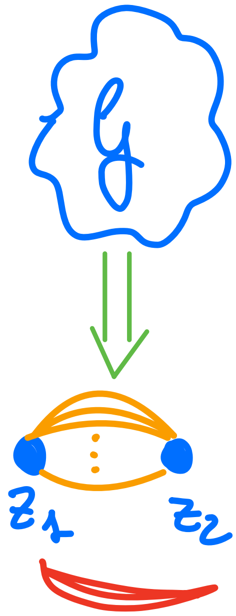
RECAP: KARGER'S MIN-CUT ALGORITHM

FULL_CONTRACTION($\mathcal{G}=(V,E)$)
$\quad \mathcal{G}'=(V',E') \leftarrow \mathcal{G}=(V,E)$
$\quad$ for $i \leftarrow 1$ to $|V|-2$ do
$\qquad e \leftarrow$ RANDOM(E') {account for multiplicities}
$\qquad \mathcal{G}' \leftarrow \mathcal{G}'/e$
$\quad$ return $|E'|$ $\qquad O(n^2)$



KARGER($\mathcal{G}$, s)
$\quad$ min $\leftarrow +\infty$
$\quad$ repeat s times
$\qquad t \leftarrow$ FULL_CONTRACTION($\mathcal{G}$)
$\qquad$ if ($t <$ min) then min $\leftarrow t$
$\quad$ return min

IDEA OF THE ANALYSIS If the random edge selections in FULL-CONTRACT "miss" the edges of C^*, then C^* survives and is returned

CRUX: $|C^*|$ is "small" w.r.t. $|E|$.

Conditional probabilities are important in the analysis of randomized algorithms since they allow to evaluate the probability of a sequence of choices

DEF For $\mathcal{G}(V,\mathcal{E})$, $\forall v \in V$ let $d(v) = \sum_{v \in e} m_e$

$d(v)$ corresponds to the degree of v in a simple graph

PROPERTY 1 $\sum_{v \in V} d(v) = 2|\mathcal{E}|$

Each edge multiplicity is counted twice (if $e = \{u,v\}$: once in $d(u)$, once in $d(v)$)

PROPERTY 2 Let $\boxed{\mathcal{G} = (V,\mathcal{E})}$, and let $\boxed{C^*}$ be a min-cut of $\mathcal{G}$. Then $\forall v \in V$

$$\boxed{|C^*| \leq d(v)}$$

The multiset of edges $\{\{ \{v,x\} \in \mathcal{E} \}\}$ disconnects v from all other nodes.

LEMMA Let $\boxed{\mathcal{G} = (V,\mathcal{E})}$, and let $\boxed{C^*}$ be a min-cut of $\mathcal{G}$. Then

$$\boxed{|\mathcal{E}| \geq |C^*| \cdot |V| / 2}$$

PROOF:

$$2|\mathcal{E}| = \sum_{v \in V} d(v) \quad \text{(PROP 1)}$$

$$\geq \sum_{v \in V} |C^*| \quad \text{(PROP 2)}$$

$$= |C^*||V|$$

$\frac{|C^*|}{|E|} \leq \frac{2}{|V|}$

Thus $|\mathcal{E}| \geq |C^*||V|/2$

ANALYSIS

Given $G = (V, E)$, let us fix a min-cut C^* of G and study the probability that FULL-CONTRACTION (G) does not sample any edge (corresponding to an edge) of C^* in the sequence of $|V|-2$ contractions. (Worst case: there may be many min-cuts!)

Let $\boxed{|V| = n, \ |C^*| = t}$. Define the event

$$\boxed{E_i := \text{"the } i\text{-th contraction avoids } C^*\text{"}}$$

We want to lower bound

$$\boxed{\Pr\left(\bigcap_{i=1}^{n-2} E_i\right)}$$

Consider E_1. By the LEMMA, we have

$$\boxed{|E| \geq t \cdot n/2 \quad (|C^*||V|/2)}$$

Therefore

$$\boxed{\Pr(E_1) = 1 - t/|E| \geq 1 - t/(tn/2) = 1 - 2/n}$$

Assume that E_1 occurs. What is the probability $\Pr(E_2 | E_1)$?

Since E_1 occurs, C^* is still the min-cut in $G_1 = (V_1, E_1) = G / e_1$.

Recall that $\boxed{|V_1| = |V| - 1 = n - 1}$,
thus
$$\boxed{|E_1| \geq t(n-1)/2} \quad \text{(lemma)}$$

Therefore
$$\boxed{\Pr(E_2 \mid E_1) \geq 1 - t/[t(n-1)/2] = 1 - \frac{2}{n-1}}$$

Inductively, at iteration i:
$$\boxed{\Pr(E_i \mid E_1 \cap E_2 \cap \ldots \cap E_{i-1}) \geq 1 - \frac{2}{n-i+1}}$$

(since $|V_{i-1}| = n-i+1 \Rightarrow |E_{i-1}| \geq t(n-i+1)/2$)

Therefore, C^* survives all $n-2$ contractions with probability

$$\Pr\left(\bigcap_{i=1}^{n-2} E_i\right) = \Pr(E_1) \prod_{i=2}^{n-2} \Pr(E_i \mid E_1 \cap \ldots \cap E_{i-1})$$

$$\geq \prod_{i=1}^{n-2} \left(1 - \frac{2}{n-i+1}\right) = \prod_{i=1}^{n-2} \frac{n-i-1}{n-i+1} =$$

$$= \underbrace{\frac{\cancel{n-2}}{n} \cdot \frac{\cancel{n-3}}{n-1}} \cdot \frac{\cancel{n-4}}{\cancel{n-2}} \cdots \overbrace{\frac{2}{\cancel{4}} \cdot \frac{1}{\cancel{3}}} =$$ TELESCOPING PRODUCT

$$= 2/n(n-1) \geq 2/n^2$$

REMARK The probability of correctly returning C^* is rather small. We can amplify it by running $s = d\,\frac{n^2}{2}\ln n$ instances of FULL-CONTRACTION:

$$\Pr(\text{All } s \text{ instances do not return } |C^*|) \le \underbrace{\left(1-\frac{2}{n^2}\right)^{\frac{n^2}{2}}}_{\le e^{-1}}{}^{\cdot d \ln n} < e^{-d\ln n} = \frac{1}{n^d}$$

With $s = O(n^2 \log n)$ executions of FULL-CONTRACTION, we get the min-cut w.h.p.

RUNNING TIME: $O(n^4 \log n)$ SLOWER than the best deterministic algorithm!

We will discuss an optimized approach (based on the same idea) due to Karger and Stein (1993) achieving running time

$$O(n^2 \log^3 n)$$

(much faster than the deterministic algorithm)

OBSERVATION The algorithm also works for weighted graphs (min-weight edge-cut): they can be seen as multigraphs, where $\forall e \in E: m_e = w(e)$

KARGER-STEIN ALGORITHM

Recall that FULL-CONTRACTION executes $|V|-2$ contractions and

$\Pr(\text{FC returns min-cut}) \geq$

$$\frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdots \frac{n-i-1}{n-i+1} \cdots \frac{2}{4} \cdot \frac{1}{3} \geq \frac{2}{n^2}$$

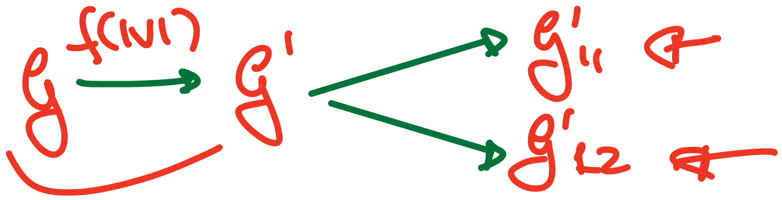
$\Pr(C^*$ "survives" i-th contraction $\mid C^*$ survived previous $i-1)$

To get high probability we need $O(n^2 \log n)$ FULL-CONTRACTIONS!

OBSERVATION: In FULL-CONTRACTION the early contractions avoid C^* with very high probability $\left(1-\frac{2}{n}, 1-\frac{2}{n-1}, \text{etc} \ldots\right)$. However the later contractions "destroy" C^* with larger and larger probability (up to $\ldots \frac{1}{2}, \frac{1}{3}$)

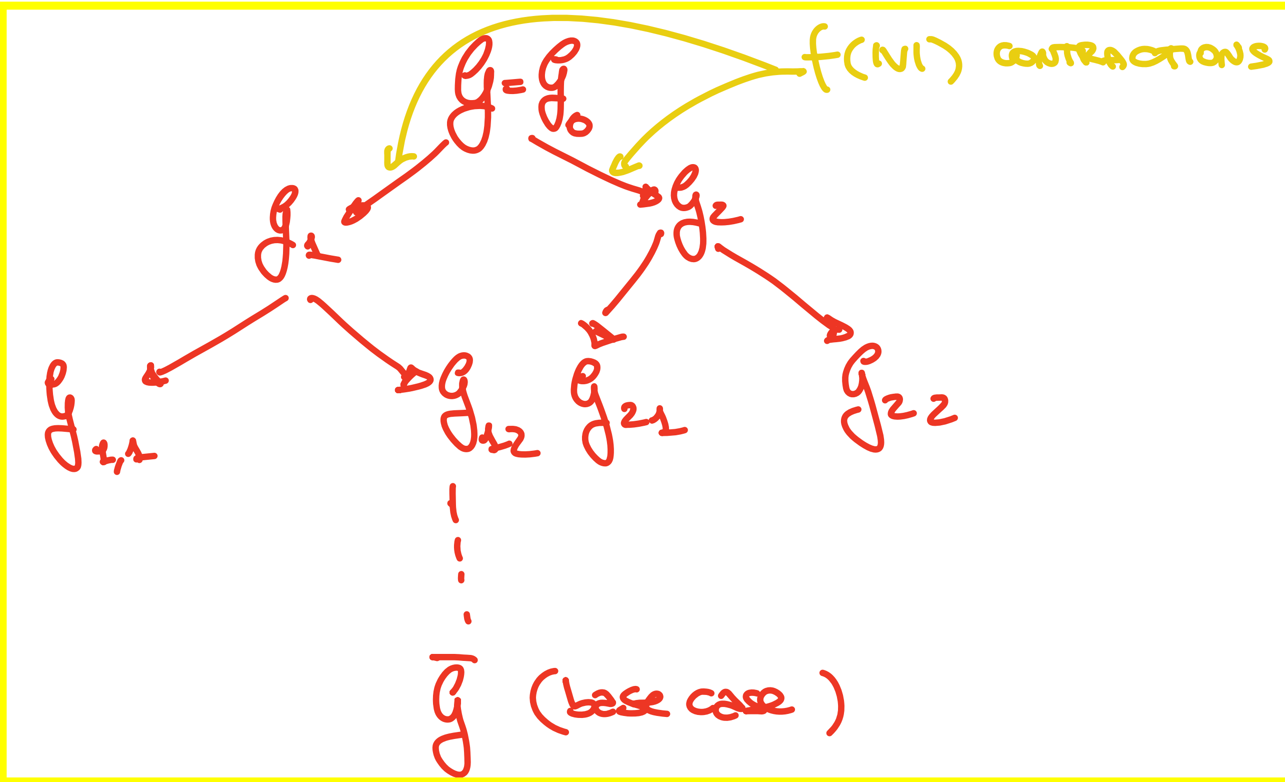
HENCE! FC is "more efficient" during the initial segment of contractions: we want to make better use of them! How: reuse them!

IDEA: "Share" an initial segment of $f(|V|)$ contractions among different instances of FC: "shrink" G to G_1 of a given size and then apply two independent sequences of contractions to G_1 to yield fully contracted G'_{11}, G'_{12}.



$\mathcal{G}'_{11}, \mathcal{G}'_{12}$ share the same "good" sequence of initial contractions $\mathcal{G} \to \mathcal{G}'$

We can implement a recursive strategy based on the above idea



At each level, the multigraphs "share" different segments of contractions

Let $\boxed{\text{PARTIAL\_CONTRACTION}(\mathcal{G}, K)}$ (PC) execute K consecutive random edge contractions on $\mathcal{G}$ ($\to FC(\mathcal{G}) = PC(\mathcal{G}, |V|-2)$)

NOTE $PC(\mathcal{G}, K) \to \mathcal{G}' = (V', E')$ with $|V'| = |V| - K$ nodes and $|E'| \le |E| - K$ edges

The recursive algorithm is the following

R_KS ($\mathcal{G} = (V, \mathcal{E})$)

$\mathcal{B}$ if ($|V| \le 8$) then * solve directly *

$K \leftarrow f(|V|)$ {see analysis}

$\mathcal{G}_1 \leftarrow PC(\mathcal{G}, K)$; $t_1 \leftarrow$ R_KS($\mathcal{G}_1$) {T1}

$\mathcal{G}_2 \leftarrow PC(\mathcal{G}, K)$; $t_2 \leftarrow$ R_KS($\mathcal{G}_2$) {T2}

{two trials, T1 and T2, to get min cut}

return MIN (t_1, t_2)

INDEPENDENT TRIALS T_1, T_2

We generate two subinstances $\mathcal{G}_1, \mathcal{G}_2$ by running two separate partial contractions and then recurse on $\mathcal{G}_1, \mathcal{G}_2$: all instances in $\mathcal{G}_1$'s (resp., $\mathcal{G}_2$'s) subtree will "share" those initial $f(|V|)$ reliable contractions (that avoid edges of C^* with larger probability)

RUNNING TIME

To simplify the analysis, we "upper bound" the cost of $PC(\mathcal{G}, K)$ by $O(n^2)$. Then

$$T_{RKS}(n) = \begin{cases} c & n \le 8 \\ 2T_{RKS}(n-K) + O(n^2) & n > 8 \end{cases}$$

$(= |V|)$

Let us set $K = n - \lceil n/\sqrt{2} + 1 \rceil \simeq n\left(1 - \frac{1}{\sqrt{2}}\right) = \Theta(n)$

Then

$$T_{RKS}(n) = \begin{cases} 2T_{RKS}\left(\left\lceil \frac{n}{\sqrt{2}} + 1 \right\rceil\right) + O(n^2) & n > 8 \\ c & n \leq 8 \end{cases}$$

behaves like the MT recurrence (PROVE IT)

$$T(n) = 2T(n/\sqrt{2}) + O(n^2) \quad \omega(n)$$

$b = \sqrt{2}$ $a = 2$ $\log_b a = \log_{\sqrt{2}} 2 = 2$

$n^{\log_b a} = n^2 = \Theta(\omega(n))$

We are in Case 2: $T_{RKS}(n) = O(n^2 \log n)$

(much less than $O(n^2 \cdot n^2 \cdot \log n)$ of KARGER$(G, d\, \frac{n^2}{2} \log n) \leftarrow$ to get h.p.)

R_KS$(G = (V, E))$; $n \leftarrow |V|$
if $(n \leq 8)$ then * solve directly *
$K \leftarrow n - \lceil n/\sqrt{2} + 1 \rceil$ $\{K \simeq (1 - \frac{1}{\sqrt{2}}) n \simeq n/3.4\}$
$G_1 = (V_1, E_1) \leftarrow PC(G, K)$; $t_1 \leftarrow$ R_KS(G_1) {T1}
$G_2 = (V_2, E_2) \leftarrow PC(G, K)$; $t_2 \leftarrow$ R_KS(G_2) {T2}
{two independent trials, T1 and T2}
$\{|V_1|, |V_2| \simeq n/\sqrt{2} \simeq n/1.41 \ (> n/2)\}$
return MIN(t_1, t_2)

CORRECTNESS ANALYSIS

What is $\boxed{\Pr(\text{RKS}(G) \text{ returns } |\text{min-cut}|)}$?

Consider a single execution of $\text{RC}(G, k)$ with $\boxed{k = n - \lceil n/\sqrt{2} + 1 \rceil}$:

We repeat the argument behind Karger's algorithm and evaluate $\Pr(E_1 \cap E_2 \cap \ldots \cap E_k)$ using conditional probabilities:

Independently:

$$\boxed{\begin{aligned} &\Pr(e^* \text{ "survives" in } G_1) = \\ &= \Pr(e^* \text{ "survives" in } G_2) = \\ &\geq \prod_{i=1}^{n-\lceil n/\sqrt{2}+1\rceil} \frac{n-i-1}{n-i+1} = \frac{\cancel{n-2}}{n} \cdot \frac{\cancel{n-3}}{n-1} \cdot \ldots \cdot \frac{\lceil n/\sqrt{2}+1\rceil}{\lceil n/\cancel{\sqrt{2}+3}\rceil} \frac{\lceil n/\sqrt{2}\rceil}{\lceil n/\cancel{\sqrt{2}+2}\rceil} \\ &= \frac{\lceil n/\sqrt{2}+1\rceil}{n} \frac{\lceil n/\sqrt{2}\rceil}{n-1} \geq \frac{\cancel{n^2}}{2} \frac{1}{\cancel{n}(n-1)} \geq \frac{n}{2(n-1)} \geq \frac{1}{2} \end{aligned}}$$

After the initial contractions, e^* survives independently in G_1 and G_2 with probability $\geq \frac{1}{2}$

Let $p(n) = \Pr(\mathrm{RKS}(G=(V,E)) \text{ correct when } |V| = n)$

Clearly, $p(n) = 1$ for $n \leq 8$ (base case)

For $n > 8$ we have:

$$p(n) = 1 - \Pr(\text{"T1 fails"} \cap \text{"T2 fails"}) =$$

independent events
with identical probability

(By "T1/2 fails" we mean that $t_1, t_2 > |C^*|$)

$$= 1 - \Pr(\text{"T1 fails"}) \cdot \Pr(\text{"T2 fails"})$$

{independent events}

$$= 1 - [\Pr(\text{"T1 fails"})]^2 \quad \{\text{same } \Pr\}$$

$$= 1 - [1 - \Pr(\text{"T1 successful"})]^2$$

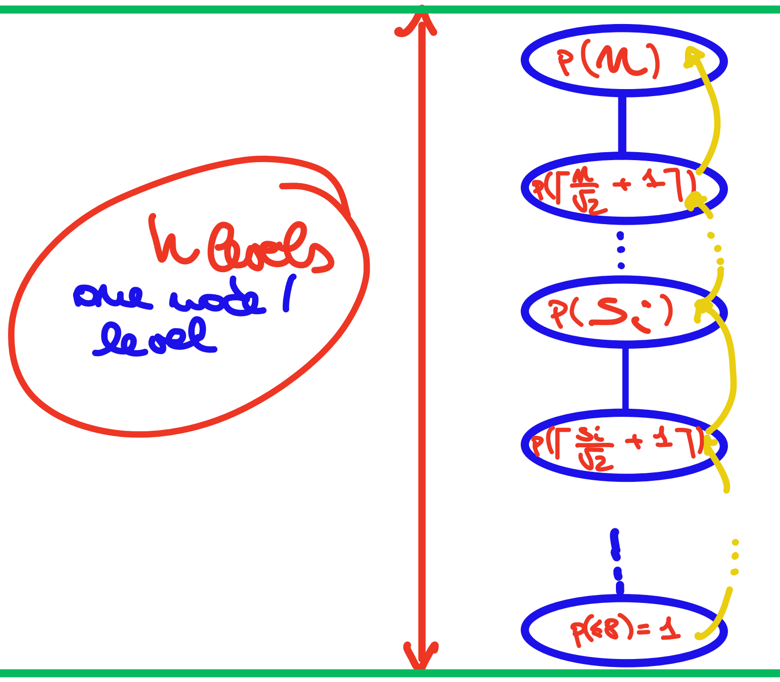
$$= 1 - [1 - \Pr(\text{"}C^* \text{ survives in } G_1\text{"}) \times \Pr(\mathrm{RKS}(G_1) \text{ correct})]^2$$

$$\geq 1 - \left(1 - \frac{1}{2} \cdot p\left(\left\lceil \frac{n}{\sqrt{2}} + 1 \right\rceil\right)\right)^2$$

We have obtained a recurrence on $p(n)$!

$$p(n) \geq \begin{cases} 1 & n \leq 8 \\ 1 - \left(1 - \frac{1}{2} p\left(\lceil n/\sqrt{2} + 1 \rceil\right)\right)^2 & n > 8 \end{cases}$$

We can approach the recurrence using some sort of recursion tree:



$S_0 = n$

$S_1 = \left\lceil \frac{n}{\sqrt{2}} + 1 \right\rceil$

$P(S_i) =$

$= 1 - \left(1 - \frac{1}{2} P(S_{i+1})\right)^2$

$S_{i+1} = \left\lceil \frac{S_i}{\sqrt{2}} + 1 \right\rceil$

How many levels? Let us upperbound S_i, the "size" on i-th level:

LEMMA $\boxed{S_i \le \left\lceil \frac{n}{(\sqrt{2})^i} + 7 \right\rceil}$

PROOF by induction on i. True for $i=0$ $(n \le n+7)$. HP: True for $k = i-1$

TH: $k = i$

$$S_i = \left\lceil \frac{S_{i-1}}{\sqrt{2}} + 1 \right\rceil \overset{HP}{\le} \left\lceil \left(\frac{n}{(\sqrt{2})^{i-1}} + 7 \right) \Big/ \sqrt{2} + 1 \right\rceil$$

$$\le \left\lceil \frac{n}{(\sqrt{2})^i} + \underbrace{\frac{7}{\sqrt{2}} + 1}_{<6} \right\rceil \le \left\lceil \frac{n}{(\sqrt{2})^i} + 7 \right\rceil$$

Note that for $\boxed{i = \lceil 2 \log_2 n \rceil}$ we have

$$\boxed{S_i \le \left\lceil \frac{n}{(\sqrt{2})^{\lceil 2\log_2 n \rceil}} + 7 \right\rceil \le \lceil 1 + 7 \rceil = 8}$$ (Base cases!)

Therefore $\boxed{h \leq \bar{\iota} + 1 = \lceil 2 \log_2 n \rceil + 1}$

We can rewrite the recurrence $p(n)$ as a function of the number of levels h:

$$\boxed{p'(h) \geq \begin{cases} 1 & h = 1 \\ 1 - \left(1 - \frac{1}{2} p'(h-1)\right)^2 \end{cases}}$$

(one level: base case)

Observe that:

$$\boxed{1 - \left(1 - \frac{1}{2} p'(h-1)\right)^2 = p'(h-1) - \frac{1}{4}\left(p'(h-1)\right)^2}$$

We will prove:

LEMMA $\boxed{p'(h) \geq \frac{1}{h}}$