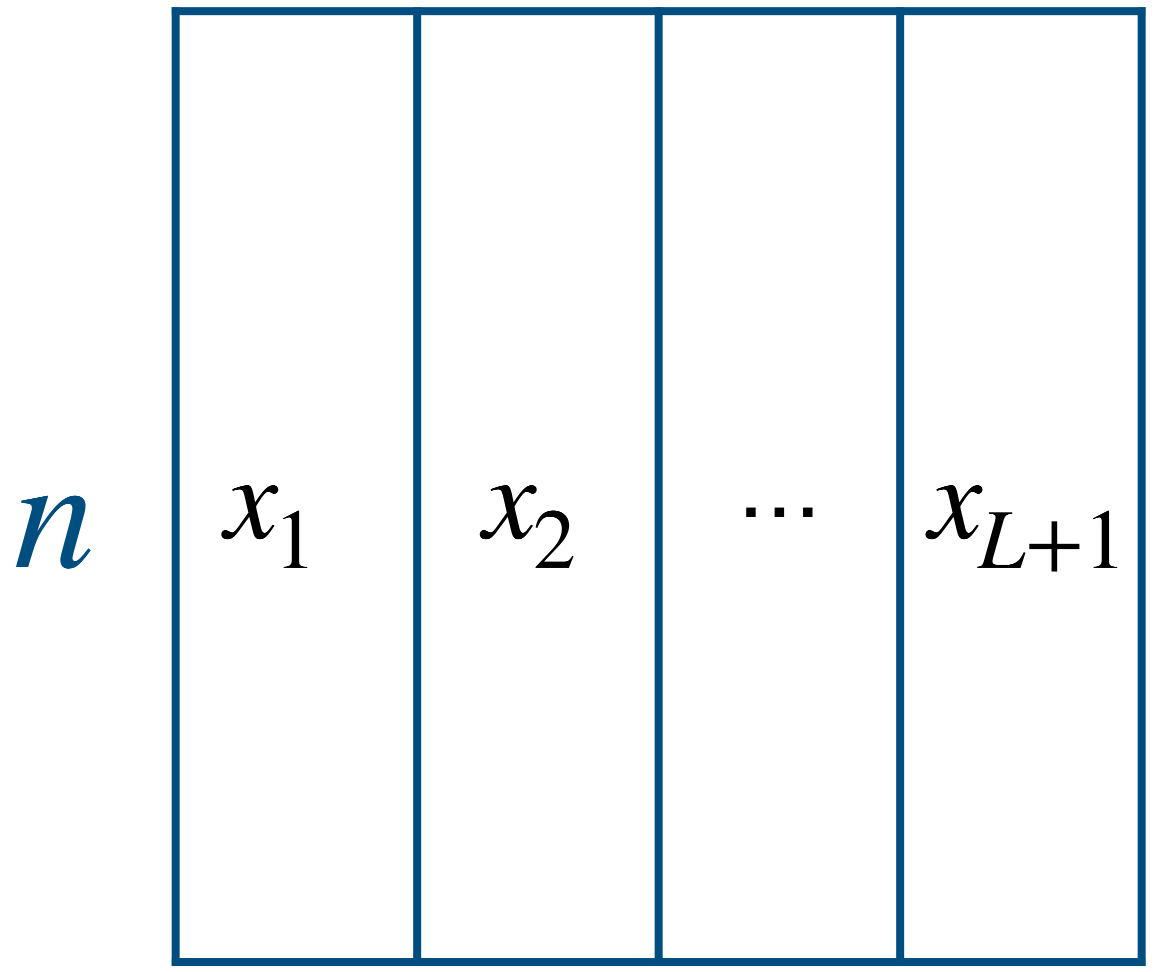
Types for large alphabet



- Suppose the columns of $A \in \mathbb{F}_q^{n\times(L+1)}$ are **ρ-clustered**.

- The row distribution of A contains **too much information**.

Types for large alphabet



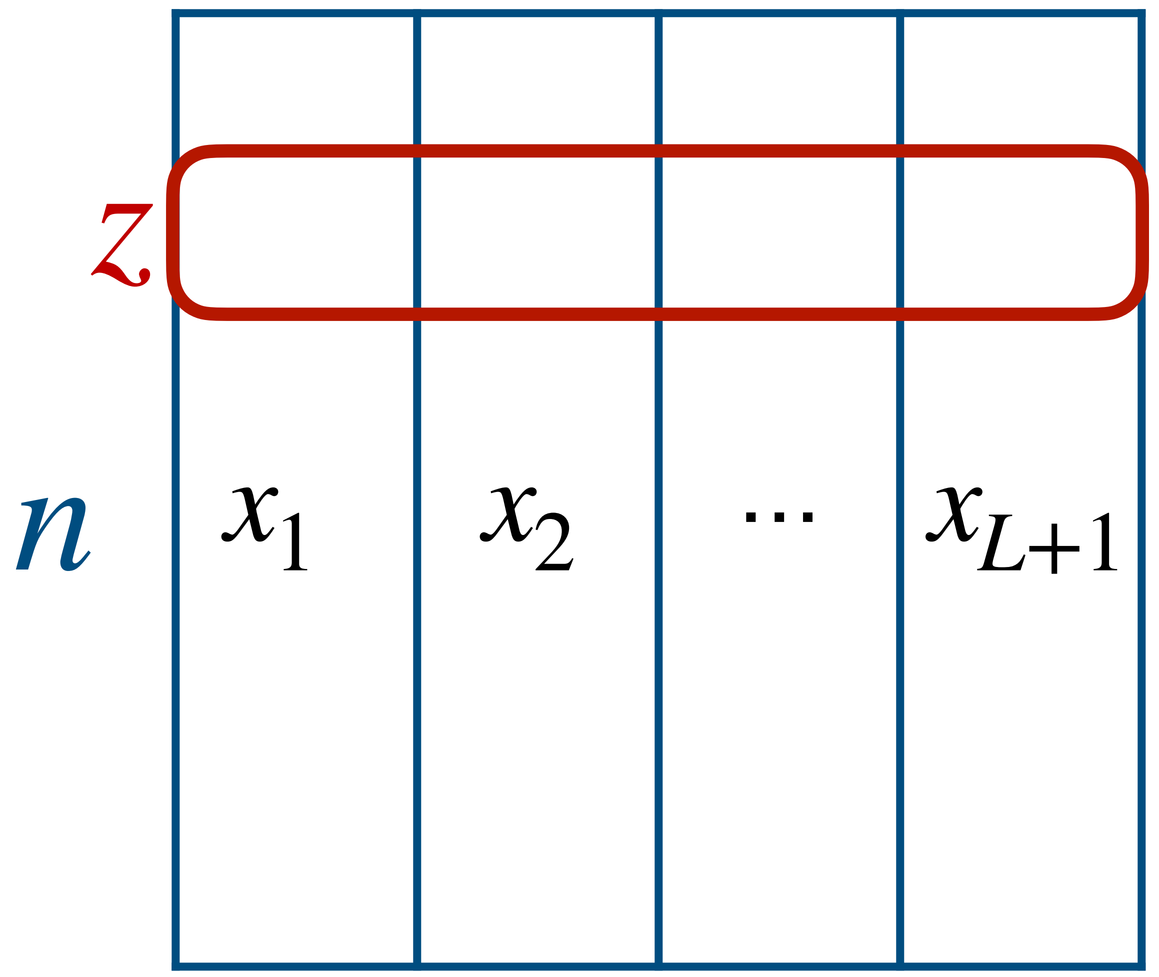
- Suppose the columns of $A \in \mathbb{F}_q^{n\times(L+1)}$ are **ρ-clustered**.
- The row distribution of A contains **too much information**.
- For a given row, we only care about the **identity relation**.

Types for large alphabet

Types for large alphabet

Given $z \in \mathbb{F}_q^{L+1}$ let P_z denote the partition of $\{1,\ldots,L+1\}$ where

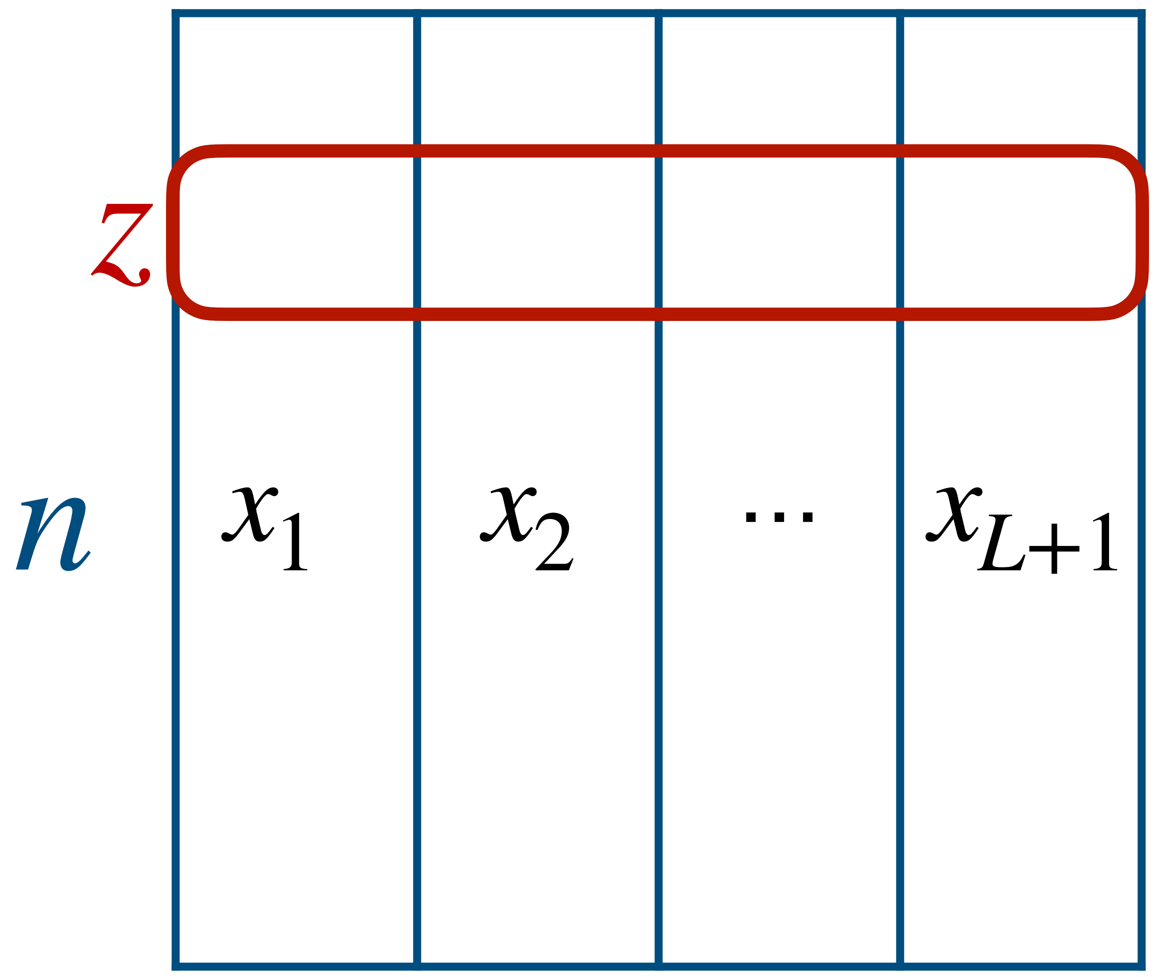
$$i \sim_{P_z} j \quad \Longleftrightarrow \quad z_i = z_j$$



Types for large alphabet

Given $z \in \mathbb{F}_q^{L+1}$ let P_z denote the partition of $\{1,\ldots,L+1\}$ where

$$i \sim_{P_z} j \quad \iff \quad z_i = z_j$$



The **type** of a matrix $A \in \mathbb{F}_q^{n\times(L+1)}$ is a pair consisting of:

1. A list of partitions $\left(P_{A_i}\right)_{i=1}^{n}$

2. The row-span of A.

Types for large alphabet

Types for large alphabet

Observation:

If a matrix A is ρ**-clustered** then so are all matrices of the **same type**.

So the **witnesses for non-list-decodability** are a union of **type classes**.

List-recoverability can also be expressed this way. A property expressible by **type classes** is called a **local identity property.**

$$n \quad \begin{array}{|c|c|c|c|} \hline x_1 & x_2 & \cdots & x_{L+1} \\ \hline \end{array}$$

Types for large alphabet

Types for large alphabet

- How many **types** are there?

Types for large alphabet

- How many **types** are there?
  - There are at most $(L+1)^{L+1}$ equivalence relations.

Types for large alphabet

- How many **types** are there?
  - There are at most $(L+1)^{L+1}$ equivalence relations.
  - So at most $q^{L^2} \cdot (L+1)^{n(L+1)}$ **types**.

Types for large alphabet

- How many **types** are there?
  - There are at most $(L+1)^{L+1}$ equivalence relations.
  - So at most $q^{L^2} \cdot (L+1)^{n(L+1)}$ **types**.
  - For constant L and $q \geq L^{\frac{L}{\epsilon}}$, the above is at most $q^{\epsilon n}$ which is **tiny**!

Types for large alphabet

- How many **types** are there?
    - There are at most $(L+1)^{L+1}$ equivalence relations.
    - So at most $q^{L^2} \cdot (L+1)^{n(L+1)}$ **types**.
    - For constant L and $q \geq L^{\frac{L}{\epsilon}}$, the above is at most $q^{\epsilon n}$ which is **tiny**!
    - We can union bound over the **ρ-clustered types**.

Large alphabet types in RLCs - Intuition

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- Consider the **type** $T = \left(P = \left(P_i\right)_{i=1}^{n}, \mathbb{F}_3^n\right)$

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- Consider the **type** $T = \left(P = (P_i)_{i=1}^{n}, \mathbb{F}_3^n\right)$

- Will an **RLC** of rate R **contain a matrix of type T**?

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- Consider the **type** $T = \left(P = \left(P_i\right)_{i=1}^{n}, \mathbb{F}_3^n\right)$
- Will an **RLC** of rate R **contain a matrix of type T**?
- There are q^{3Rn} triplets x_1, x_2, x_3 of words in C.

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- Consider the **type** $T = \left(P = \left(P_i\right)_{i=1}^{n}, \mathbb{F}_3^n\right)$

- Will an **RLC** of rate R **contain a matrix of type T**?

- There are q^{3Rn} triplets x_1, x_2, x_3 of words in C.

- Each P_i imposes $3 - |P_i|$ **linear conditions**.

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- Consider the **type** $T = \left(P = \left(P_i\right)_{i=1}^{n}, \mathbb{F}_3^n\right)$

- Will an **RLC** of rate R **contain a matrix of type** T?

- There are q^{3Rn} triplets x_1, x_2, x_3 of words in C.

- Each P_i imposes $3 - |P_i|$ **linear conditions**.

- Let $\deg(P, \mathbb{F}_q^3) = 3Rn - \sum_i (3 - |P_i|)$.

	3
n	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- Consider the **type** $T = \left(P = \left(P_i\right)_{i=1}^{n}, \mathbb{F}_3^n\right)$
- Will an **RLC** of rate R **contain a matrix of type T**?
- There are q^{3Rn} triplets x_1, x_2, x_3 of words in C.
- Each P_i imposes $3 - |P_i|$ **linear conditions**.
- Let $\deg(P, \mathbb{F}_q^3) = 3Rn - \sum_i (3 - |P_i|)$.
- If $\deg(T) < 0$ then there is **probably no type T matrix in C**.

	3
	{1,2},{3}
	{1,2,3}
n	{1},{2},{3}
	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

$$\begin{array}{cc} & 3 \\ n & \begin{array}{|c|}\hline \{1,2\},\{3\} \\ \hline \{1,2,3\} \\ \hline \{1\},\{2\},\{3\} \\ \hline \{1,2,3\} \\ \hline \cdots \\ \hline \{1,3\},\{2\} \\ \hline \end{array}\end{array}$$

Large alphabet types in RLCs - Intuition

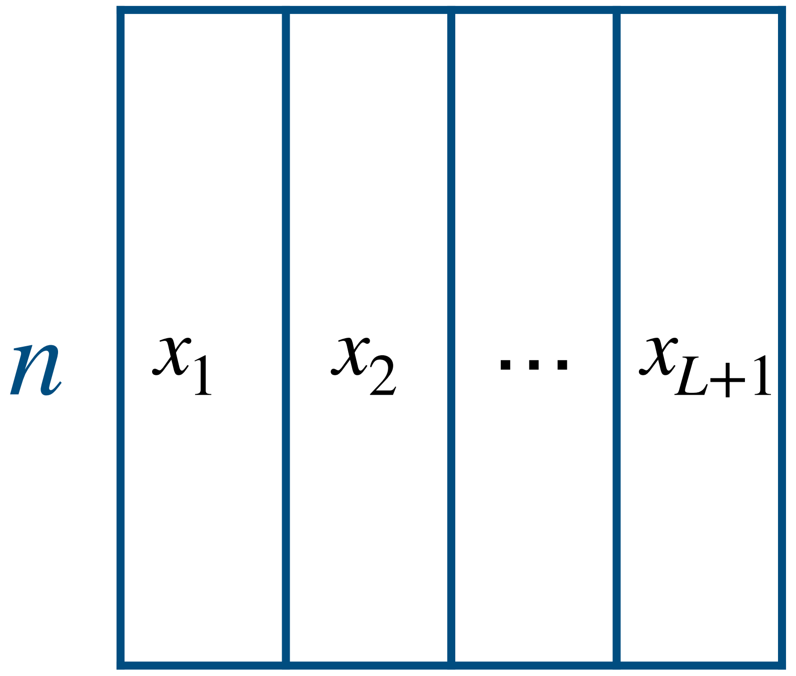
- What if $\deg(P, \mathbb{F}_q^3) > 0$?

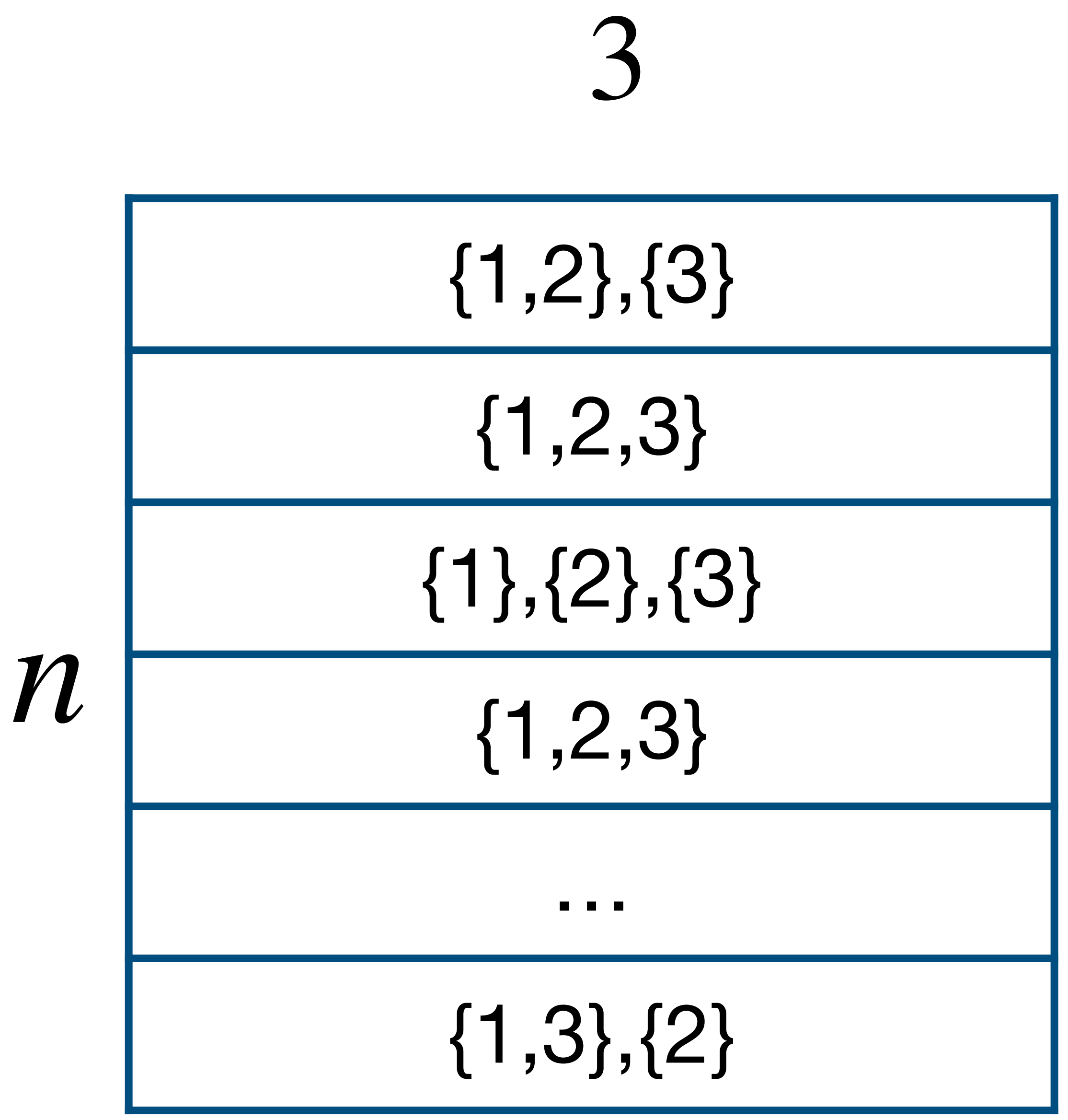
	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- What if $\deg(P, \mathbb{F}_q^3) > 0$?

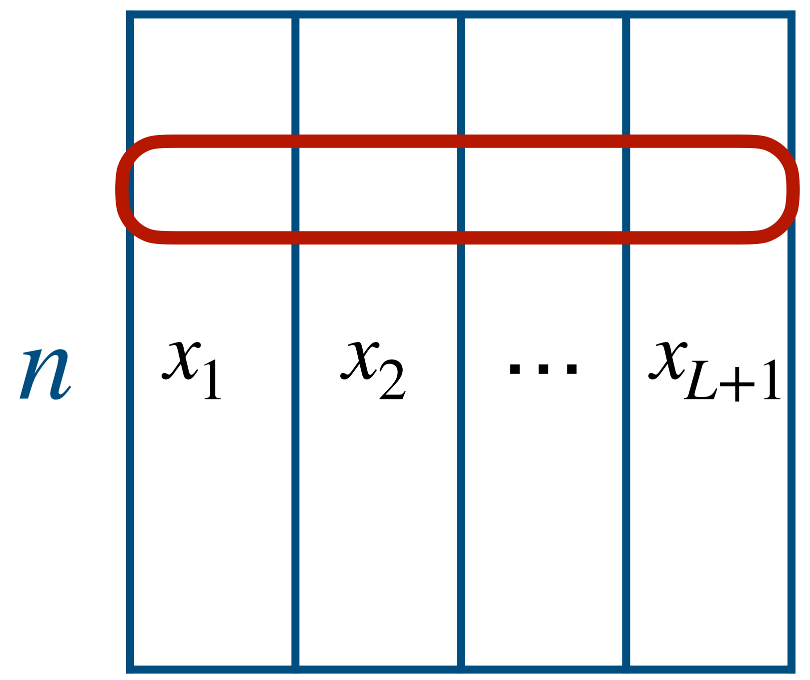
- Then **must** be non trivial triplets $x_1, x_2, x_3 \in C$ satisfying P.

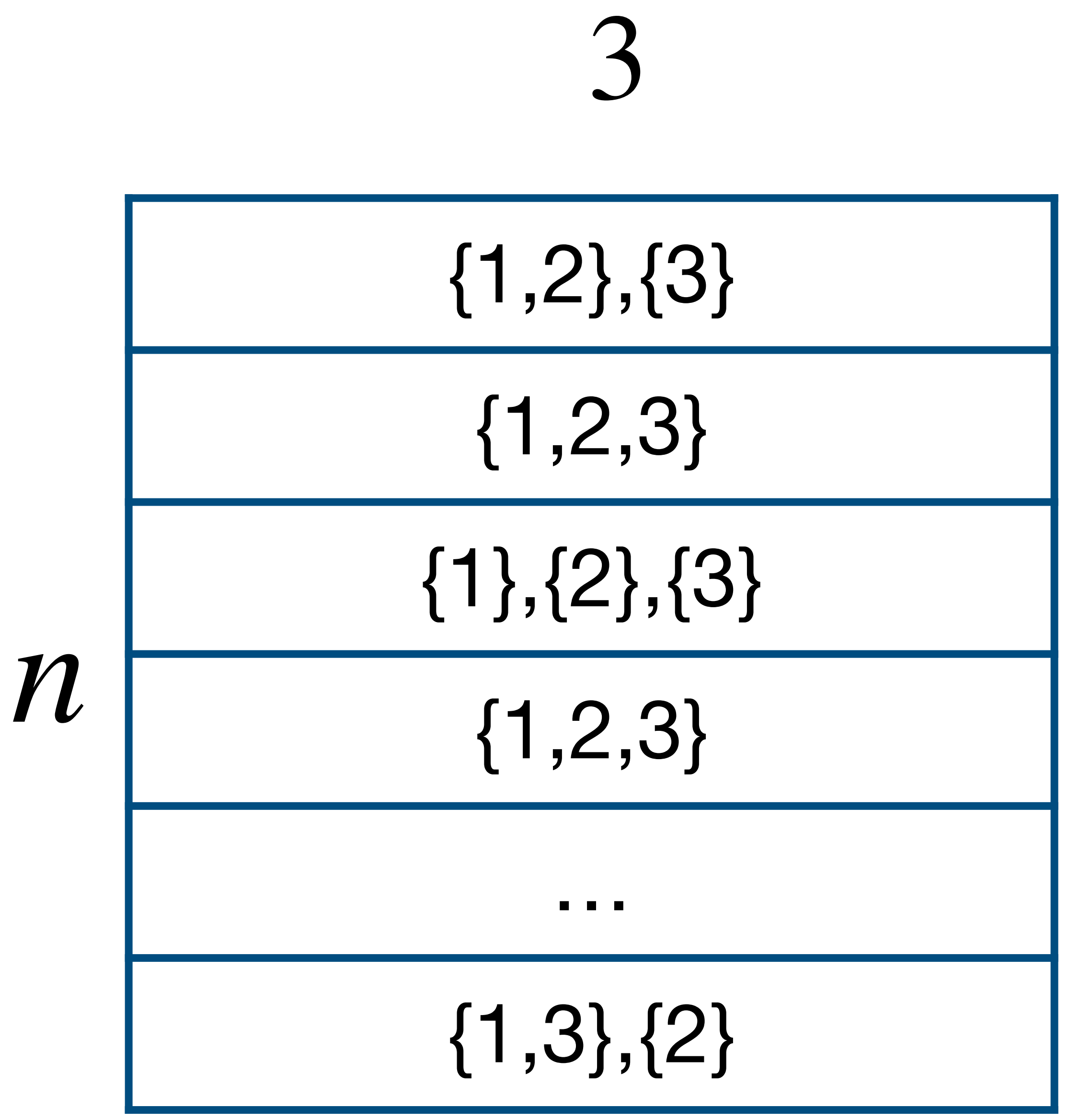




Large alphabet types in RLCs - Intuition

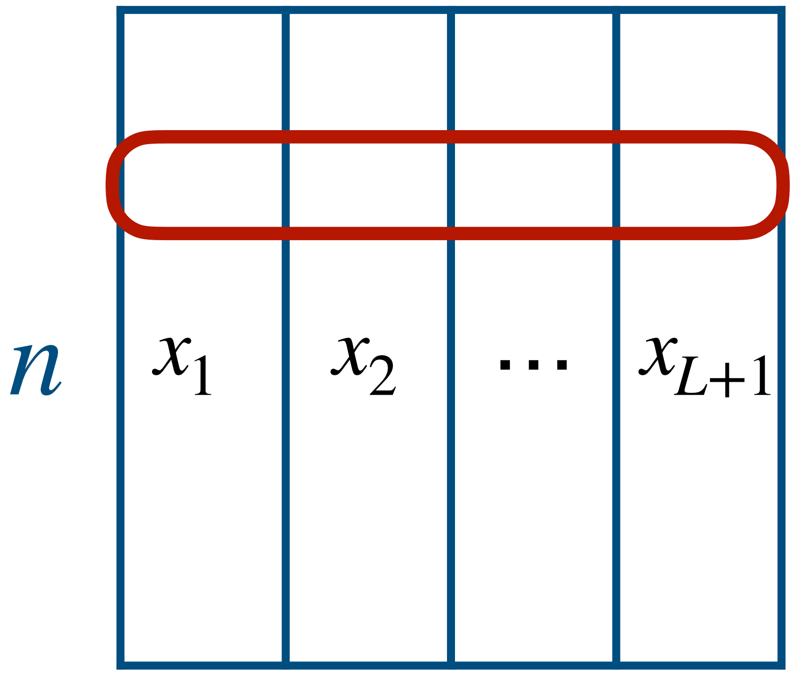
- What if $\deg(P, \mathbb{F}_q^3) > 0$?
- Then **must** be non trivial triplets $x_1, x_2, x_3 \in C$ satisfying P.
- But is their row span $\mathbb{F}_3^n$?

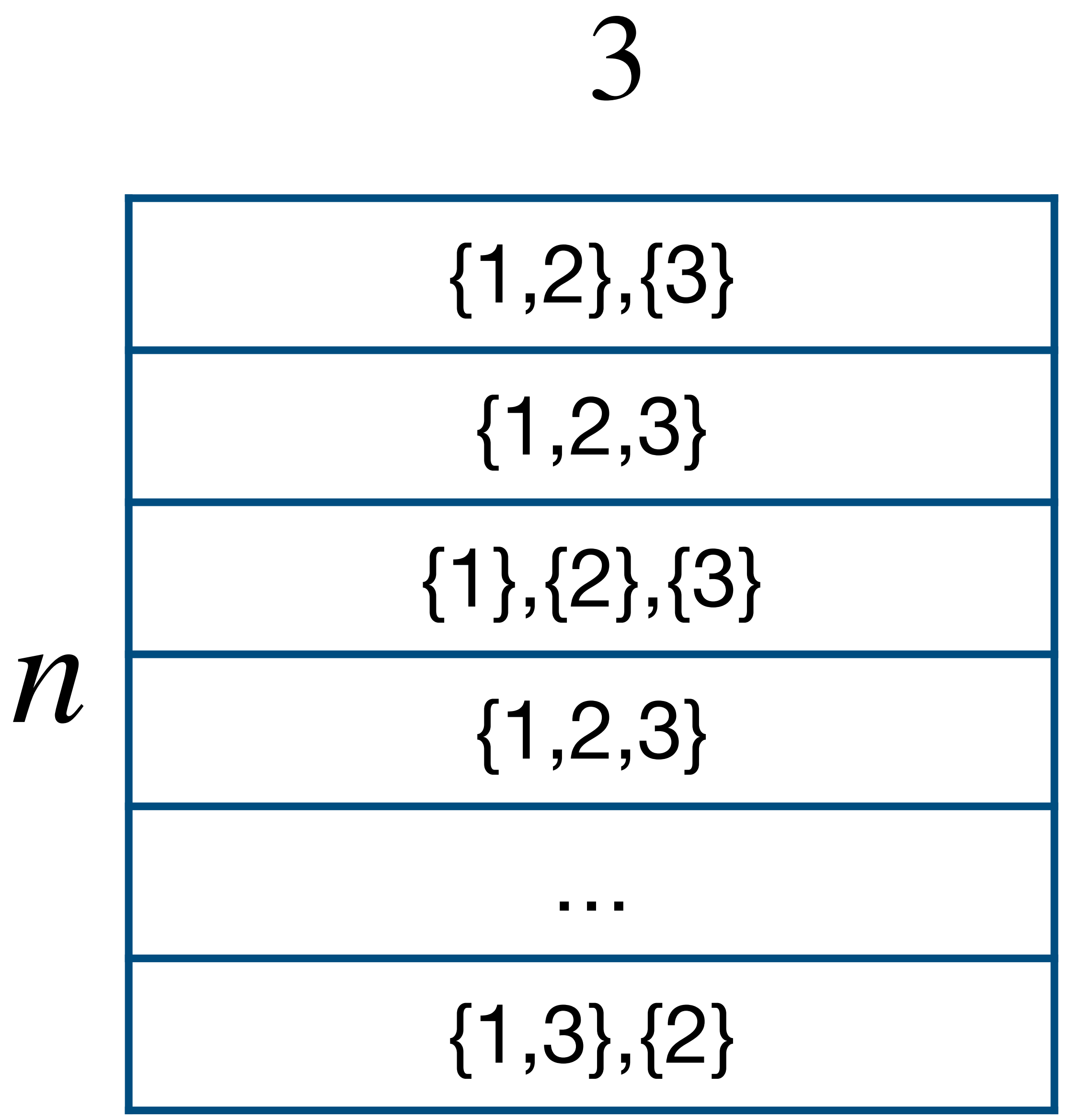




Large alphabet types in RLCs - Intuition

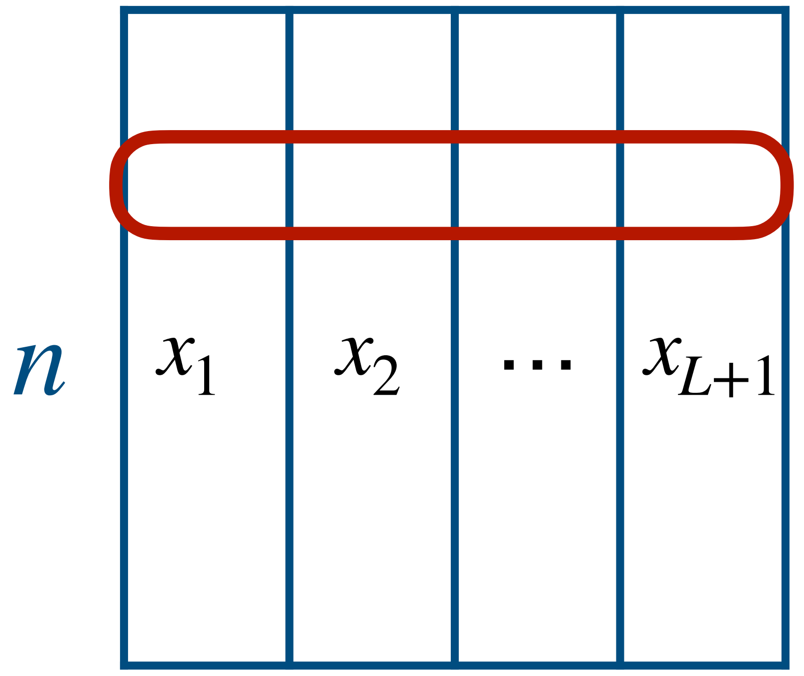
- What if $\deg(P, \mathbb{F}_q^3) > 0$?
- Then **must** be non trivial triplets $x_1, x_2, x_3 \in C$ satisfying P.
- But is their row span $\mathbb{F}_3^n$?
- **Maybe not!**

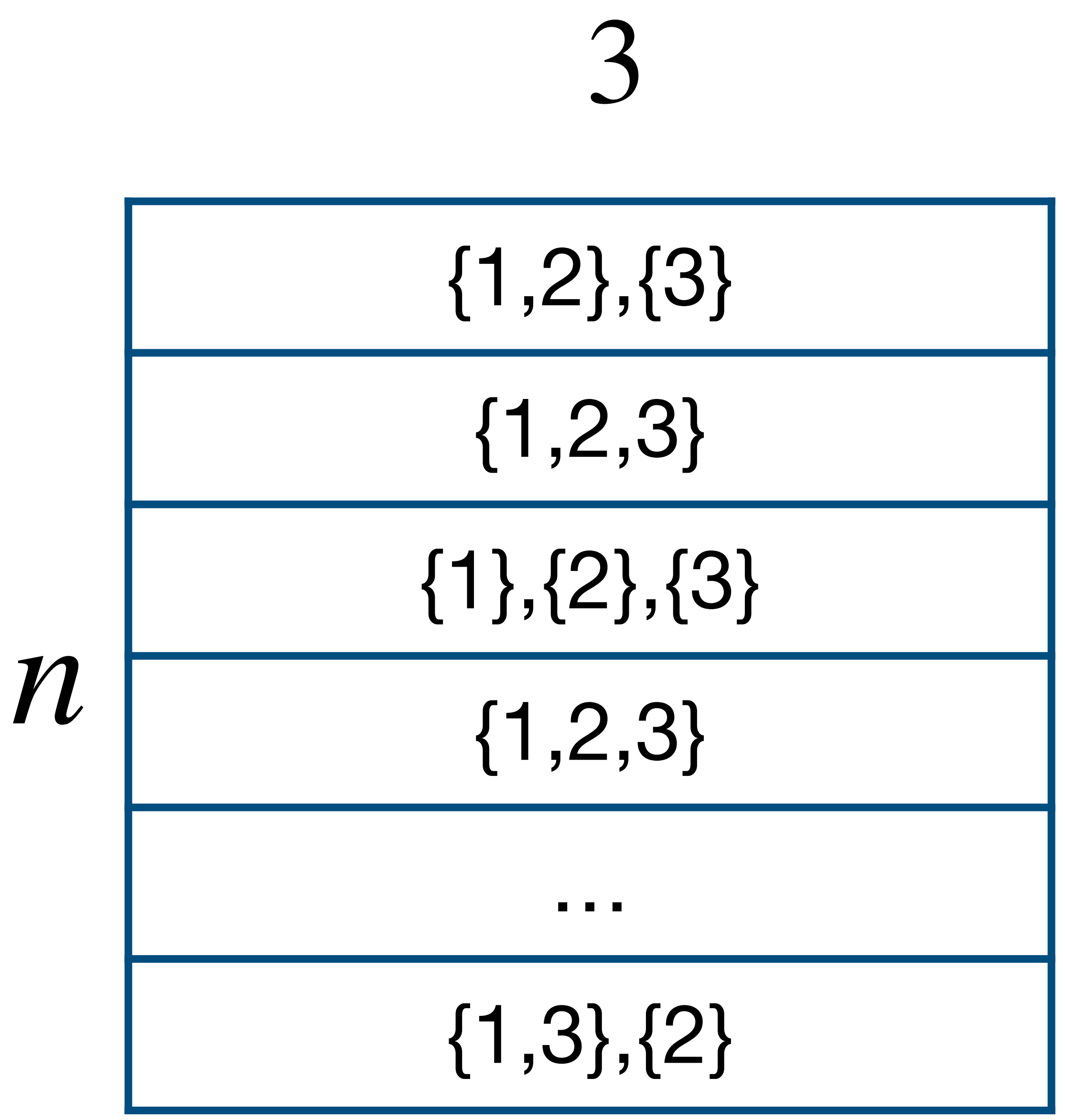




Large alphabet types in RLCs - Intuition

- What if $\deg(P, \mathbb{F}_q^3) > 0$?
- Then **must** be non trivial triplets $x_1, x_2, x_3 \in C$ satisfying P.
- But is their row span $\mathbb{F}_3^n$?
- **Maybe not!**
- It's possible that these x_1, x_2, x_3 are **not even distinct!**





Large alphabet types in RLCs - Intuition

3

n

{1,2},{3}
{1},{2},{3}
{1,2},{3}
{1,2},{3}
...
{1,2},{3}

- In this example we have $\deg(P, \mathbb{F}_q^3) > 0$.
- However, it's likely that all solutions will have $x_1 = x_2$!

Large alphabet types in RLCs - Intuition

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- What about the **type** $\left(\left(P_i\right)_{i=1}^{n}, V \right)$

3

n

{1,2},{3}
{1,2,3}
{1},{2},{3}
{1,2,3}
...
{1,3},{2}

Large alphabet types in RLCs - Intuition

- What about the **type** $\left(\left(P_i\right)_{i=1}^{n}, V\right)$
  - We take $V = \left\{ z \in \mathbb{F}_q^3 \mid z_1 + z_2 - 2z_3 = 0 \right\}$

	3
	{1,2},{3}
	{1,2,3}
n	{1},{2},{3}
	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- What about the **type** $\left(\left(P_i\right)_{i=1}^{n}, V\right)$
    - We take $V = \left\{ z \in \mathbb{F}_q^3 \mid z_1 + z_2 - 2z_3 = 0 \right\}$
- z_3 is determined by z_1, z_2 so we only have $2Rn$ **degrees of freedom**.

	3
	{1,2},{3}
	{1,2,3}
n	{1},{2},{3}
	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- What about the **type** $\left(\left(P_i\right)_{i=1}^{n}, V\right)$
  - We take $V = \left\{z \in \mathbb{F}_q^3 \mid z_1 + z_2 - 2z_3 = 0\right\}$
- z_3 is determined by z_1, z_2 so we only have $2Rn$ **degrees of freedom**.
- On the other hand, $z_1 = z_2 \Rightarrow z_1 = z_3$

	3
n	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
	{1,2,3}
	...
	{1,3},{2}

Large alphabet types in RLCs - Intuition

- What about the **type** $\left(\left(P_i\right)_{i=1}^{n}, V\right)$
    - We take $V = \left\{ z \in \mathbb{F}_q^3 \mid z_1 + z_2 - 2z_3 = 0 \right\}$
- z_3 is determined by z_1, z_2 so we only have $2Rn$ **degrees of freedom**.
- On the other hand, $z_1 = z_2 \Rightarrow z_1 = z_3$
    - So $\{1,2,3\}$ is just **1 constraint instead of 2**.

3

{1,2},{3}
{1,2,3}
{1},{2},{3}
{1,2,3}
...
{1,3},{2}

n

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

$$\deg(P, V) = \dim V \cdot Rn - \sum_{i=1}^{n} \left(\dim V - \dim V \cap V_{P_i} \right)$$

Where

$$V_{P_i} = \left\{ z \in \mathbb{F}_q^3 \mid z \text{ satisfies the equalities asserted by } P_i \right\}$$

	3
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

$$\deg(P, V) = \dim V \cdot Rn - \sum_{i=1}^{n} \left(\dim V - \dim V \cap V_{P_i} \right)$$

Where

$$V_{P_i} = \left\{ z \in \mathbb{F}_q^3 \mid z \text{ satisfies the equalities asserted by } P_i \right\}$$

Theorem [RLC thresholds for large alphabet]:

An **RLC** is **likely to contain a type (P, V) matrix** if and only if

$$\deg(P, V) > \deg(P, U)$$

For all $U \subseteq V$.

3

n	
	{1,2},{3}
	{1,2,3}
	{1},{2},{3}
n	{1,2,3}
	...
	{1,3},{2}

$$\deg(P, V) = \dim V \cdot Rn - \sum_{i=1}^{n} \left(\dim V - \dim V \cap V_{P_i}\right)$$

Where

$$V_{P_i} = \left\{ z \in \mathbb{F}_q^3 \mid z \text{ satisfies the equalities asserted by } P_i \right\}$$

Theorem [RLC thresholds for large alphabet]:

An **RLC** is **likely to contain a type (P, V) matrix** if and only if

$$\deg(P, V) > \deg(P, U)$$

For all $U \subseteq V$.

In particular
$\deg(P, V) > \deg(P, \{0\}) = 0$

Theorem [List-Decodability of RLC]
(previously proven by **[AGL]**):

For $q \geq 2^{\Omega(L)}$, an **RLC** in $\mathbb{F}_q^n$ **achieves the list-decoding GV bound**.

Theorem [List-Decodability of RLC]
(previously proven by **[AGL]**):

For $q \geq 2^{\Omega(L)}$, an **RLC** in $\mathbb{F}_q^n$ **achieves the list-decoding GV bound**.

Theorem [Reduction from RLC to random RS codes]:

Let $\mathscr{P}$ be a **local identity property achieved with high probability** by an **RLC**.

Then, $\mathscr{P}$ is also **achieved with high probability** by a **random RS code** with $q = O_L(n)$.

Theorem [Reduction from RLC to random RS codes] (Levi-M-Shagrithaya):

Let $\mathscr{P}$ be a **local identity property achieved with high probability** by an **RLC**.

Then, $\mathscr{P}$ is also **achieved with high probability** by a **random RS code** with $q = O_L(n)$.

Theorem [Reduction from RLC to random RS codes] (Levi-M-Shagrithaya):

Let $\mathscr{P}$ be a **local identity property** **achieved with high probability** by an **RLC**.

Then, $\mathscr{P}$ is also **achieved with high probability** by a **random RS code** with $q = O_L(n)$.

Corollary:

A **random RS code** **achieves the list-decoding GV-bound**.
(Already proven by **[AGL]** using the **GM-MDS theorem**)

A **random RS code** is at **least as list-recoverable** as an **RLC**.

Proof sketch: Reduction from random RS to RLC

By the **threshold theorem**, it suffices to solve the following problem:

Proof sketch: Reduction from random RS to RLC

By the **threshold theorem**, it suffices to solve the following problem:

Fix partitions $P = \left(P_i\right)$.

Suppose that $\deg(P, \mathbb{F}_2^{L+1}) \leq -\epsilon n$.

We need to prove:

$$\Pr\left[\text{A random RS code contains a type } (P, \mathbb{F}_2^{L+1}) \text{ matrix}\right] \leq q^{-\Omega(n)}$$

Proof sketch: Reduction from random RS to RLC

By the **threshold theorem**, it suffices to solve the following problem:

Fix partitions $P = \left(P_i\right)$.

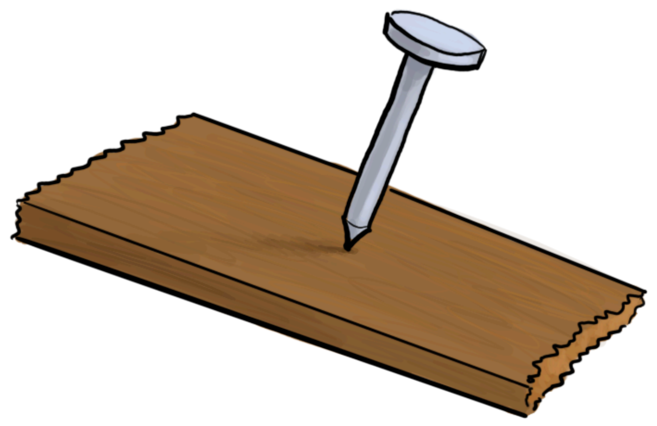
Suppose that $\deg(P, \mathbb{F}_2^{L+1}) \leq -\epsilon n$.

We need to prove:

$$\Pr\left[\text{A random RS code contains a type } (P, \mathbb{F}_2^{L+1}) \text{ matrix}\right] \leq q^{-\Omega(n)}$$

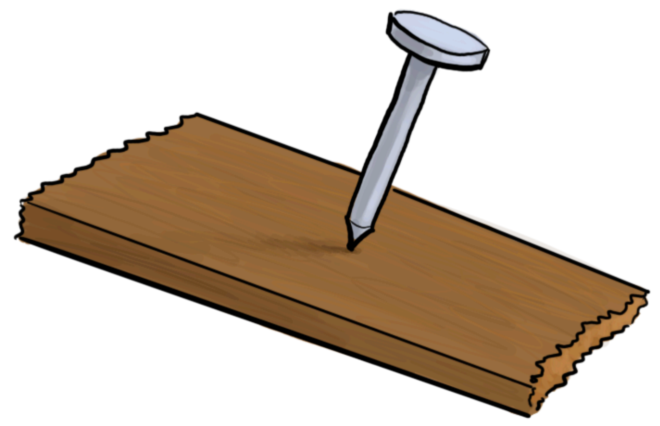
On the Board

Open Problems



Open Problems

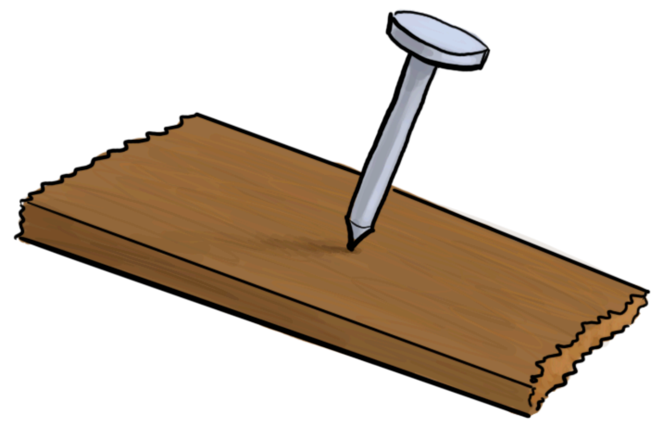
- Fully understand list-Recoverability of **RLC** and **random RS**.



Open Problems

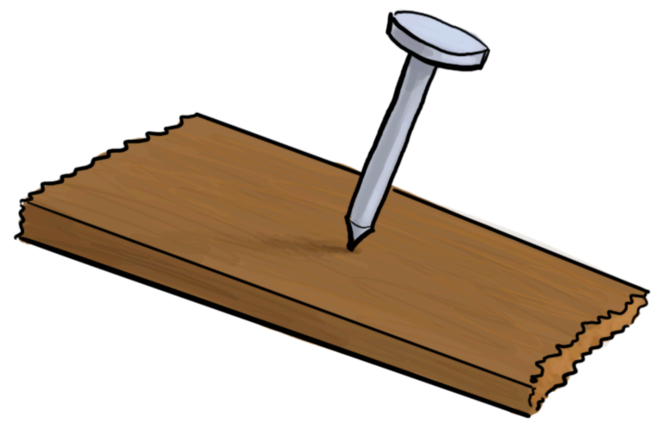
- Fully understand list-Recoverability of **RLC** and **random RS**.

- Break $\Omega(n)$ randomness barrier



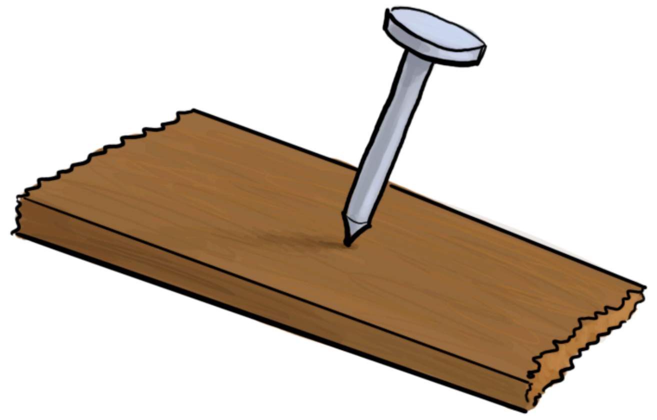
Open Problems

- Fully understand list-Recoverability of **RLC** and **random RS**.
- Break $\Omega(n)$ randomness barrier
- Handle **non-local properties**



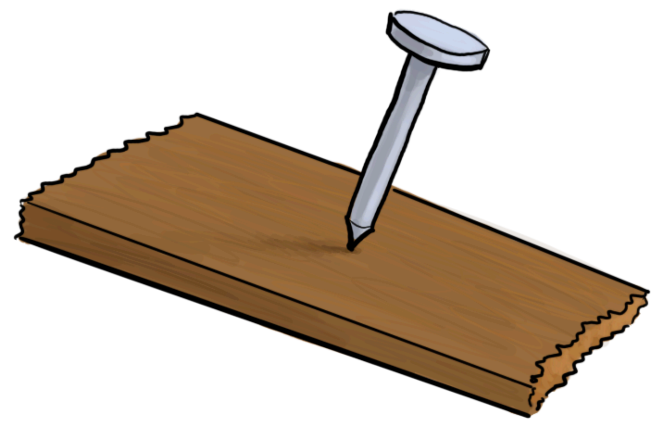
Open Problems

- Fully understand list-Recoverability of **RLC** and **random RS**.
- Break $\Omega(n)$ randomness barrier
- Handle **non-local properties**
    - **List-recoverability** with **large list size**.



Open Problems

- Fully understand list-Recoverability of **RLC** and **random RS**.
- Break $\Omega(n)$ randomness barrier
- Handle **non-local properties**
  - **List-recoverability** with **large list size**.
- Handle Π_2 **properties**



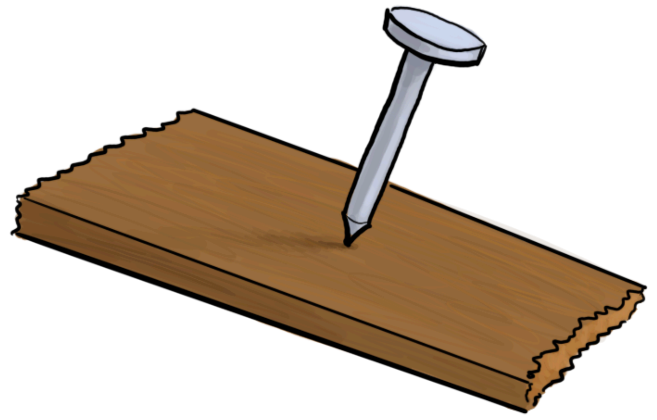
Open Problems

- Fully understand list-Recoverability of **RLC** and **random RS**.

- Break $\Omega(n)$ randomness barrier

- Handle **non-local properties**

    - **List-recoverability** with **large list size**.

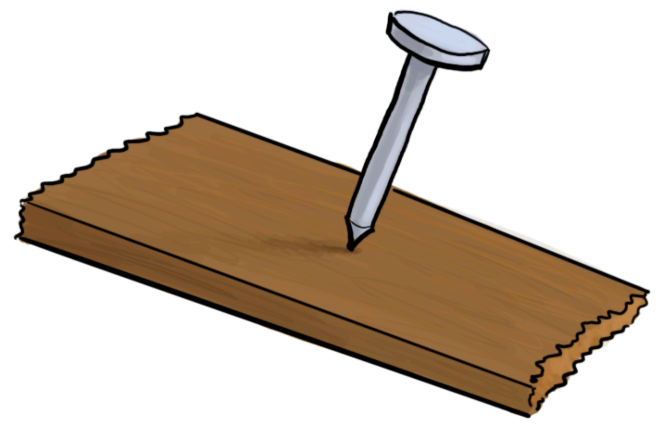


- Handle Π_2 **properties**

    - (ρ, L)**-covering**

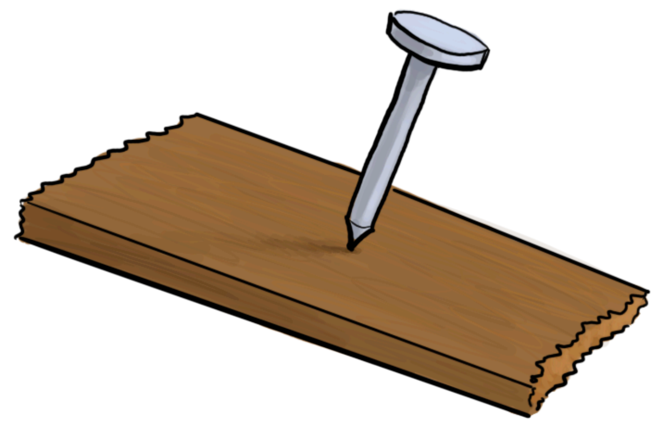
Open Problems

- Fully understand list-Recoverability of **RLC** and **random RS**.
- Break $\Omega(n)$ randomness barrier
- Handle **non-local properties**
    - **List-recoverability** with **large list size**.
- Handle Π_2 **properties**
    - (ρ, L)**-covering**
- Find **limit objects** for codes.



Open Problems

- Fully understand list-Recoverability of **RLC** and **random RS**.
- Break $\Omega(n)$ randomness barrier
- Handle **non-local properties**
    - **List-recoverability** with **large list size**.
- Handle Π_2 **properties**
    - (ρ, L)**-covering**
- Find **limit objects** for codes.



Thank you!