Near-linear Time Decoding of Ta-Shma's Codes via Splittable Regularity

Fernando Granha Jeronimo
(UChicago)

joint work with

Shashank Srivastava (TTIC) and
Madhur Tulsiani (TTIC)

STOC 2021

Goal of the Talk

Goal

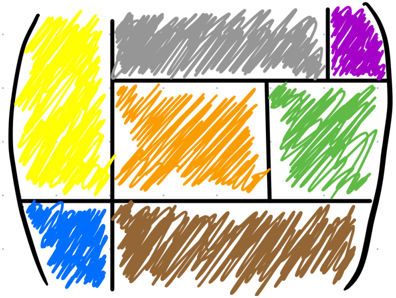
Present a near-linear time decoding algorithm for Ta-Shma's codes

Goal of the Talk

Goal

Present a near-linear time decoding algorithm for Ta-Shma's codes

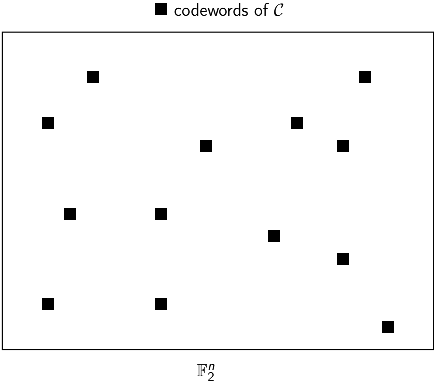
Weak Regularity Technique



Coding Theory Concepts

Code

A binary code is a subset $\mathcal{C} \subseteq \mathbb{F}_2^n$



Coding Theory Concepts

Two fundamental parameters

Distance

The distance $\Delta(\mathcal{C})$ of $\mathcal{C}$ is $\Delta(\mathcal{C}) := \min_{z,z' \in \mathcal{C}:\ z \neq z'} \Delta(z, z')$

Coding Theory Concepts

Two fundamental parameters

Distance

The distance $\Delta(\mathcal{C})$ of $\mathcal{C}$ is $\Delta(\mathcal{C}) := \min_{z,z' \in \mathcal{C}:\ z \neq z'} \Delta(z, z')$

Rate

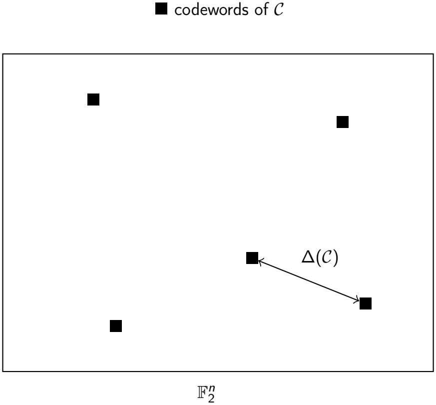
The rate $r(\mathcal{C})$ of $\mathcal{C}$ is $\frac{\log_2(|\mathcal{C}|)}{n}$ (the fraction of information symbols)

Tension between Rate and Distance of a Code

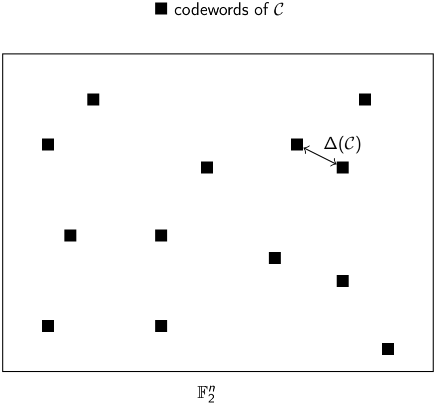
Tension

- Higher rate $r(\mathcal{C})$, lower distance $\Delta(\mathcal{C})$
- Higher distance $\Delta(\mathcal{C})$, lower rate $r(\mathcal{C})$

Tension between Rate and Distance of a Code



Lower rate $r(\mathcal{C})$



Higher rate $r(\mathcal{C})$

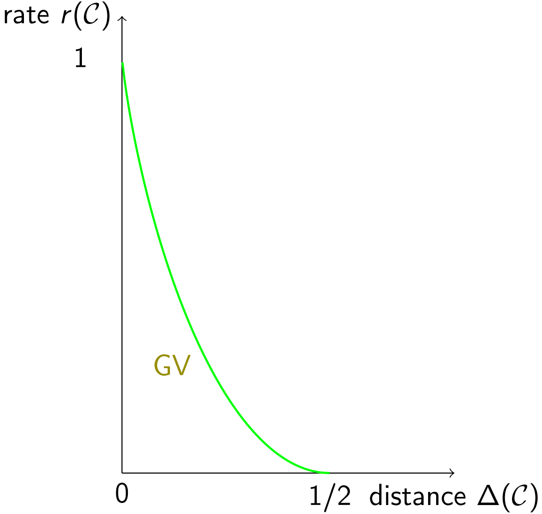
Coding Theory Concepts

Question

What is the best trade-off between rate $r(\mathcal{C})$ and distance $\Delta(\mathcal{C})$?

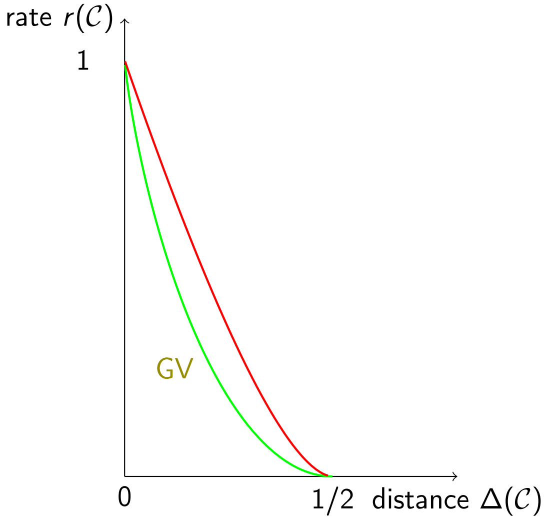
Coding Theory Concepts

Gilbert–Varshamov existential bound (Gilbert'52,Varshamov'57)

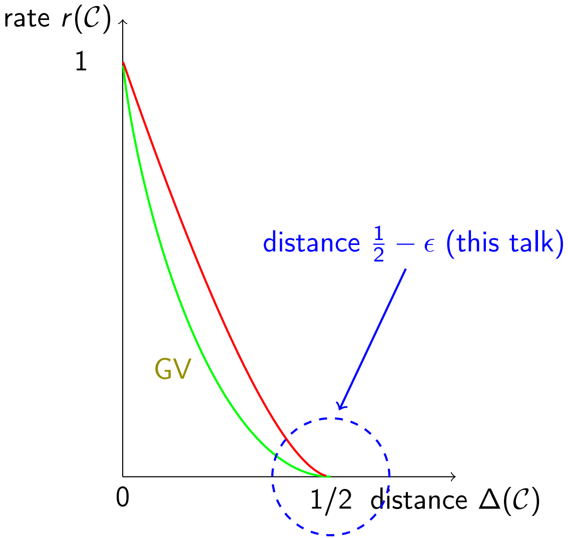


Coding Theory Concepts

McEliece–Rodemich–Rumsey–Welch'77 impossibility bound



Coding Theory Concepts



Coding Theory Concepts

Why is the Gilbert–Varshamov bound interesting?

The Gilbert–Varshamov (GV) bound is *"nearly"* optimal

For distance $1/2 - \epsilon$

- rate $\Omega(\epsilon^2)$ is achievable (Gilbert–Varshamov bound)
- rate better than $O(\epsilon^2 \log(1/\epsilon))$ is impossible (McEliece *et al.*)

Coding Theory Concepts

For distance $1/2 - \epsilon$

- rate $\Omega(\epsilon^2)$ is achievable (Gilbert–Varshamov bound)
- rate better than $O(\epsilon^2 \log(1/\epsilon))$ is impossible (McEliece *et al.*)

Ta-Shma's Codes (60 years later!)

First **explicit** binary codes near the GV bound are due to Ta-Shma'17 with

- distance $1/2 - \epsilon/2$ (actually ϵ-balanced), and
- rate $\Omega(\epsilon^{2+o(1)})$.

Coding Theory Concepts

Ta-Shma's Codes (60 years later!)

First **explicit** binary codes near the GV bound are due to Ta-Shma'17 with

- distance $1/2 - \epsilon/2$ (actually ϵ-balanced), and
- rate $\Omega(\epsilon^{2+o(1)})$.

Question

How efficiently can we decode Ta-Shma codes?

Coding Theory Concepts

Question

How efficiently can we decode Ta-Shma codes?

Theorem (this talk)

Ta-Shma's codes are near-linear time unique decodable

Our Contribution

Theorem (Near-linear Time Decoding)

For every $\epsilon > 0$, $\exists$ explicit binary linear Ta-Shma codes $\mathcal{C}_{N,\epsilon} \subseteq \mathbb{F}_2^N$ with

1. *distance at least $1/2 - \epsilon/2$ (actually ϵ-balanced),*
2. *rate $\Omega(\epsilon^{2+o(1)})$, and*
3. *a unique decoding algorithm with running time $\widetilde{O}_\epsilon(N)$.*

Our Contribution

Pseudorandomness approach

Theorem (Near-linear Time Decoding)

For every $\epsilon > 0$, $\exists$ explicit binary linear Ta-Shma codes $\mathcal{C}_{N,\epsilon} \subseteq \mathbb{F}_2^N$ with

1. *distance at least $1/2 - \epsilon/2$ (actually ϵ-balanced),*
2. *rate $\Omega(\epsilon^{2+o(1)})$, and*
3. *a unique decoding algorithm with running time $\widetilde{O}_\epsilon(N)$.*

Previous Efficient Decoder for Ta-Shma's Codes

Sum-of-Squares SDP hierarchy approach (SOS approach)

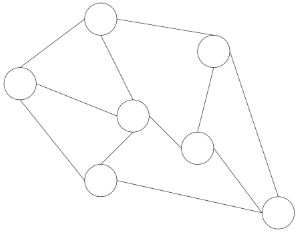
Theorem (J-Quintana-Srivastava-Tulsiani'20)

Ta-Shma's codes are unique decodable in $N^{O_\epsilon(1)}$ *time*

Towards Ta-Shma's Codes

Expander Graphs and Codes

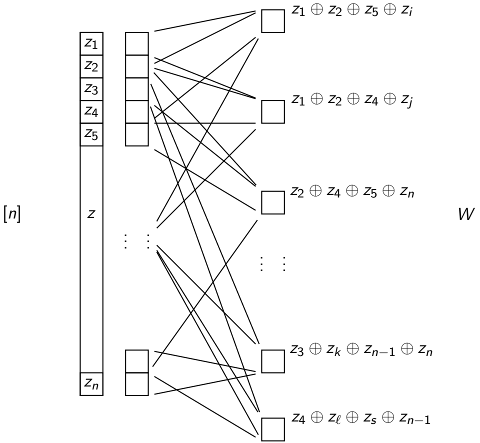
Expanders can amplify the distance of a not so great base code $\mathcal{C}_0$



Expansion and Distance Amplification

Fix a bipartite graph between $[n]$ and $W \subseteq [n]^k$. Let $z \in \mathbb{F}_2^n$.

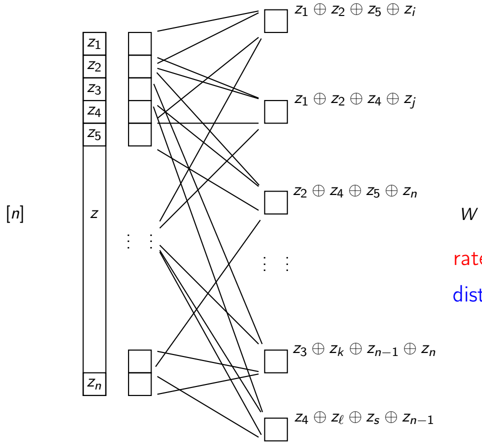
Direct Sum



Expansion and Distance Amplification

Fix a bipartite graph between $[n]$ and $W \subseteq [n]^k$. Let $z \in \mathbb{F}_2^n$.

Direct Sum



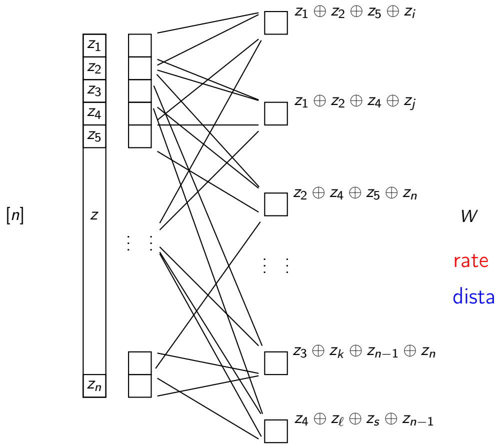
rate loss factor $n/|W|$

distance amplification needs to be worth this loss

Expansion and Distance Amplification

Fix a bipartite graph between $[n]$ and $W \subseteq [n]^k$. Let $z \in \mathbb{F}_2^n$.

Direct Sum



rate loss factor $n/|W|$

distance amplification needs to be worth this loss

Alon–Brooks–Naor–Naor–Roth & Alon–Edmonds–Luby *style* distance amplification

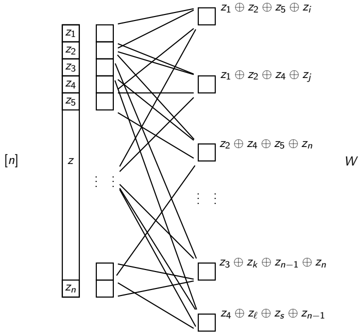
Expansion and Distance Amplification

Direct Sum

Let $z \in \mathbb{F}_2^n$ and $W \subseteq [n]^k$. The *direct sum* of z is $y \in \mathbb{F}_2^W$ defined as

$$y_{(i_1,\ldots,i_k)} = z_{i_1} \oplus \cdots \oplus z_{i_k},$$

for every $(i_1, \ldots, i_k) \in W$. We denote $y = \mathrm{dsum}_W(z)$.



Expansion and Distance Amplification

Bias

- Let $z \in \mathbb{F}_2^n$. Define $\text{bias}(z) := |\mathbf{E}_{i \in [n]} (-1)^{z_i}|$
- $\text{bias}(\mathcal{C}) = \max_{z \in \mathcal{C} \setminus 0} \text{bias}(z)$
- If $\text{bias}(\mathcal{C}) \leq \epsilon$, then $\Delta(\mathcal{C}) \geq 1/2 - \epsilon/2$ (assuming $\mathcal{C}$ linear)

Expansion and Distance Amplification

Bias

- Let $z \in \mathbb{F}_2^n$. Define $\mathrm{bias}(z) := \left| \mathbf{E}_{i \in [n]} (-1)^{z_i} \right|$
- $\mathrm{bias}(\mathcal{C}) = \max_{z \in \mathcal{C} \setminus 0} \mathrm{bias}(z)$
- If $\mathrm{bias}(\mathcal{C}) \leq \epsilon$, then $\Delta(\mathcal{C}) \geq 1/2 - \epsilon/2$ (assuming $\mathcal{C}$ linear)

$$\mathrm{bias}(\underbrace{00\ldots0}_{n}) = \mathrm{bias}(\underbrace{11\ldots1}_{n}) = 1$$

$$\mathrm{bias}(\underbrace{0\ldots0}_{n/2}\underbrace{1\ldots1}_{n/2}) = 0$$

Expansion and Distance Amplification

Bias

- Let $z \in \mathbb{F}_2^n$. Define $\text{bias}(z) := \left|\mathbf{E}_{i\in[n]}(-1)^{z_i}\right|$
- $\text{bias}(\mathcal{C}) = \max_{z\in\mathcal{C}\setminus 0} \text{bias}(z)$
- If $\text{bias}(\mathcal{C}) \leq \epsilon$, then $\Delta(\mathcal{C}) \geq 1/2 - \epsilon/2$ (assuming $\mathcal{C}$ linear)

Definition (Parity Sampler, c.f. Ta-Shma'17)

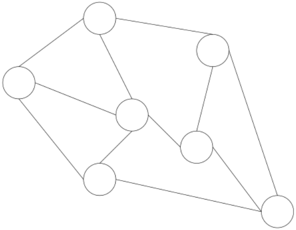
Let $W \subseteq [n]^k$. We say that dsum_W is (ϵ_0, ϵ)-**parity sampler** iff

$$(\forall z \in \mathbb{F}_2^n)\,(\text{bias}(z) \leq \epsilon_0 \implies \text{bias}(\text{dsum}_W(z)) \leq \epsilon)\,.$$

Explicit Constructions of Parity Samplers

Solution (Alon and Rozenman–Wigderson)

Take $W \subseteq [n]^k$ to be the collection of **all** length-$(k-1)$ walks on a sparse expander graph $G = (V = [n], E)$



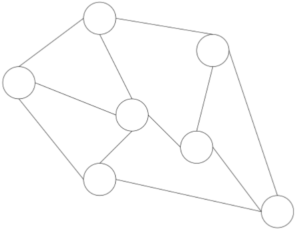
Explicit Constructions of Parity Samplers

Solution (Alon and Rozenman–Wigderson)

Take $W \subseteq [n]^k$ to be the collection of **all** length-$(k-1)$ walks on a sparse expander graph $G = (V = [n], E)$

Solution (good but not near optimal)

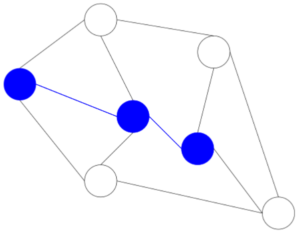
This yields codes of distance $1/2 - \epsilon$ and rate $\Omega(\epsilon^{4+o(1)})$



Explicit Constructions of Parity Samplers

Solution of Ta-Shma'17

Take $W \subseteq [n]^k$ to be a **carefully chosen** collection of length-$(k-1)$ walks on a structured sparse expander graph $G = (V = [n], E)$



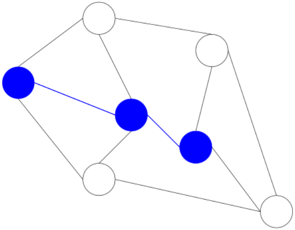
Explicit Constructions of Parity Samplers

Solution of Ta-Shma'17

Take $W \subseteq [n]^k$ to be a **carefully chosen** collection of length-$(k-1)$ walks on a structured sparse expander graph $G = (V = [n], E)$

Solution (near optimal)

This yields codes of distance $1/2 - \epsilon$ and rate $\Omega(\epsilon^{2+o(1)})$



General Techniques for Decoding

Decoding Direct Sum

What does decoding look like for direct sum?

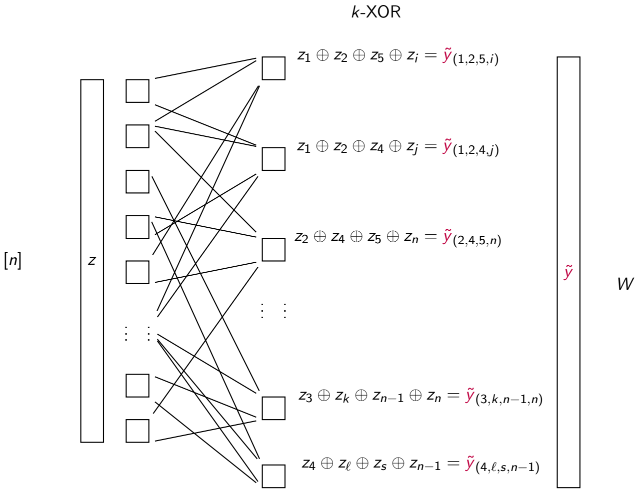
Decoding by Solving a k-CSP

Setup (informal)

- $\mathcal{C}_0 \subseteq \mathbb{F}_2^n$ is a code of small distance
- $W \subseteq [n]^k$ for direct sum
- $\mathcal{C} = \mathrm{dsum}_W(\mathcal{C}_0)$ is a code of large distance

Decoding by Solving a k-CSP

Suppose $y^* \in \mathcal{C}$ is corrupted into some $\tilde{y} \in \mathbb{F}_2^W$ in the unique decoding ball centered at y^*.



Decoding by Solving a k-CSP

Unique Decoding Scenario: k-XOR like

Unique decoding $\tilde{y}$ amounts to solving

$$\arg\max_{z \in \mathcal{C}_0} \mathbf{E}_{(i_1,\ldots,i_k)\in W} 1[z_{i_1} \oplus \cdots \oplus z_{i_k} = \tilde{y}_{(i_1,\ldots,i_k)}].$$

Decoding by Solving a k-CSP

Unique Decoding Scenario: k-XOR like

Unique decoding $\tilde{y}$ amounts to solving

$$\arg\max_{z \in \mathcal{C}_0} \mathbf{E}_{(i_1,\ldots,i_k)\in W} 1[z_{i_1} \oplus \cdots \oplus z_{i_k} = \tilde{y}_{(i_1,\ldots,i_k)}].$$

A Relaxation

Suppose that we can find $\tilde{z} \in \mathbb{F}_2^n$ (rather than in $\mathcal{C}_0$) satisfying

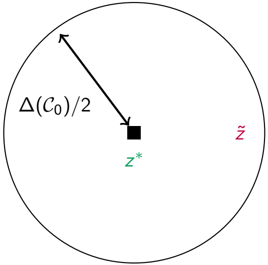
$$\mathbf{E}_{(i_1,\ldots,i_k)\in W} 1[\tilde{z}_{i_1} \oplus \cdots \oplus \tilde{z}_{i_k} = \tilde{y}_{(i_1,\ldots,i_k)}] \approx \mathrm{OPT}.$$

Decoding by Solving a k-CSP

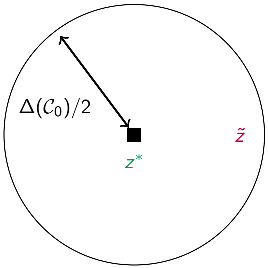
Say $y^* = \mathrm{dsum}(z^*)$ for some $z^* \in \mathcal{C}_0$

Claim (Informal)

If the parity sampler is *strong enough*, then $\tilde{z}$ lies in the unique decoding ball centered at $z^* \in \mathcal{C}_0$.



Decoding by Solving a k-CSP



Moral

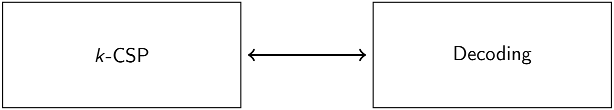
- Find approx. optimal solution $\tilde{z} \in \mathbb{F}_2^n$ (rather than in $\mathcal{C}_0$) is enough
- Use unique decoder of $\mathcal{C}_0$ to correct $\tilde{z}$ into z^*

What do we have so far?

k-CSP $\longleftrightarrow$ Decoding

What do we have so far?

Why can we efficiently approximate these k-CSPs?



What do we have so far?

Why can we efficiently approximate these k-CSPs?

k-CSP $\longleftrightarrow$ Decoding

$+$

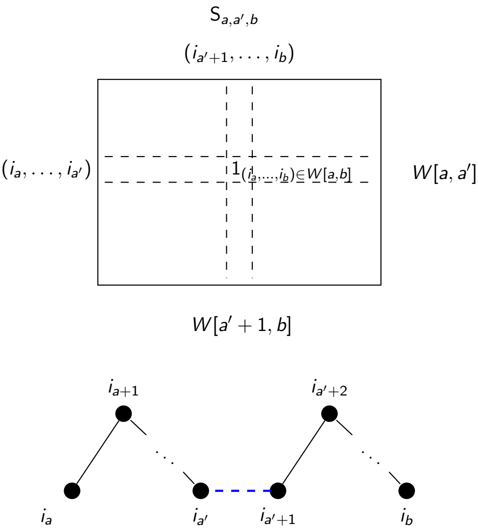
high-dimensional expansion

A Notion of High-dimensional Expansion

Let $W \subseteq [n]^k$. Define $W[a, b]$ for $1 \leq a \leq b \leq k$ as

$$W[a, b] = \{(i_a, \ldots, i_b) \mid (i_1, \ldots, i_k) \in W\}.$$

A Notion of High-dimensional Expansion



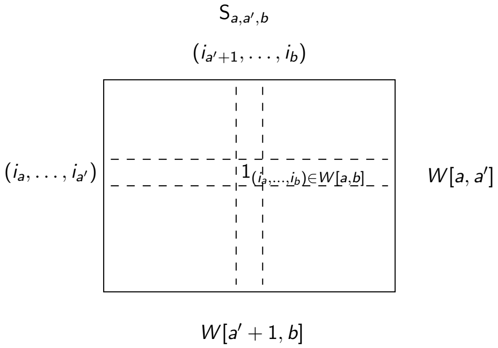
$$W[a, b] = \{(i_a, \ldots, i_b) \mid (i_1, \ldots, i_k) \in W\}$$

A Notion of High-dimensional Expansion

Definition (Splittability (informal) Mossel'10)

A collection $W \subseteq [n]^k$ is said to be τ-splittable, if $k = 1$ or for every $1 \leq a \leq a' < b \leq k$:

1. The (normalized) matrix $\mathsf{S}_{a,a',b} \in \mathbb{R}^{W[a,a'] \times W[a'+1,b]}$ defined as $\mathsf{S}_{a,a',b}(w, w') = 1_{ww' \in W[a,b]}$ satisfy $\sigma_2(\mathsf{S}_{a,a',b}) \leq \tau$

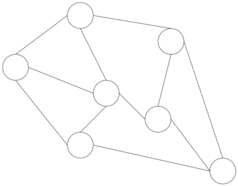


A Notion of High-dimensional Expansion

Example of τ-splittable structures

Lemma (Alev–J–Quintana–Srivastava–Tulsiani'20)

*The collection $W \subseteq [n]^k$ of **all** walks on τ-two-sided spectral expander graph $G = (V = [n], E)$ is τ-splittable*

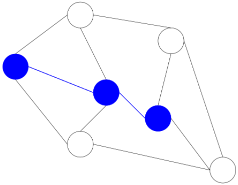


A Notion of High-dimensional Expansion

Example of τ-splittable structures

Lemma (J–Quintana–Srivastava–Tulsiani'20)

A simple modification of Ta-Shma's parity sampler $W \subseteq [n]^k$ is τ-splittable



Previous Approach via Sum-of-Squares

Sum-of-Squares

k-CSP $\longleftrightarrow$ Decoding

$+$

high-dimensional expansion

Previous Approach via Sum-of-Squares

Using the Sum-of-Squares (SOS) semi-definite programming hierarchy:

Theorem (Alev–J–Tulsiani'19 (informal))

Instances of k-XOR supported on expanding (τ-splittable) tuples $W \subseteq [n]^k$ can be efficiently approximated

(building on Barak–Raghavendra–Steurer'11)

Previous Approach via Sum-of-Squares

Using the Sum-of-Squares (SOS) semi-definite programming hierarchy:

Theorem (Alev–J–Tulsiani'19)

Let $W \subseteq [n]^k$ be τ-splittable. Suppose $\mathfrak{I}$ is a k-XOR instance on W. If $\tau \leq \mathrm{poly}(\delta/2^k)$, then we can find a solution $z \in \mathbb{F}_2^n$ satisfying

$$\mathrm{OPT}(\mathfrak{I}) - \delta,$$

fraction of the constraints of $\mathfrak{I}$ in time $n^{\mathrm{poly}(2^k/\delta)}$.

(building on Barak–Raghavendra–Steurer'11)

Pseudorandomness Approach

Pseudorandomness

k-CSP $\longleftrightarrow$ Decoding

$+$

high-dimensional expansion

Pseudorandomness Approach

Using pseudorandomness techniques (weak regularity decompositions):

Theorem (J–Srivastava–Tulsiani'20)

Let $W \subseteq [n]^k$ be τ-splittable. Suppose $\mathfrak{I}$ is a k-XOR instance on W. If $\tau \leq \mathrm{poly}(\delta/k)$, then we can find a solution $z \in \mathbb{F}_2^n$ satisfying

$$\mathrm{OPT}(\mathfrak{I}) - \delta,$$

fraction of the constraints of $\mathfrak{I}$ in time $\widetilde{O}_\delta(|W|)$.

Weak Regularity Decomposition: Notation

Cut Matrix

$$S, T \subseteq [n]$$

$$\begin{array}{cc} & \begin{array}{ccc} & \overbrace{}^{T} & \end{array} \\ S & \left[\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] \end{array}$$

$$1_S \otimes 1_T$$

Weak Regularity Decomposition: Dense Graphs

We recall Frieze and Kannan'96 approach.
Let A be the adjacency matrix of a **dense** graph $G = ([n], E)$. Suppose we have $A \approx \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{S_2^\ell}$ such that

$$\max_{S,T \subseteq [n]} \left| \langle A - \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{S_2^\ell}, 1_S \otimes 1_T \rangle \right| \leq \delta \cdot n^2,$$

and $L = O(1/\delta^2)$.

Weak Regularity Decomposition: Dense Graphs

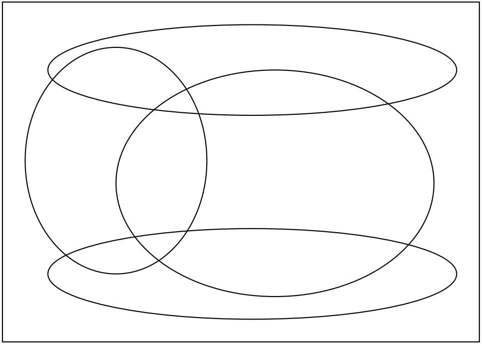
Frieze and Kannan use $\sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{S_2^\ell}$ to approximate the **maximum cut** value of G within additive error $\delta \cdot n^2$

$$\begin{aligned}|E(S,\overline{S})| = \langle A, 1_S \otimes 1_{\overline{S}} \rangle &\approx \langle \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{S_2^\ell}, 1_S \otimes 1_{\overline{S}} \rangle, \\ &= \sum_{\ell=1}^{L} c_\ell \cdot |S_1^\ell \cap S||S_2^\ell \cap \overline{S}|,\end{aligned}$$

Weak Regularity Decomposition: Dense Graphs

$$|E(S, \overline{S})| \approx \sum_{\ell=1}^{L} c_\ell \cdot |S_1^\ell \cap S||S_2^\ell \cap \overline{S}|$$

Venn diagram of sets S_1^ℓ, S_2^ℓ for $\ell \in [L]$

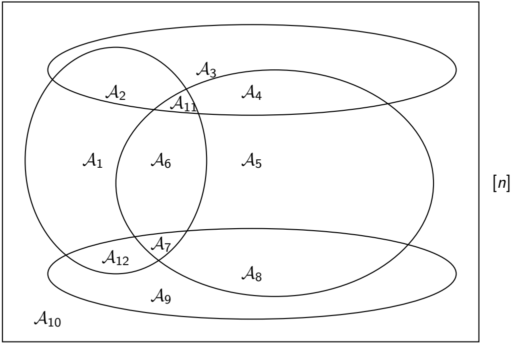


$[n]$

Weak Regularity Decomposition: Dense Graphs

$|E(S, \overline{S})| \approx \sum_{\ell=1}^{L} c_\ell \cdot |S_1^\ell \cap S||S_2^\ell \cap \overline{S}|$

Venn diagram of sets S_1^ℓ, S_2^ℓ for $\ell \in [L]$

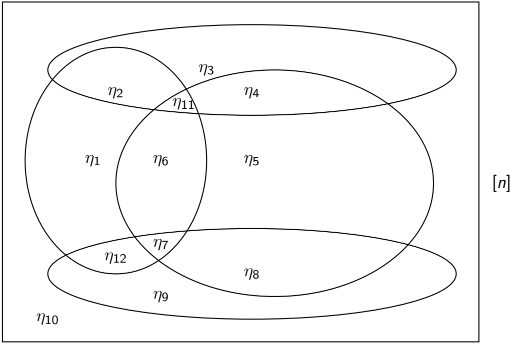


$\exp(L) = \exp(1/\delta^2)$ atoms $\mathcal{A}_1, \mathcal{A}_2, \ldots$

Weak Regularity Decomposition: Dense Graphs

$|E(S, \overline{S})| \approx \sum_{\ell=1}^{L} c_\ell \cdot |S_1^\ell \cap S||S_2^\ell \cap \overline{S}|$

Venn diagram of sets S_1^ℓ, S_2^ℓ for $\ell \in [L]$



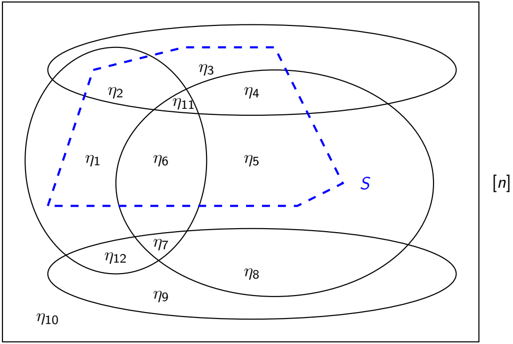
$\exp(L) = \exp(1/\delta^2)$ atoms $\mathcal{A}_1, \mathcal{A}_2, \ldots$

$\eta_j = \frac{|\mathcal{A}_j \cap S|}{|\mathcal{A}_j|}$ for atom $\mathcal{A}_j$

Weak Regularity Decomposition: Dense Graphs

$$|E(S,\overline{S})| \approx \sum_{\ell=1}^{L} c_\ell \cdot |S_1^\ell \cap S||S_2^\ell \cap \overline{S}|$$

Venn diagram of sets S_1^ℓ, S_2^ℓ for $\ell \in [L]$



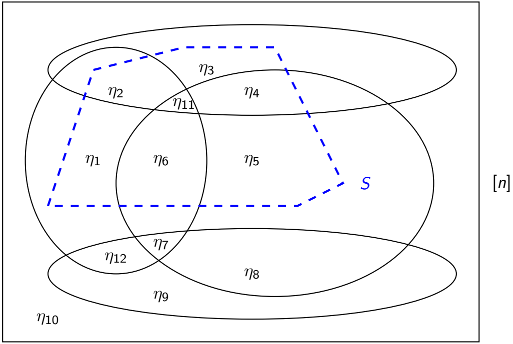
$\exp(L) = \exp(1/\delta^2)$ atoms $\mathcal{A}_1, \mathcal{A}_2, \ldots$

$$\eta_j = \frac{|\mathcal{A}_j \cap S|}{|\mathcal{A}_j|} \text{ for atom } \mathcal{A}_j$$

Weak Regularity Decomposition: Dense Graphs

$|E(S,\overline{S})| \approx \sum_{\ell=1}^{L} c_\ell \cdot |S_1^\ell \cap S||S_2^\ell \cap \overline{S}|$

Venn diagram of sets S_1^ℓ, S_2^ℓ for $\ell \in [L]$



$\exp(L) = \exp(1/\delta^2)$ atoms $\mathcal{A}_1, \mathcal{A}_2, \ldots$

$\eta_j = \frac{|\mathcal{A}_j \cap S|}{|\mathcal{A}_j|}$ for atom $\mathcal{A}_j$

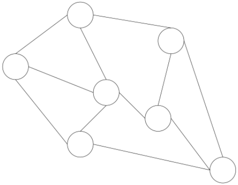
To find best S brute-force over a fine enough discretization of η_j's

Weak Regularity Decomposition: Sparse Graphs

Theorem (Oveis Gharan and Trevisan'13)

Expander graphs admit efficient weak regularity decompositions, so MaxCut can be approximated on them

(their result also holds for low threshold rank graphs)



Weak Regularity Decomposition: Sparse Tensors

Sparse Tensors on Splittable Structures

Let $W \subseteq [n]^k$ and $g\colon W \to [-1,1]$. We want to find $g \approx \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell}$ such that

$$\max_{S_1,\ldots,S_k \subseteq [n]} \left| \left\langle g - \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell}, 1_{S_1} \otimes \cdots \otimes 1_{S_k} \right\rangle \right| \leq \delta \cdot |W|,$$

and $L = O(1/\delta^2)$.

Weak Regularity Decomposition: Sparse Tensors

Sparse Tensors on Splittable Structures

Let $W \subseteq [n]^k$ τ-splittable and $g\colon W \to [-1,1]$. If $\tau \leq \text{poly}(\delta/k)$, there exists $\sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell}$ such that

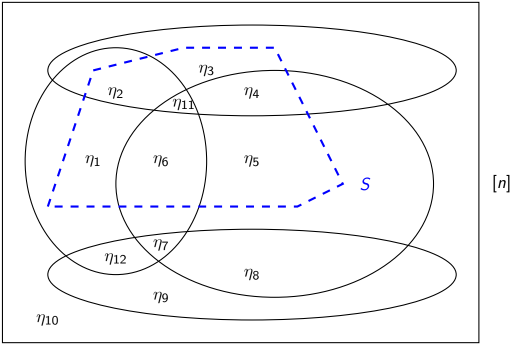
$$\max_{S_1,\ldots,S_k \subseteq [n]} \left| \left\langle g - \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell}, 1_{S_1} \otimes \cdots \otimes 1_{S_k} \right\rangle \right| \leq \delta \cdot |W|,$$

and $L = O(1/\delta^2)$.

Weak Regularity Decomposition: Sparse Tensors

Similar strategy works for k-CSPs (FK'96 and even to list decoding JST'21)

Venn diagram of sets $S_1^\ell, \ldots, S_k^\ell$ for $\ell \in [L]$



$\exp(kL) = \exp(k/\delta^2)$ atoms $\mathcal{A}_1, \mathcal{A}_2, \ldots$

$$\eta_j = \frac{|\mathcal{A}_j \cap S|}{|\mathcal{A}_j|} \text{ for atom } \mathcal{A}_j$$

Weak Regularity Decomposition: Sparse Tensors

Existential regularity decomposition for splittable tensors

Showing the existence of $\sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell} \approx g$ is not too hard

Reingold, Trevisan, Tulsiani and Vadhan [RTTV'08,TTV'09]

Weak Regularity Decomposition: Sparse Tensors

$\mathsf{CUT}^{\otimes k} = \{\pm 1_{S_1} \otimes \cdots \otimes 1_{S_k} \mid S_1, \ldots, S_k \subseteq [n]\}$
Let μ be a probability measure on W

function ExistentialWeakRegularityDecomposition($g \colon W \to [-1, 1]$)
    $h \leftarrow 0$
    while $\exists f \in \mathsf{CUT}^{\otimes k} \colon \langle g - h, f \rangle_\mu \geq \delta$ **do**
        $h \leftarrow h + \delta \cdot f$
    end while
    return h
end function

Weak Regularity Decomposition: Sparse Tensors

function ExistentialWeakRegularityDecomposition($g\colon W \to [-1,1]$)
    $h \leftarrow 0$
    while $\exists f \in \mathrm{CUT}^{\otimes k} \colon \langle g - h, f \rangle_\mu \geq \delta$ **do**
        $h \leftarrow h + \delta \cdot f$
    end while
    return h
end function

Claim: $\|g - h\|_\mu^2$ decreases by δ^2 at each iteration

$$\langle g - h - \delta \cdot f, g - h - \delta \cdot f \rangle_\mu = \langle g - h, g - h \rangle_\mu - 2\delta \langle g - h, f \rangle_\mu + \delta^2 \langle f, f \rangle_\mu$$

Weak Regularity Decomposition: Sparse Tensors

```
function ExistentialWeakRegularityDecomposition(g : W → [−1, 1])
    h ← 0
    while ∃f ∈ CUT^{⊗k} : ⟨g − h, f⟩_μ ≥ δ do
        h ← h + δ · f
    end while
    return h
end function
```

Claim: $\|g-h\|_\mu^2$ decreases by δ^2 at each iteration

$$\langle g-h-\delta\cdot f, g-h-\delta\cdot f\rangle_\mu = \langle g-h, g-h\rangle_\mu - 2\delta\underbrace{\langle g-h, f\rangle_\mu}_{\geq\delta} + \delta^2\underbrace{\langle f, f\rangle_\mu}_{\leq 1}$$

Weak Regularity Decomposition: Sparse Tensors

---

function ExistentialWeakRegularityDecomposition($g : W \to [-1, 1]$)
    $h \leftarrow 0$
    while $\exists f \in \mathrm{CUT}^{\otimes k} : \langle g - h, f \rangle_\mu \geq \delta$ **do**
        $h \leftarrow h + \delta \cdot f$
    end while
    return h
end function

Claim: $\|g - h\|_\mu^2$ decreases by δ^2 at each iteration

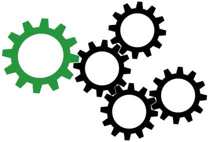
$$\langle g - h - \delta \cdot f, g - h - \delta \cdot f \rangle_\mu = \langle g - h, g - h \rangle_\mu - 2\delta \underbrace{\langle g - h, f \rangle_\mu}_{\geq \delta} + \delta^2 \underbrace{\langle f, f \rangle_\mu}_{\leq 1}$$

$$\leq \langle g - h, g - h \rangle_\mu - \delta^2$$

Weak Regularity Decomposition: Sparse Tensors

Near-linear time regularity decomposition for splittable tensors

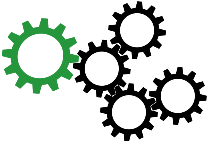
The more challenging steps are related to algorithmically finding a decomposition $\sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell} \approx g$ in time $\widetilde{O}_\delta(|W|)$ (and also proving list decoding)



Weak Regularity Decomposition: Sparse Tensors

Near-linear time regularity decomposition for splittable tensors

The more challenging steps are related to algorithmically finding a decomposition $\sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell} \approx g$ in time $\widetilde{O}_\delta(|W|)$ (and also proving list decoding)



We will give a (simplified) high-level description of the algorithmic ideas

Weak Regularity Decomposition: Algorithmic Ideas

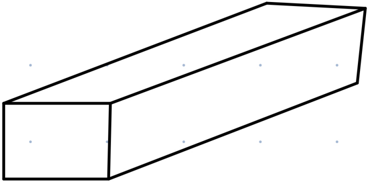
Iteratively "Splitting" the function $g \colon W \subseteq [n]^k \to [-1, 1]$

At step 1, find $h_1 = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{T^\ell} \approx g$
where $S_1^\ell \subseteq [n]$, $T^\ell \subseteq [n]^{k-1}$

Weak Regularity Decomposition: Algorithmic Ideas

Iteratively "Splitting" the function $g \colon W \subseteq [n]^k \to [-1, 1]$

At step 1, find $h_1 = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{T^\ell} \approx g$
where $S_1^\ell \subseteq [n]$, $T^\ell \subseteq [n]^{k-1}$
At step 2, find $h_2 = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{S_2^\ell} \otimes 1_{T^\ell} \approx h_1$
where $S_1^\ell, S_2^\ell \subseteq [n]$, $T^\ell \subseteq [n]^{k-2}$



(The sets and L change at each step and splittability is being crucially used)

Weak Regularity Decomposition: Algorithmic Ideas

Iteratively "Splitting" the function $g\colon W \subseteq [n]^k \to [-1,1]$

At step 1, find $h_1 = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{T^\ell} \approx g$
where $S_1^\ell \subseteq [n]$, $T^\ell \subseteq [n]^{k-1}$

At step 2, find $h_2 = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes 1_{S_2^\ell} \otimes 1_{T^\ell} \approx h_1$
where $S_1^\ell, S_2^\ell \subseteq [n]$, $T^\ell \subseteq [n]^{k-2}$

$$\vdots$$

At step $k-1$, find $h_{k-1} = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_k^\ell} \approx h_{k-2}$
where $S_1^\ell, \ldots, S_k^\ell \subseteq [n]$

(The sets and L change at each step and splittability is being crucially used)

Weak Regularity Decomposition: Algorithmic Ideas

Suppose we are at the begining of the sth step and we have

$$h_{s-1} = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_{s-1}^\ell} \otimes 1_{T^\ell},$$

where $S_1^\ell, \ldots, S_{s-1}^\ell \subseteq [n]$, $T^\ell \subseteq [n]^{k-(s-1)}$

Weak Regularity Decomposition: Algorithmic Ideas

Suppose we are at the begining of the sth step and we have

$$h_{s-1} = \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_{s-1}^\ell} \otimes 1_{T^\ell},$$

where $S_1^\ell, \ldots, S_{s-1}^\ell \subseteq [n]$, $T^\ell \subseteq [n]^{k-(s-1)}$

1: **function** WeakRegularityDecomposition($g = h_{s-1}$)
2: $\quad h \leftarrow 0$
3: $\quad$**while** $\exists f = \pm 1_{S_1'} \otimes \cdots \otimes 1_{S_s'} \otimes 1_{T'}, S_i' \subseteq [n], T' \subseteq [n]^{k-s} : \langle g - h, f \rangle \geq \delta$ **do**
4: $\quad\quad h \leftarrow h + \delta \cdot f$
5: $\quad$**end while**
6: $\quad$**return** h
7: **end function**

Weak Regularity Decomposition: Algorithmic Ideas

```
function WeakRegularityDecomposition(g = h_{s-1})
    h ← 0
    while ∃f = ±1_{S'_1} ⊗ ··· ⊗ 1_{S'_s} ⊗ 1_{T'}, S'_i ⊆ [n], T' ⊆ [n]^{k-s} : ⟨g − h, f⟩ ≥ δ do
        h ← h + δ · f
    end while
    return h
end function
```

function WeakRegularityDecomposition($g = h_{s-1}$)

$h \leftarrow 0$

while $\exists f = \pm 1_{S'_1} \otimes \cdots \otimes 1_{S'_s} \otimes 1_{T'}, S'_i \subseteq [n], T' \subseteq [n]^{k-s} : \langle g - h, f \rangle \geq \delta$ **do**

$h \leftarrow h + \delta \cdot f$

end while

return h

end function

Can we make the red step efficient?

Weak Regularity Decomposition: Algorithmic Ideas

---

function WeakRegularityDecomposition($g = h_{s-1}$)
  $h \leftarrow 0$
  while $\exists f = \pm 1_{S'_1} \otimes \cdots \otimes 1_{S'_s} \otimes 1_{T'}, S'_i \subseteq [n], T' \subseteq [n]^{k-s} : \langle g - h, f \rangle \geq \delta$ **do**
    $h \leftarrow h + \delta \cdot f$
  end while
  return h
end function

Can we make the red step efficient?
Difficulty: How can we find sets $S'_1, \ldots, S'_s \subseteq [n]$, $T' \subseteq [n]^{k-s}$?

Weak Regularity Decomposition: Algorithmic Ideas

function WeakRegularityDecomposition($g = h_{s-1}$)
  $h \leftarrow 0$
  while $\exists f = \pm 1_{S'_1} \otimes \cdots \otimes 1_{S'_s} \otimes 1_{T'}, S'_i \subseteq [n], T' \subseteq [n]^{k-s} : \langle g - h, f \rangle \geq \delta$ **do**
    $h \leftarrow h + \delta \cdot f$
  end while
  return h
end function

Can we make the red step efficient?
Suppose Merlin tells us (all but the last two sets)

$S'_1, \ldots, S'_{s-1} \subseteq [n]$

Weak Regularity Decomposition: Algorithmic Ideas

We still need to find $S'_s \subseteq [n]$ and $T' \subseteq [n]^{k-s}$

$$\max_{S'_s, T'} \quad |\langle \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_{s-1}^\ell} \otimes 1_{T^\ell}, 1_{S'_1} \otimes \cdots \otimes 1_{S'_s} \otimes 1_{T'} \rangle|$$

Weak Regularity Decomposition: Algorithmic Ideas

We still need to find $S'_s \subseteq [n]$ and $T' \subseteq [n]^{k-s}$

$$\max_{S'_s, T'} \quad |\langle \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_{s-1}^\ell} \otimes 1_{T^\ell}, 1_{S'_1} \otimes \cdots \otimes 1_{S'_s} \otimes 1_{T'} \rangle| =$$

$$\max_{S'_s, T'} \quad |\sum_{\ell=1}^{L} \underbrace{c_\ell \prod_{i=1}^{s-1} \langle 1_{S_i^\ell}, 1_{S'_i} \rangle}_{\gamma_\ell} \langle 1_T^\ell, 1_{S'_s} \otimes 1_{T'} \rangle||$$

Weak Regularity Decomposition: Algorithmic Ideas

We still need to find $S'_s \subseteq [n]$ and $T' \subseteq [n]^{k-s}$

$$\max_{S'_s, T'} \quad |\langle \sum_{\ell=1}^{L} c_\ell \cdot 1_{S_1^\ell} \otimes \cdots \otimes 1_{S_{s-1}^\ell} \otimes 1_{T^\ell}, 1_{S'_1} \otimes \cdots \otimes 1_{S'_s} \otimes 1_{T'} \rangle| =$$

$$\max_{S'_s, T'} \quad |\sum_{\ell=1}^{L} \underbrace{c_\ell \prod_{i=1}^{s-1} \langle 1_{S_i^\ell}, 1_{S'_i} \rangle}_{\gamma_\ell} \langle 1_T^\ell, 1_{S'_s} \otimes 1_{T'} \rangle| =$$

$$\max_{S'_s, T'} \quad |\langle \underbrace{\sum_{\ell=1}^{L} \gamma_\ell \cdot 1_T^\ell}_{M}, 1_{S'_s} \otimes 1_{T'} \rangle|$$

Weak Regularity Decomposition: Algorithmic Ideas

$$\max_{S'_s, T'} \quad |\langle M, 1_{S'_s} \otimes 1_{T'} \rangle|$$

Good News 1

The problem above is the well-known CUT norm optimization which has an SDP-based approximation algorithm by Alon–Naor'04

Weak Regularity Decomposition: Algorithmic Ideas

$$\max_{S'_s, T'} \quad |\langle M, 1_{S'_s} \otimes 1_{T'} \rangle|$$

Good News 1

The problem above is the well-known CUT norm optimization which has an SDP-based approximation algorithm by Alon–Naor'04

Good News 2

M is an $n \times n^{k-s}$ matrix, but it is very sparse in our case ($O_\epsilon(n)$ non-zero entries). We can use near-linear time SDP solvers such as Arora–Kale'06

Weak Regularity Decomposition: Algorithmic Ideas

Question

Do we really need Merlin to tell us $S'_1, \ldots, S'_{s-1}$?

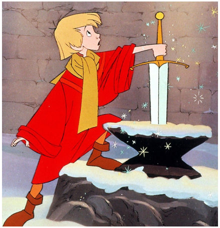
Weak Regularity Decomposition: Algorithmic Ideas

Question

Do we really need Merlin to tell us $S'_1, \ldots, S'_{s-1}$?

Good News 3

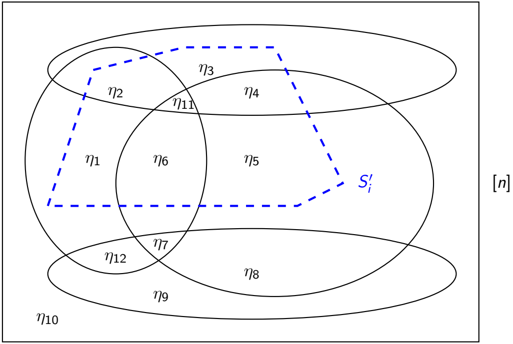
No! Arthur is enough!



Weak Regularity Decomposition: Algorithmic Ideas

We have a low-complexity set system, so we can (approximately) generate $S'_1, \ldots, S'_{s-1}$ ourselves!

Venn diagram of sets $S_1^\ell, \ldots, S_{s-1}^\ell$ for $\ell \in [L]$



$\exp((s-1)L) \leq \exp(k/\delta^2)$ atoms $\mathcal{A}_1, \mathcal{A}_2, \ldots$

That's all.

Thank you!

Questions?