$\boxed{\text{12 Nov 2021}}$ More on Sparsest Cut

Wednesday: Multicommodity flow linear program

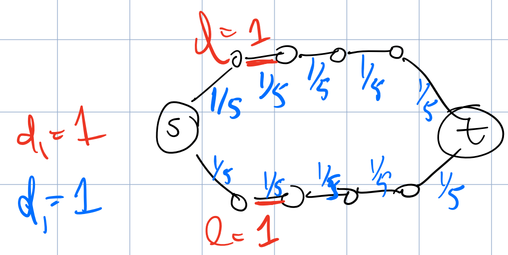
maximize concurrent flow rate
$$\left\{\begin{array}{llll} \max & \rho & & \\ \text{s.t.} & \rho - \sum_{P \in Path(s_i,t_i)} f_P & \leq 0 & \forall i \in [k] \\ & \sum_{P:\,(u,v) \in P} f_P & \leq c(u,v) & \forall (u,v) \in E \\ & f_P & \geq 0 & \forall P \\ & \rho & \geq 0 & \end{array}\right.$$

$\boxed{\text{DUAL}}$

$$\begin{array}{llll} \min & \sum c(u,v) \cdot l_{uv} & & \\ \text{s.t.} & \sum_{(u,v) \in P} l_{uv} - d_i & \geq 0 & \forall i \ \forall P \in Path(s_i,t_i) \\ & \sum_{i=1}^{k} d_i & \geq 1 & \\ & l_{uv},\ d_i & \geq 0 & \end{array}$$

Postprocessing the dual to extract a sparse cut?

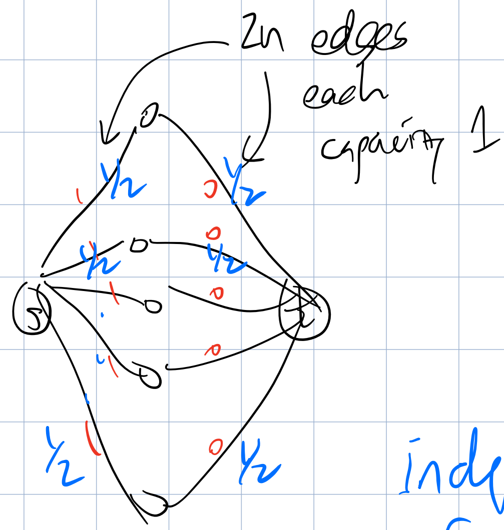
Example $k=1$ (ordinary s-t flow)



Max flow = 2.

optimal dual

another optimal dual



max flow = n.

single-cut opt dual

another opt dual

independent samplings of edges with probability ℓ_{uv} is bad

$\Longrightarrow$ expected cap(A) is exactly the opt value, but expected $|sep(A)|$ can be exponentially small.

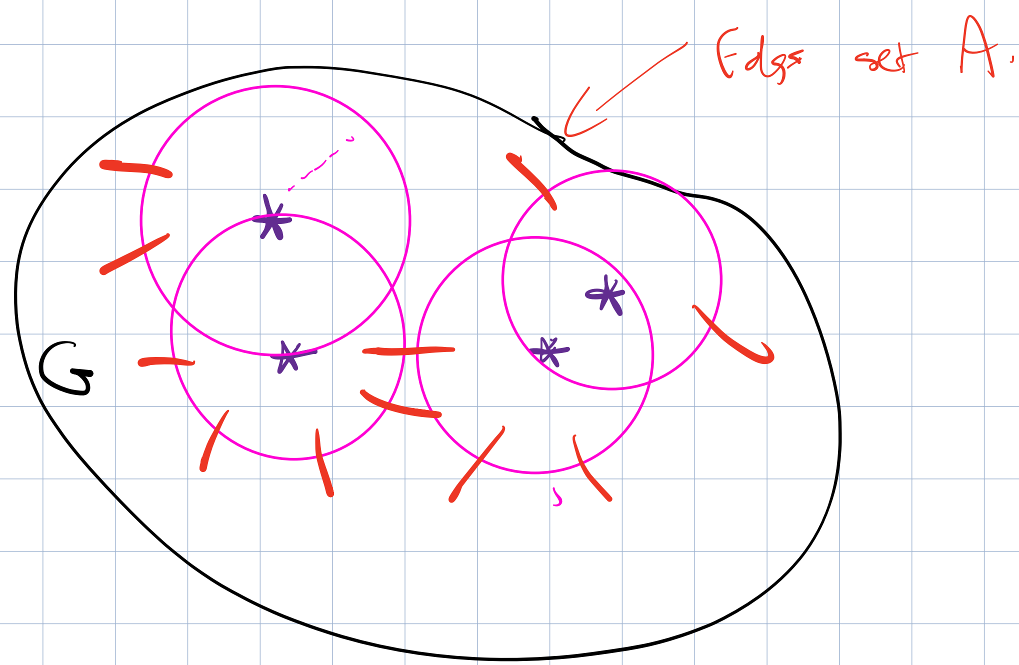
It's too hard to ensure s,t are separated without correlating choices on different paths.

Given a flow network with edge lengths l_{uv}, such that $d(s,t) \geq 1$, there is a very natural way to define a random edge set A s.t. $\Pr(A \text{ separates } s,t) = 1$ and $\Pr((u,v) \in A) \leq l_{uv}$ $\forall (u,v)$.

Choose r uniformly random in $(0,1)$. For edge (u,v) include it in A if and only if
$$d(s,u) \leq r < d(s,v) \text{ or } d(s,v) \leq r < d(s,u).$$

$$\Pr((u,v) \text{ included in } A)$$
$$= \Pr(r \text{ is strictly between } d(s,u), d(s,v))$$
$$= |d(s,u) - d(s,v)|$$
$$\leq l_{uv} \quad \text{(by triangle inequality)}$$

With multicomm flow, we'll do the same thing but using a set of sources, W.



Edge (u,v) gets included in A
if and only if

$$d(W,u) \leq r < d(W,v) \quad \text{or}$$
$$d(W,v) \leq r < d(W,u),$$

where $r \in (0,1)$ unif. random,
and as before,

$$\forall (u,v) \in E \quad \Pr\big((u,v) \in A \big)$$
$$= \big| d(W,u) - d(W,v) \big|$$
$$\leq l_{uv}.$$

$$\mathbb{E}[\mathrm{cap}(A)] = \sum_{(u,v)\in E} c(u,v) \cdot \Pr((u,v) \in A)$$

(*)

$$\leq \sum_{(u,v)\in E} c(u,v)\, \ell_{uv} = \text{value of dual solution.}$$

$$= \text{maximum concurrent flow rate.}$$

If we show $\mathbb{E}[\#\mathrm{sep}(A)] \geq \frac{C}{\log(k)}$ the approx max-flow min-cut theorem follows.

Proof by contra: suppose sparsity of every cut is $> (\text{MCF rate}) \cdot \left(\frac{\log k}{C}\right)$.

Then

$$\mathbb{E}[\mathrm{cap}(A)] = \sum_{A} \mathrm{cap}(A) \cdot \Pr(A)$$

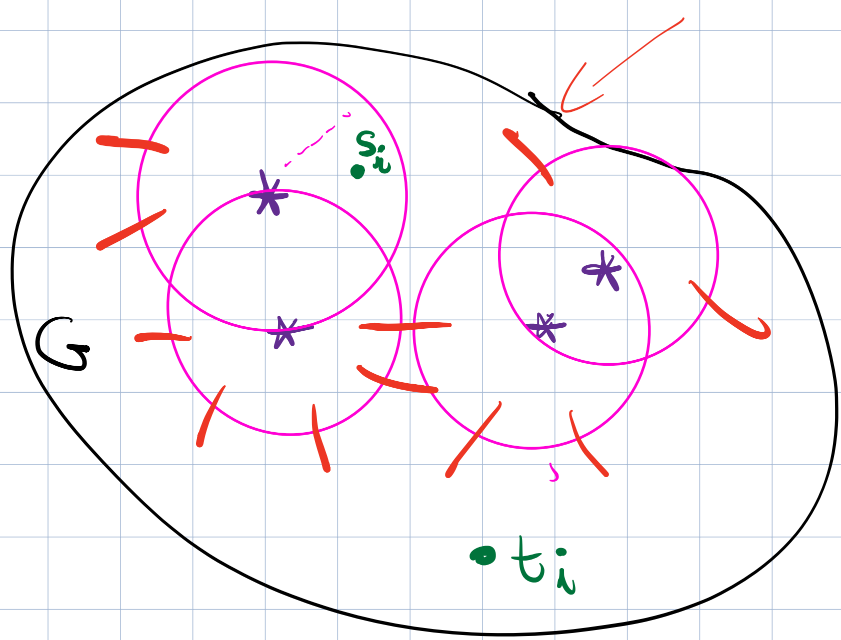
$$= \sum_{A} (\#\mathrm{sep}(A))(\mathrm{sparsity}(A)) \cdot \Pr(A)$$

$$> (\text{MCF rate})\left(\frac{\log k}{C}\right) \sum_{A} (\#\mathrm{sep}(A)) \cdot \Pr(A)$$

$$= (\text{MCF rate})\left(\frac{\log k}{C}\right) \mathbb{E}[\#\mathrm{sep}(A)]$$

$$\geq \text{MCF rate}, \text{ contradicting } (*).$$

$$E[\#\,\mathrm{sep}(A)] = \sum_{i=1}^{k} \Pr(A \text{ separates } s_i, t_i)$$



Obs. $\forall i, r$ A always separates s_i from t_i when W contains a point within distance r from one of $\{s_i, t_i\}$ but not the other.

Let $X = \{s_1, t_1, s_2, t_2, \dots, s_k, t_k\}$,

$S = \{$elements of X within dist. r from $s_i\}$

$T = \{$elements of X in dist r from $t_i\}$

$\forall i, r$

A separates s_i, t_i when W intersects S but not T or vice versa.

<u>Obs 2.</u> $S \cap T = \emptyset$ if $0 < r < \frac{1}{2} d(s_i, t_i)$.

<u>Obs. 3.</u> S, T nonempty.

$s_i \in S$, $t_i \in T$.

Finally: a method for randomly sampling W s.t. $\forall S, T$ non-empty, disjoint,

$$\Pr\left(W \text{ intersects one but not both of } S, T\right) \geq \frac{c}{\log k}.$$

Sample $i \in \{1, 2, \ldots, \lceil \log_2 k \rceil\}$.
Sample every element of X indep'ly with prob 2^{-i}.

If $S \cup T$ has m elements,

$$\Pr\left(i = \lceil \log_2(m) \rceil\right) = \frac{1}{\lceil \log_2 k \rceil}.$$

Then $\frac{1}{2m} < 2^{-i} \le \frac{1}{m}$.

In expectation we sample between $\frac{1}{2}$ and 1 element of $S \cup T$.
With const. probability $|W \cap (S \cup T)| = 1$.

We've shown: for all i,
if $r < \frac{1}{2} d(s_i, t_i)$
then

$$\Pr(A \text{ sep. } s_i, t_i) \geq \frac{c}{\log k}.$$

So

$$\Pr(A \text{ sep. } s_i, t_i) \geq \left(\frac{c}{\log k}\right) \cdot \frac{1}{2} \cdot d(s_i, t_i).$$

$$\mathbb{E}\left[\# \operatorname{sep}(A)\right] \geq \frac{1}{2} \frac{c}{\log k} \sum_{i=1}^{k} d(s_i, t_i).$$

$$\geq \frac{c}{2 \log k}$$ (by dual LP constraints).