Cut Sparsification and Succinct Representation of Submodular Hypergraphs

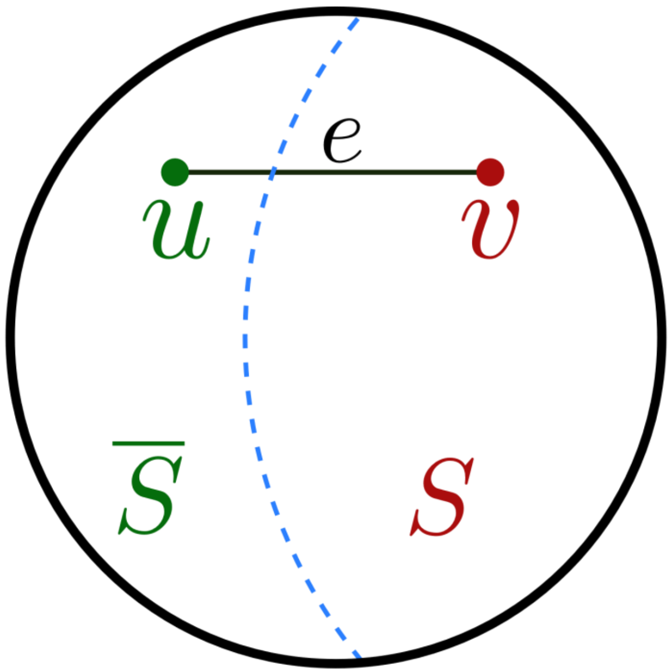
Robert Krauthgamer, Weizmann Institute of Science

Joint work with Yotam Kenneth

Simons Institute, July 2024

Cut Intuitions: Graphs

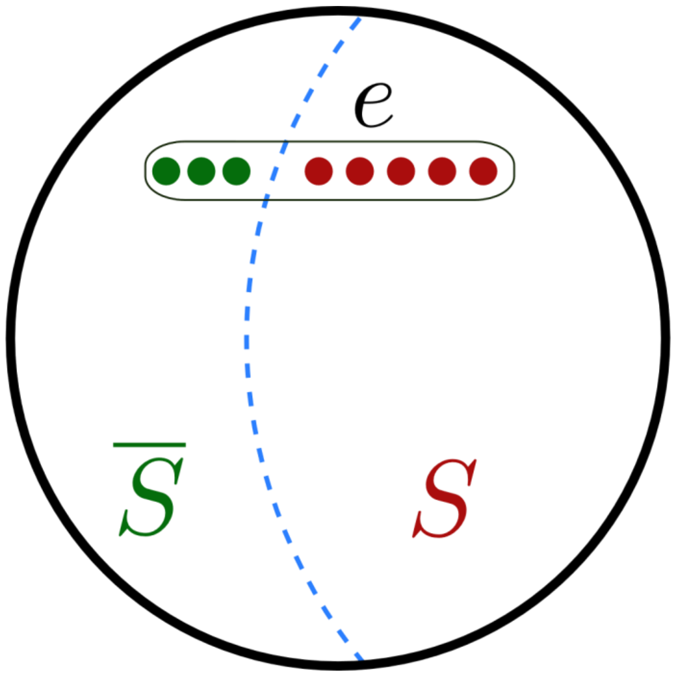
- $G = (V, E, w)$ is a graph



$$\mathrm{cut}_G(S) = \sum_{e \in E} \mathbb{1}_{e \in S \times \overline{S}} w_e$$

Cut Intuitions: Hypergraphs

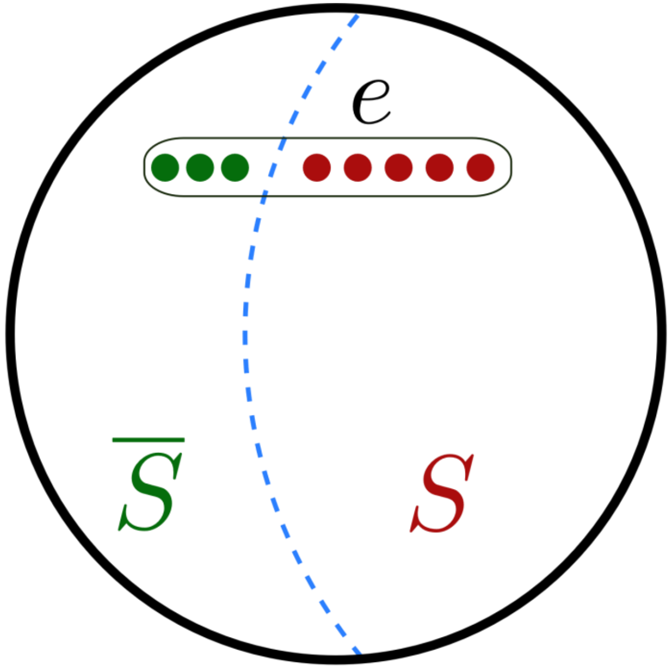
- $\mathrm{H} = (V, E, w)$ is a hypergraph



$$\mathrm{cut}_H(S) = \sum_{e \in E} \underbrace{\mathbb{1}_{0<|e \cap S|<|e|}}_{\text{Why binary?}} \cdot w_e$$

Cut Intuitions: Submodular Hypergraphs

- Associate each hyperedge $e \in E$ with a splitting function $g_e: 2^e \rightarrow \mathbb{R}_+$



$$\text{cut}_H(S) = \sum_{e \in E} g_e(S \cap e)$$

Properties of Splitting Functions

- Splitting functions should have two properties
  - Submodularity (diminishing returns):

$$\forall S, T \subseteq e, \qquad g_e(S \cup T) + g_e(S \cap T) \leq g_e(S) + g_e(T)$$

  - "Irrelevance": $g_e(\emptyset) = 0$

- Examples of Splitting functions:
  - All-or-Nothing: $g_e(S) = 1_{\{e:0<|S\cap e|<|e|\}}$
  - Small Side: $g_e(S) = \min(|S|, |e \setminus S|)$
  - Capped Small Side: $g_e(S) = \min(|S|, |e \setminus S|, \boldsymbol{c})$ for some $c > 0$
  - Budget Additive: $g_e(S) = \min(|S|, \boldsymbol{c})$ for some $c > 0$

Uses of Submodular Hypergraphs

- Clustering [Li & Milenkovich‘17; Li & Milenkovich‘18]
- Data Summarization [Gomese & Krause‘10; Lin & Bilmes‘10; Tschiatschek, Iyer, Wei & Bilmes‘14]

Model:

Valuation/Similarity function

Subset selection

$$\max_{S \subseteq V, |S| \leq k} \sum_i f_i(S)$$

- Welfare Maximization
  - Approximation Algorithms [Feige‘09, Feige & Vondrak‘06]
  - Mechanism Design [Dobzinski & Schapira‘06, Assadi & Singla‘20]

Model: Decomposable submodular function

Research Questions

- **Goal:** find a <u>small</u> $H' = (\boldsymbol{V}, E', g')$ such that

$$\forall S \subseteq V, \qquad cut_{H'}(S) \in (1 \pm \epsilon) cut_H(S)$$

  - Small: number of hyperedges; or storage complexity

- **Hyperedge Sparsification:** $E' \subseteq E$ with small $|E'|$

Today

1. Graphs admit sparsifiers with $O(\epsilon^{-2} n)$ edges [BK'96,BSS'14]; what is the analogue for submodular hypergraphs [RY22]?
2. Better bounds for specific families?

- **Succinct Representation:** encoding using few bits

3. Store all cut values more efficiently than a subgraph [ACKQWZ'16]?

Hyperedge Sparsification

Known Results

Splitting Functions	Lower Bound	Upper Bound	Comments	Reference
Specific Functions				
All-or-Nothing	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[KK15, CKN20]
Small Side	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[AGK14, ADKKP16]
Directed Hypergraph	$\Omega(\epsilon^{-1}n^2)$	$\tilde{O}(\epsilon^{-2}n^2)$		[SY19, KKTY21,OST23]

Known Results

Splitting Functions	Lower Bound	Upper Bound	Comments	Reference
Specific Functions				
All-or-Nothing	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[KK15, CKN20]
Small Side	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[AGK14, ADKKP16]
Directed Hypergraph	$\Omega(\epsilon^{-1}n^2)$	$\tilde{O}(\epsilon^{-2}n^2)$		[SY19, KKTY21,OST23]
General Families				
General Submodular	$\Omega(\epsilon^{-1}n^2)$	$\tilde{O}(\epsilon^{-2}n^2B_{\mathrm{H}})$	B_{H} can be exponential in n	[RY22]
Monotone Functions	-			[RY22, KZ23]
Symmetric Functions	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[JLLS23]

Known Results + Ours

Splitting Functions	Lower Bound	Upper Bound	Comments	Reference
Specific Functions				
All-or-Nothing	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[KK15, CKN20]
Small Side	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[AGK14, ADKKP16]
Directed Hypergraph	$\Omega(\epsilon^{-1}n^2)$	$\tilde{O}(\epsilon^{-2}n^2)$		[SY19, KKTY21,OST23]
General Families				
General Submodular	$\Omega(\epsilon^{-1}n^2)$	$\tilde{O}(\epsilon^{-2}n^3)$		[*here*]
Monotone Functions	-	$\tilde{O}(\epsilon^{-2}n^2)$	now $\tilde{O}_\epsilon(\mathrm{n})$ [KPS24]	[*here*]
Symmetric Functions	$\Omega(\epsilon^{-2}n)$	$\tilde{O}(\epsilon^{-2}n)$		[JLLS23]
Finite-Spread	-	$\tilde{O}(\epsilon^{-2}n\,\mu_H)$	$\mu_H = \max_{e\in E} \frac{\max_{T\subseteq e} g_e(T)}{\min_{S\subseteq e} g_e(S)}$	[*here*]

Main Result

Theorem 1:

Every $H = (V, E, g)$ admits a sparsifier with $O(\epsilon^{-2} n^3)$ edges

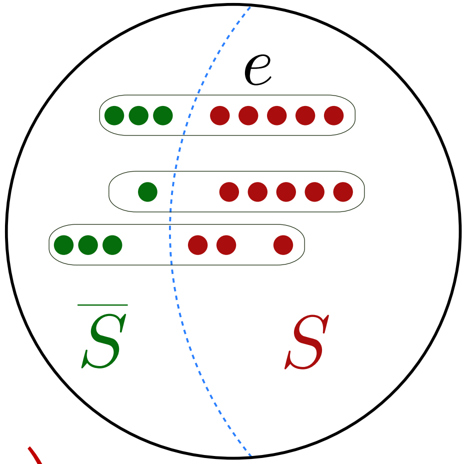
- **Need to prove**
  - Approximation guarantee
  - Sparsifier Size

Proof Overview

- **Approach:** Importance Sampling
  - Quantify for every $e \in E$ its “importance” $\sigma_e \in [0,1]$
    - Intuitively – its relative contribution to a specific/any cut
  - Sample each $e \in E$ with probability $p_e = \min(1, M\sigma_e)$ for parameter $M > 0$
    - Scale each sampled hyperedge by p_e^{-1} and add it to $H' = (V, E', g')$

- **Need to prove**
  - Approximation Guarantee – by Chernoff bound
  - Sparsifier Size – by its expectation $\mathbb{E}[|E'|] = \sum_e p_e$

Sparsifying a Single Cut



- Fix $S \subseteq V$. Define importance of e to $\mathrm{cut}_H(S)$ as

$$\sigma_e(S) := \frac{g_e(S)}{\sum_{f \in E} g_f(S)} = \frac{g_e(S)}{cut_H(S)}$$

- By Chernoff

$$\Pr\left(cut_{H'}(S) \notin (1 \pm \epsilon) cut_H(S)\right) \leq \exp\left(-\frac{\epsilon^2 cut_H(S)}{3 \max_{e \in E} p_e^{-1} g_e(S)}\right) \leq e^{-\Omega(\epsilon^2 M)}$$

  - Suitable $M = O(\epsilon^{-2})$ suffices

- Sparsifier size:

$$\mathbb{E}[|E'|] \leq M \sum_{e \in E} \sigma_e(S) = M$$

Sparsifying All Cuts [RY22]

- Importance of e overall (= to all cuts)

$$\sigma_e = \max_{S \subseteq e} \sigma_e(S) = \max_{S \subseteq e} \frac{g_e(S)}{cut_H(S)}$$

- For all $S \subseteq V$ we have $\sigma_e \geq \sigma_e(S)$ and thus

$$\Pr\left(cut_{H'}(S) \notin (1 \pm \epsilon)cut_H\,(S)\right) \leq e^{-\Omega(\epsilon^2 M)}$$

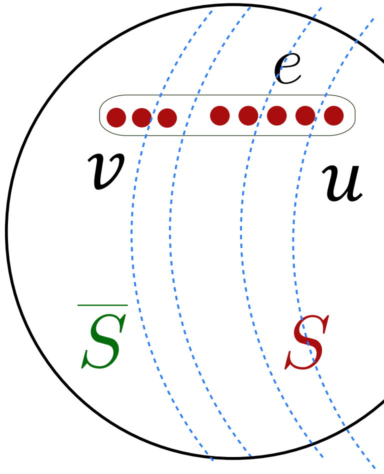
  - Suitable $M = O(\epsilon^{-2} n)$ suffices for union bound over 2^n cuts
- Sparsifier size:

$$\mathbb{E}[|E'|] \leq M \sum_{e \in E} \sigma_e = O(\epsilon^{-2} n^2 B_H)$$

  - Where B_H is number of extreme points of polytope of g_e
  - Unfortunately, B_H can be exponential in n

Sparsifying All Cuts: Our Bound

- **Main idea**: Bound σ_e by something easier to analyze



- **Definition**: The *minimum directed* $u \to v$ *cut* on e

$$g_e^{u \to v} = \min_{\substack{S \subseteq e \\ u \in S, v \notin S}} g_e(S)$$

- **Lemma**: Can approximate $g_e(S)$ by sum of minimum directed cuts

$$\max_{u \in S, v \in e \setminus S} g_e^{u \to v} \leq g_e(S) \leq \sum_{u \in S, v \in e \setminus S} g_e^{u \to v}$$

Factor $\leq |S|^2$

Sparsifying All Cuts: Our Bound

- Set the approximate importance by

$$\rho_e = \sum_{(u,v)\in V\times V} \frac{g_e^{u\to v}}{\sum_{f\in E} g_f^{u\to v}}$$

- By lemma, for all $S \subseteq V$

$$\sigma_e(S) = \frac{g_e(S)}{\sum_{f\in E} g_f(S)}$$

- **Size ana**

Lemma:

$$\max_{u\in S, v\in e\setminus S} g_e^{u\to v} \le g_e(S) \le \sum_{u\in S, v\in e\setminus S} g_e^{u\to v}$$

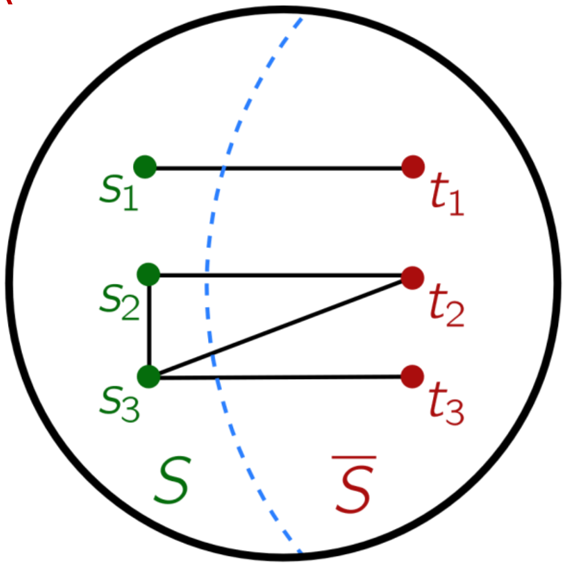
Lemma Intuition

- **Lemma**: Can approximate $g_e(S)$ by sum of minimum directed cuts

$$\max_{u \in S, v \in e \setminus S} g_e^{u \to v} \le g_e(S) \le \sum_{u \in S, v \in e \setminus S} g_e^{u \to v}$$

- Lower Bound – trivial
- Upper Bound – submodularity of optimal cuts

  Intuition – bounding a cut by all pairwise flows



$\mathrm{cut}_G(S) \le \sum_{s \in S, t \in \overline{S}} \mathrm{cut}(\{\boldsymbol{s}\}, \{\boldsymbol{t}\})$

Improved Bound for Monotone Case

Theorem 2:

Every $H = (V, E, g)$ with monotone splitting functions admits a sparsifier with $O(\epsilon^{-2}n^2)$ edges

- Similar approach but with different lemma:

$$\max_{v \in V} g_e(\{v\} \cap S) \leq g_e(S) \leq \sum_{v \in V} g_e(\{v\} \cap S)$$

Succinct Representation

Succinct Encoding of All Cut Values

- **Question:** Is there a more succinct encoding than sparsifiers?
  - For graphs: No! [ACKQWZ'16]
  - Possible approaches: non-subgraph sparsifiers? use different hyperedges/splitting functions?

Theorem 3: For budget-additive splitting $g_e(S) = \min(|S|, K)$ with $K = \Omega(|e|)$,

(1) encoding a reweighted-subgraph sparsifier requires $\Omega(n^2)$ bits;

(2) but non-subgraph sparsifiers can be encoded with $\tilde{O}(\epsilon^{-6} n)$ bits.

Encoding of Budget-Additive Splitting

- **Lower bound:** "encode" $\Omega(n^2)$ bits into hypergraphs H that must have distinct subgraph $(1+\epsilon)$-sparsifiers H'

- **Upper Bound:** "Break" large hyperedges into small hyperedge
- Outline: Two steps of $(1+\epsilon)$-approximation:
  - Deform each g_e into many small hyperedges (also budget-additive)
    - Small is $O\left(\epsilon^{-2}\left(\frac{|E|}{K}\right)\log|e|\right)$; it implies low spread as well
    - Many is $O(\epsilon^{-2}|e|^2)$
    - Generating small hyperedges: Subsample vertices at rate p and scale by $1/p$
  - Apply our sparsification for low spread μ_H → straightforward encoding

Conclusion

Conclusion

- All submodular hypergraphs admit size $O(\epsilon^{-2}n^3)$ sparsifiers
  - Some admit even smaller ones (details in paper)

- **Open Questions:**
  - Close the gap between $\Omega(n^2)$ lower bound and $O(n^3)$ upper bound?
  - Families with smaller sparsifiers (other than symmetric and monotone)?
  - Characterize a smooth tradeoff between those families?