Streaming CSPs

Madhu Sudan
Harvard University

Based on joint works with Chi-Ning Chou, Alexander Golovnev, Noah Singer, Ameya Velingker and Santhoshini Velusamy.
This Talk

- CSPs (& approximation & streaming)
- Background
- Our results
- Future directions
Constraint Satisfaction Problems (CSPs)

- Class of infinitely many problems.
Constraint Satisfaction Problems (CSPs)

- Class of infinitely many problems.

- Specified by \( q, k \) and family \( F \subseteq \{ f : [q]^k \rightarrow \{0,1\} \} \).

- Instance of MaxCSP\((F)\):
  \[ \Psi = (X_1, ..., X_n; C_1, ..., C_m) \]
  Constraint \( C_j(X) = f_j(X_{i_1(j)}, ..., X_{i_k(j)}) ; f_j \in F \)

- \( \text{opt}(\Psi) \equiv \max_{a \in [q]^n} \{ \sum_j C_j(a) \} \)
Constraint Satisfaction Problems (CSPs)

Class of infinitely many problems. ▶️ MaxCSP(F)

Specified by $q, k$ and family $F \subseteq \{f: [q]^k \rightarrow \{0,1\}\}$.

Instance of MaxCSP(F): $\Psi = (X_1, ..., X_n; C_1, ..., C_m)$ ;
Constraint $C_j(X) = f_j(X_{i_1(j)}, ..., X_{i_k(j)})$; $f_j \in F$

$\text{opt}(\Psi) \equiv \max_{a \in [q]^n} \{\sum_j C_j(a)\}$

Examples:

MaxCut ($F = \{\oplus\}$)
MaxDicut ($F = \{x \land \overline{y}\}$)
Max 3SAT ($|F| = 11$), Max Exact 3SAT ($|F| = 4$), Max $q$-Colorability.
Constraint Satisfaction Problems (CSPs)

“Special Case”: Boolean MaxCSP(F):

- \( q = 2 \);
- Constraints \( C = f(X_1, \overline{X_2}, ..., \overline{X_k}) \)
  - Constraints applied to literals.

Warning: MaxCSP(F) ≠ Boolean MaxCSP(F)

\( \forall F \exists G \) s.t. Boolean MaxCSP(F) = MaxCSP(G). (converse not true. MaxCut, MaxDicut ...)

April 8, 2022
Streaming CSPs
Streaming & Approximation

- Streaming input
  - Streaming: \( s(n) \)-space algorithm – gets one constraint at a time.
  - Sketching: Maintains sketch \( S(\sigma) \) with \(|S(\sigma)| \leq s(n)\);
    - Restriction: \( (S(\sigma), S(\tau)) \rightarrow S(\sigma \circ \tau) \)
  - Space milestones: polylog, sqrt, or (nearly-) linear.

- Approximations:
  - Usual notion: \( \alpha \)-approximation:
    - Output \( v \) s.t. \( \alpha \cdot \text{opt}(\Psi) \leq v \leq \text{opt}(\Psi) \)
  - Refined notion: \( (\gamma, \beta) \)-distinguishability:
    - Output Yes if \( \text{opt}(\Psi) \geq \gamma \), No if \( \text{opt}(\Psi) \leq \beta \)
  - Equivalence: \( \alpha = \min_{\gamma} \max_{\beta} \beta / \gamma \)
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut

April 8, 2022

Streaming CSPs
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut

[Diagram of points connected by lines]

April 8, 2022
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut

April 8, 2022
Streaming CSPs
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut
Streaming MaxCut

[Diagram showing a graph with points connected by lines]
Brief History

- 2011 Bertinoro W’shop (P. 45): “We know nothing +/-”
- 2015-19: Lower bounds for MaxCut [KKS, KKSv, KK]
  - Kapralov-Krachun: $\frac{1}{2} + \epsilon$-approximation requires $\Omega(n)$-space (in $n$ vertex graph) (streaming).
- 2017-20: Algorithms for Max DiCut, Max SAT
  - Guruswami-Velingker-Velusamy: DiCut
  - Chou-Golovnev-Velusamy: DiCut, Max SAT
- 2020: Sketching Classification
  - Chou-Golovnev-Velusamy: Classify all Boolean MaxCSP with $k = q = 2$
Why study CSPs

- Contain some problems of direct interest
  - Max Cut, Max Dicut, Max colorability
- Allow possibility of classification!
- Highlight general algorithms
  - Norm approximations (already used)
  - Local Exploration
  - Crude snapshots
- Identify Phenomena:
  - No $2^{\sqrt{\log n}}$-space algorithms?
Our Results

- [CGSV21]:
  - Dichotomy for sketching (polylog vs. sqrt)
  - Implications for dynamic CSPs [Li, Nguyen, Woodruff, Ai Hu Li Woodruff]
  - Non-trivial (polylog) algorithms for infinitely many dynamic CSPs
  - General sqrt lower bounds for broad classes (“one-wise ind.”, “padded one-wise ind.”) in insertion-only setting.

- [CGSVVV 21]: Linear space lower bounds for subclass of “one-wise-ind”.
  - Pins approximability of all MaxCSP($F$) to within q-factor (trivial alg vs. linear space).

- [SSV21]: No sublinear algorithms for “Ordering CSPs”
GVV+CGV Algorithms for Max DiCut

- Define \( \text{Bias}(v) \defeq \text{indeg}(v) - \text{outdeg}(v) \)
- \( \text{Bias}(G) \defeq \frac{1}{2} \sum_v |\text{Bias}(v)| \)
- Claim 1 [GVV]:
  - Bias can be estimated arbitrarily closely in polylog space.
  - \( \text{Dicut}(G) \leq \frac{\text{Bias}(G) + m}{2} \)
  - \( \text{Bias}(G) \leq \text{Dicut}(G) \) (Greedy rounding)
  - Output: \( \max \left\{ \frac{m}{4}, \text{Bias}(G) \right\} \Rightarrow 2/5\)-approx.

- Claim 2 [CGV]: \( \text{Bias}(G) \leq \frac{m}{3} \Rightarrow \text{Dicut}(G) \geq \frac{m}{4} + \frac{\text{Bias}(G)^2}{4(m-2\text{Bias}(G))} \)
  (Rand. Rounding w.p. \( \frac{1}{2} - \frac{\text{Bias}(G)}{2(m-2\text{Bias}(G))} \)) \( \Rightarrow 4/9\)-approx.
Generalizing to other CSPs: Challenges

- What is bias, say for $f(x, y, z) = x \land (y \oplus z)$
  - E.g.: $x \land (a_1 \oplus a_2), a_3 \land (x \oplus y), y \land (x \oplus a_4)$

- Why is $\ell_1$-estimation useful in Dicut approximation
  - $\ell_1(x_1, \ldots, x_n) = \max_{a_1 \ldots a_n \in \{-1,1\}} \{\sum_i a_ix_i\}$
  - Useful generalization: $||M||_{1,\infty} := \max_{a \in [q]^n} \{\sum_i M_{i,a(i)}\}$
  - Computable in polylog space

- Dicut Analysis: Graph theory, some 3-var. calculus, Why did rounding end up optimal?
Algorithms - 1

- Abstracting GVV, CGV algorithm (|F| = 1)
  - For every \( i \in [n] \) maintain (non-normalized) distribution \( D_i \) over \([q] \).
  - In the end, round using max. likelihood over \( D_i \).
  - Update?
    - If \( X_i \) appears as \( j \)th var in constraint, need to perturb \( D_i \). How?
    - Idea:
      - (Initially) Guess & fix some \( \lambda_{j\sigma}, j \in [k], \sigma \in [q] \)
      - (Update on \((i, j)\)) Add \( \lambda_{j\sigma} \) to \( D_i(\sigma) \) for every \( \sigma \in [q] \)
Algorithms - 2

- Streaming implementation?
  - Can’t output rounding of $D_i$ for $i \in [n]$
  - But can compute $\frac{1}{m} \sum_i \max_{\sigma} D_i(\sigma)$
    - $(1, \infty)$ norm of $D \in \mathbb{R}^{n \times q}$. [Andoni Krauthgamer Onak]

- What $\lambda_{j\sigma}$ to use?
  - Don’t know! (e.g., $f(x_1, \ldots, x_k) = \text{sign}(\sum_j (-1.1)^j x_j)$)

- “Theorem”: Algorithm works if $\exists \lambda_{j\sigma}, \tau$ s.t.
  - $\text{opt}(\Psi) \geq \gamma \Rightarrow \frac{1}{m} \sum_i \max_{\sigma} D_i(\sigma) \geq \tau$
  - $\text{opt}(\Psi) \leq \beta \Rightarrow \frac{1}{m} \sum_i \max_{\sigma} D_i(\sigma) < \tau$

- Theorem [CGSV]: Criterion above is decidable given $F$. 
Only use
const set
by \( \overline{a} \)

\[ \text{Opt} \geq 0 \]

\[ \text{Opt} \leq \beta \]
Decidability & Criterion

- Main insight: Suffices to look at instances on $kq$ variables.
- Instance = distribution of constraints over $kq$ variables.
- ...
- $\exists$ convex sets $K_Y^Y$ and $K_N^N \subseteq \Delta([q])^k \subseteq \mathbb{R}^{qk}$ s.t. Algorithm exists if $K_Y^Y \cap K_N^N = \emptyset$
Complexity?

- If Algorithm does not exist, then?
  - There exist distributions $D_Y$ and $D_N \in \Delta([q]^k)$ with matching (one-wise) marginals corresponding to two “interesting” instances.

- Theorem: If such $D_Y$ and $D_N$ exist then $o(\sqrt{n})$ space sketching algorithm can’t solve $(\gamma, \beta)$-distinguishability!

- One-wise Theorem: If such $D_Y$ and $D_N$ exist with $D_N = \text{Unif}([q]^k)$ then $o(\sqrt{n})$ space streaming algorithm can’t solve $(\gamma, \beta)$-distinguishability!

Extends to Padded One Wise ($k=q=2$):

$$D_Y = \tau D_0 + (1 - \tau) D_1; D_N = \tau D_0 + (1 - \tau) \text{Unif}([q]^k)$$
Underlying Communication Problem

- Randomized Mask Detection:
  - Alice $\leftarrow X \sim \text{Unif}([q]^n)$
  - Bob $\leftarrow \text{Seq. of disjoint masked projections of } X$
    - Projection: $X \rightarrow (S, X|_S)$ $S \subseteq [n]$, $|S| = k$
    - Masked: $(S, X|_S) \rightarrow (S, X|_S + m \text{ (mod } q))$
      - YES: $m \sim D_Y$; NO: $m \sim D_N$
  - Sequence: $S_1, ..., S_{0001n}$
  - Disjoint: $S_i \cap S_j = \emptyset$
  - Challenge: Distinguish YES from NO
- Thm: Requires $\Omega(\sqrt{n})$ bits Alice $\rightarrow$ Bob if marginals of $D_Y$ and $D_N$ match.
Linear Space Lower Bounds

- \( F = \{f\} \) case
- \( \rho(F) \equiv \inf_{\Psi} \text{opt}(\Psi) \)
- Trivial approximation factor = \( \rho \)
- Theorem: If \( f^{-1}(1) \) contains a “line” then \( F \) has no non-trivial approximation in sublinear space.
  - Line \( \equiv \{a + (i, \ldots, i)(\text{mod } q)|i \in [q]\}, \text{ for } a \in [q]^k \)
- Thm’: If \( f^{-1}(1) \) contains \( \omega \) fraction of a line then \( \text{MaxCSP}(F) \) has no sublinear approximation better than \( \frac{\rho}{\omega} \)
- Proof: Extends (and compresses?) Kapralov-Krachun.
Open Questions

- Sketching = Streaming in sqrt setting?
  - Extend sqrt lower bounds from sketching to streaming?
  - Challenge: Walk length algorithm!
Open Questions

- Sketching = Streaming in sqrt setting?
- Extend sqrt lower bounds from sketching to streaming!
- Challenge: Walk length algorithm!

(Suggested by Saxena et al.)

**Walk length algorithm:**

For every pattern $\rightarrow \leftarrow \cdots \rightarrow$

Pick $S \subset [n], |S| = \text{polylog}(n)$

Follow patterned path from $S$ till we run out of stream.

Look at the distribution of path lengths when we stop.

Different for $\text{opt} \geq \gamma$ vs. $\text{opt} \leq \beta$?
Open Questions

- Sketching = Streaming in sqrt setting?
  - Extend sqrt lower bounds from sketching to streaming!
  - Challenge: Walk length algorithm!

- Sqrt = Linear?
  - Is there a dichotomy at linear space?
  - Challenge: Template-based algorithms!
  - Can Dicut approximation be improved with \( o(n) \) space?
Open Questions

- Sketching = Streaming?
  - Extend sqrt lower bounds from sketching to streaming!
  - Challenge: Walk length algorithm!

- Sqrt = Linear?
  - Is there a dichotomy at linear space?
  - Challenge: Template-based algorithms!

- Can \textbf{Dicut} approximation be improved with \( o(n) \) space?

April 8, 2022

Streaming CSPs

Template algorithm (for bounded degree graphs)

Classify vertices by \((\text{indeg}, \text{outdeg})\)

Estimate fraction of edges from class \((i, j) \rightarrow (k, \ell)\)

This is the template

Different for \(\text{opt} \geq \gamma\) vs. \(\text{opt} \leq \beta\)?
Open Questions

- Sketching = Streaming?
  - Extend sqrt lower bounds from sketching to streaming!
  - Challenge: Walk length algorithm!

- Sqrt = Linear?
  - Is there a dichotomy at linear space?
  - Challenge: Template-based algorithms!
  - Can Dicut approximation be improved with o(n) space?
  - \( \omega(n) \)-lower bounds? (Trivial: \( O(n \log n) \ldots \)
Thank You!
Thank You!