APPROXIMATION ALGORITHMS

MULTI-WAY CUTS & HARDNESS OF APPROXIMATION

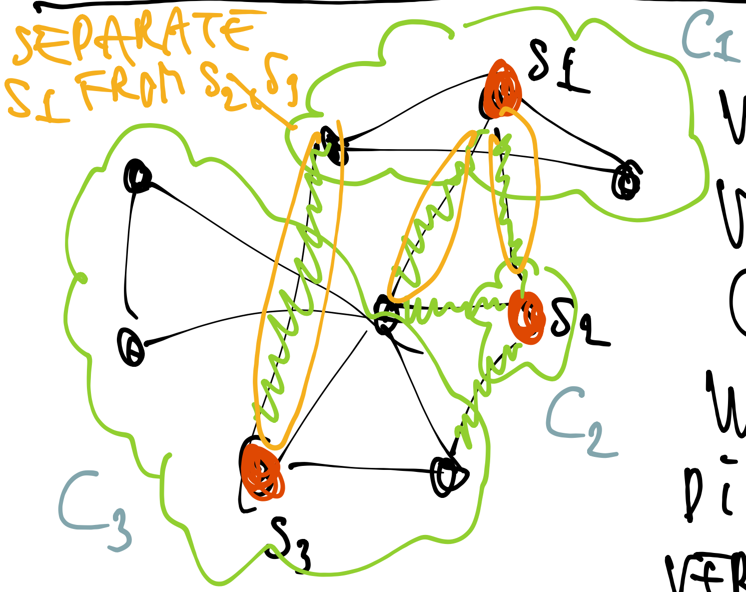
RASMUS PAGH
UNIVERSITY OF COPENHAGEN

TODAY

- PART I: MULTI-WAY CUTS
  - 2-APPROXIMATION BY REPEATED CUTS
  - 3/2-APPROXIMATION USING RANDOMIZED ROUNDING

- PART II: HARDNESS OF APPROXIMATION
  - REDUCTIONS FROM NP-HARD PROBLEMS
  - FINE-GRAINED REDUCTIONS

k-WAY CUT PROBLEM



WEIGHTED ($c_e \geq 0$) UNDIRECTED GRAPH WITH k DISTINGUISHED VERTICES $s_1, \dots, s_k$

GOAL:

REMOVE EDGES SUCH THAT $s_1, \dots, s_k$ ARE IN DIFFERENT CONNECTED COMPONENTS, MINIMIZING WEIGHT OF REMOVED EDGES

PAUSE AND THINK. WHY IS THIS ALGORITHM A 2-APPROXIMATION?

SUPPOSE F^* IS AN OPTIMAL SET OF EDGES SEPARATING $s_1, s_2, \dots$

- EACH EDGE OF F^* HAS A VERTEX IN AT MOST 2 COMPONENTS C_i
- THE EDGES OF F^* THAT TOUCH C_i IS A CUT BETWEEN s_i AND THE OTHER DISTINGUISHED VERTICES

ALGORITHM:

- FOR $i = 1, \dots, k$: COMPUTE MIN CUT F_i^* SEPARATING s_i FROM $\{s_1, \dots, s_k\} \setminus \{s_i\}$
- OUTPUT $\bigcup_i F_i^*$

K-WAY CUTS VIA LP-ROUNDING

IP FORMULATION:

$x_u^i \in \{0,1\}$ INDICATES IF VERTEX u IS PART OF CONNECTED COMPONENT C_i AROUND s_i.

$x_{s_i}^i = 1$, FOR $i = 1, \ldots, k$

$\sum_i x_u^i = 1$

$z_e^i \geq x_u^i - x_v^i$ FOR $(u,v) \in E$

$z_e^i \geq x_v^i - x_u^i$ FOR $(u,v) \in E$

$\sum_{i=1}^{k} |x_u^i - x_v^i|$ "ℓ_1 DISTANCE"

OBJECTIVE: MINIMIZE $\cancel{\frac{1}{2} \sum_{(u,v) \in E} c_e \|x_u - x_v\|_1}$ $\frac{1}{2} \sum_{(u,v) \in E} c_{(u,v)} \sum_{i=1}^{k} z_{u,v}^i$

DEF. $\boxed{x_u = (x_u^1, \ldots, x_u^k)}$

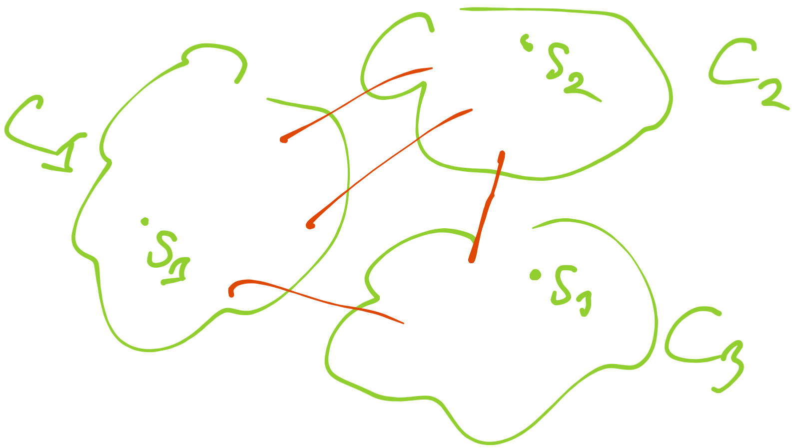
(ASSUME x, z ARE AN OPT LP SOLUTION)

ROUNDING INTUITION:

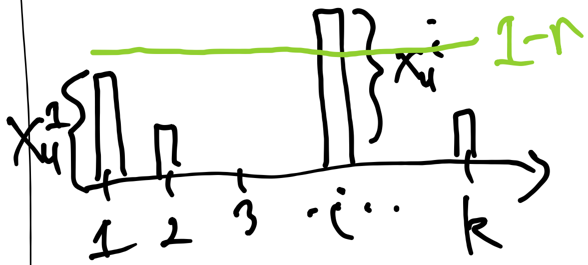
- $x_{s_1}, \ldots, x_{s_k}$ ARE CONSTRAINED TO BE STANDARD BASIS VECTORS
- WANT TO PUT u IN CONNECTED COMPONENT i WHERE x_{s_i} AND x_u ARE "CLOSE".

ROUNDING ALGORITHM

- RANDOMLY SHUFFLE THE ORDER OF $S_1, \dots, S_k$
- PICK THRESHOLD $r \in (0,1)$ AT RANDOM
- FOR $i = 1, \dots, k-1$:
  - PUT UNASSIGNED VERTICES u WITH $x_u^i \geq 1-r$ IN COMPONENT i
- PUT REMAINING UNASSIGNED VERTICES IN COMP. k
- RETURN SET OF EDGES WITH ENDPOINTS IN DIFFERENT COMPONENTS



VECTOR x_u:



$\sum_i x_u^i = 1$ (LP CONSTR.)

ASSIGNED TO C_i FOR THE FIRST i WITH $x_u^i \geq 1-r$.

ANALYSIS OF ROUNDING ALGORITHM

EXPECTED CUT VALUE:

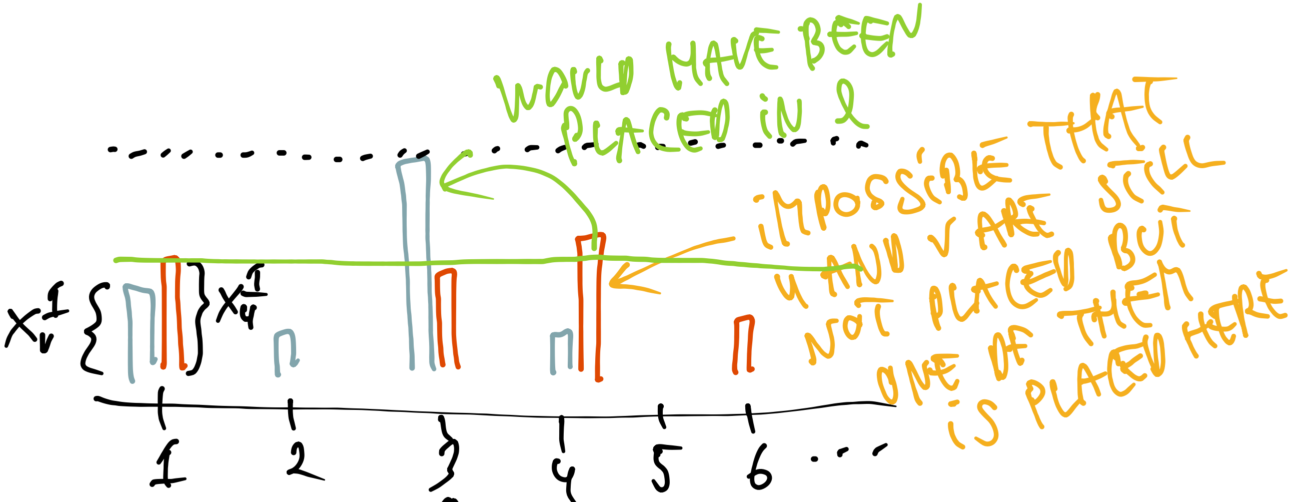
CLAIM

$$\sum_{(u,v)\in E} c_{u,v} \cdot \Pr[(u,v) \text{ in CUT}] \leq \sum_{(u,v)\in E} c_{u,v} \cdot \frac{3}{4} \| x_u - x_v \|_1 = \frac{3}{2} \cdot \left[\frac{1}{2} \sum_{(u,v)\in E} c_e \sum_{i=1}^{k} z^i_{u,v} \right]$$

$$\leq \frac{3}{2} \cdot OPT.$$

ANALYZING $\Pr[(u,v) \text{ in CUT}]$

CONSIDER x_u AND x_v



ℓ: INDEX WITH LARGEST VALUE

INGREDIENTS OF PROOF OF CLAIM:

- NEED ONLY CONSIDER CUTS THAT "HAPPEN" AT i BEFORE OR AT ℓ
- CAN ONLY GET INTO CUT IF $1-r \in (x_u^i, x_v^i)$, WHICH HAS PROB $|x_u^i - x_v^i|$
- IN EXPECTATION, HALF OF i's ARE BEFORE ℓ.

HARDNESS OF APPROXIMATION

USUAL NP-HARDNESS PROOF:

$\boxed{\text{INSTANCE } I \text{ OF NP-HARD PROBLEM } P}$

↓ POLY. TIME REDUCTION

$\boxed{\text{INSTANCE } I' \text{ OF TARGET PROBLEM } P' \text{ SUCH THAT } P(I) = P'(I')}$

MUST ALSO BE NP-HARD

HARDNESS OF APPROXIMATION PROOF:

$\boxed{\text{INSTANCE } I \text{ OF NP-HARD PROBLEM } P}$

↓ POLY. TIME REDUCTION

INSTANCE I' OF TARGET PROBLEM P' SUCH THAT FOR SOME THRESHOLD t, $\alpha > 1$

- IF $P(I)$ = 'yes' THEN $P'(I') \leq t$
- IF $P(I)$ = 'no' THEN $P'(I') \geq \alpha t$

APPROX P' WITHIN α FACTOR IS NP-HARD

"GAP" α

k-CENTER PROBLEM

NP-HARD

START WITH INSTANCE OF DOMINATING SET:

FOR GRAPH (V,E) DOES THERE EXIST $S \subseteq V$ WHERE $|S|=k$ AND EVERY $v \in V \setminus S$ IS ADJACENT TO SOME $w \in S$?

REDUCTION

DEFINE $d(u,v) = \begin{cases} 1 & \text{if } (u,v) \in E \\ 2 & \text{otherwise} \end{cases}, \quad u \neq v$

SATISFIES TRIANGLE INEQUALITY

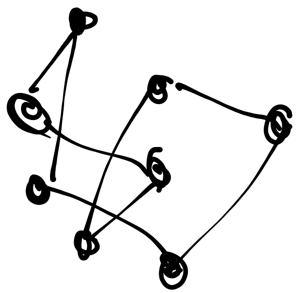
DOMINATING SET OF SIZE $k \iff$ k-CENTER OF RADIUS 1

RADIUS IS 1 OR 2

$\boxed{\text{IF WE COULD 1.99 APPROXIMATE k-CENTER, WE COULD SOLVE (IN P) DOMINATING SET IN POLY TIME}}$

TRAVELING SALESPERSON

- START WITH INSTANCE (V,E) OF HAMILTON PATH.



DOES A CLOSED CYCLE CONNECTING ALL VERTICES EXIST?

NP-HARD

$\Big\downarrow$ REDUCTION

- SET $d(u,v) = \begin{cases} 1 & \text{if } (u,v) \in E \\ 2^n & \text{otherwise} \end{cases}$

n-BIT NUMBER CAN WRITE IN LINEAR TIME

EXISTS A HAMILTON CYCLE $\Longleftrightarrow$ TSP TOUR OF LENGTH n

SHORTEST LENGTH OTHERW. IS $\geq 2^n$

EASY TO INCREASE

$\boxed{\text{APPROXIMATING TSP UP TO FACTOR } 2^n/n \text{ IS NP-HARD}}$

BIN PACKING

NP-HARD

- START WITH INSTANCE OF PARTITION PROBLEM:
  GIVEN $b_1, \dots, b_n \in \mathbb{N}$ CAN WE FIND $S \subseteq \{1, \dots, n\}$ SUCH THAT $\sum_{i \in S} b_i = B/2$, WHERE $B = \sum_i b_i$?

REDUCTION

- LET $a_i = 2b_i/B$ AND CONSIDER $a_1, \dots, a_n$ AS AN INSTANCE OF BIN PACKING.

ANSWER TO PARTITION PROBLEM IS 'yes' $\Longleftrightarrow$ POSSIBLE TO PACK $a_1, \dots, a_n$ INTO 2 BINS

OTHERWISE ≥ 3 BINS

APPROXIMATING BIN PACKING WITHIN FACTOR 1.4999 IS NP-HARD

SOME FAMOUS INAPPROXIMABILITY RESULTS $\boxed{\varepsilon > 0}$

TRIVIAL LOWER BOUND

- MAX E3SAT CANNOT BE APROXIMATED WITHIN $\frac{7}{8}+\varepsilon$ UNLESS P=NP. $\boxed{\text{PROOF VIA PCP THEOREM}}$

WE SAW AN 0.878 APPROX.

- MAX CUT CANNOT BE APPROXIMATED WITHIN $\frac{16}{17}+\varepsilon$ UNLESS P=NP.

BEST POSSIBLE ASSUMING THE "UNIQUE GAMES" PROBLEM IS HARD.

FINE-GRAINED HARDNESS

IDEA: SHOW HARDNESS FOR PROBLEMS IN POLYNOMIAL TIME.

$\boxed{\text{STRONG EXPONENTIAL TIME HYPOTHESIS: SAT REQUIRES TIME } \Omega(1.999^n)\text{, WHERE } n \text{ IS THE NUMBER OF VARIABLES}}$

ORTHOGONAL VECTORS: GIVEN SIZE-n SETS $A, B \subseteq \{0,1\}^{O(\log n)}$ DETERMINE IF THERE EXISTS $x \in A$, $y \in B$ SUCH THAT $x \cdot y = 0$

$x \cdot y = \sum_i x_i y_i$

TRIVIAL ALGORITHM: TIME $O(n^2)$.

$\boxed{\text{ALGORITHM IN TIME } O(n^{1.999}) \Rightarrow \text{SETH IS FALSE}}$

SO ASSUMING SETH, SUCH AN ALGORITHM FOR OVP IS IMPOSSIBLE!

FINE-GRAINED APPROXIMATION HARDNESS

- START WITH INSTANCE A, B OF OVP.
- WANT INSTANCE OF L_∞ CLOSEST PAIR: GIVEN $X, Y \subseteq \mathbb{R}^d$, WHAT IS $\min_{\substack{x \in X \\ y \in Y}} \underbrace{\|x-y\|_\infty}_{\text{LARGEST COORDINATE DIFFERENCE}}$, WHERE $\|z\|_\infty = \max_i |z_i|$?
- <u>REDUCTION</u>: TAKE $X = \{x_a \mid a \in A\}$, WHERE $(x_a)_i = \begin{cases} 1 & \text{if } a_i = 1 \\ 1/3 & \text{if } a_i = 0 \end{cases}$

  $Y = \{y_b \mid b \in B\}$, WHERE $(y_b)_i = \begin{cases} 0 & \text{if } b_i = 1 \\ 2/3 & \text{if } b_i = 0 \end{cases}$

$$\boxed{\|x_a - y_b\|_\infty = \begin{cases} 1 & \text{if } a \cdot b > 0 \\ 1/3 & \text{if } a \cdot b = 0 \end{cases}}$$

SOLVING CLOSEST PAIR WITH APPROX ~~FACTOR~~ 2.99 IN TIME $O(n^{1.999})$ IMPLIES SOLVING OVP IN THE SAME TIME BOUND.