
University of Michigan–Ann Arbor

Department of Electrical Engineering and Computer Science
EECS 498 004 Advanced Graph Algorithms, Fall 2021

Lecture 5: Approximate Max-flow Min-cut Theorem

September 14, 2021

Instructor: Thatchaphol Saranurak

Scribe: Yaowei Long

1. Overview and Preliminaries

We are going to prove a generalization of the max-flow min-cut theorem. Roughly speaking, for two undirected weighted graph G and H , we will show that the H -expansion of G and the maximum H -concurrent flow on G are approximately the same. In details, the gap between them is bounded by a $O(\log n)$ factor, and this factor is tight. Moreover, we will learn two partial orders based on cuts and concurrent flows respectively. The approximate max-flow min-cut theorem implies that these two orders are essentially the same.

The proof of the approximate max-flow min-cut theorem is based on the relationship between the H -expansion of G and its cut-metrics-to-all-metrics relaxation (Leighton-Rao relaxation).

1.1. Definition (H -expansion $\Phi(G, H)$). Let G, H be two undirected weighted graphs on the same vertex set with edge weight functions w_G, w_H . The H -expansion of G is defined as

$$\Phi(G, H) = \min_S \frac{w_G(\partial_G S)}{w_H(\partial_H S)} = \min_{d_S: \text{cut metric over } V} \frac{\sum_{u,v} w_G(u, v) \cdot d_S(u, v)}{\sum_{u,v} w_H(u, v) \cdot d_S(u, v)}.$$

The latter equation follows the property of the cut metric, i.e. $\omega(\partial S) = \sum_{u,v} \omega(u, v) \cdot d_S(u, v)$.

1.2. Definition (LR relaxation). Let G, H be graphs as in definition 1.1, the Leighton-Rao relaxation of $\Phi(G, H)$ is defined as

$$\text{LR}(G, H) = \min_{d: \text{metric over } V} \frac{\sum_{u,v} w_G(u, v) \cdot d(u, v)}{\sum_{u,v} w_H(u, v) \cdot d(u, v)}$$

$\text{LR}(G, H)$ is also equal to the optimal solution of the following LP.

$$\begin{aligned} & (\text{LR}) \\ & \min \sum_{u,v} w_G(u, v) \cdot d(u, v) \end{aligned} \tag{1}$$

$$\begin{aligned}
\text{s.t. } & \sum_{u,v} w_H(u,v) \cdot d(u,v) = 1 \\
& d(u,v) \leq d(u,w) + d(w,v) \quad \forall u,v,w \\
& d(u,v) \geq 0 \quad \forall \{u,v\} \in \binom{V}{2}
\end{aligned}$$

1.3. Theorem. Let G, H be undirected weighted graphs.

$$\text{LR}(G, H) \leq \Phi(G, H) \leq O(\log n) \cdot \text{LR}(G, H)$$

The detailed proof of theorem 1.3 is shown in lecture 2-1.

2. Concurrent Flow

Concurrent flows are also called multi-commodity flows, which will route on an undirected weighted graph different types of commodities each of which has independent source, sink and demands.

The single commodity flow is the special one with just one type of commodity and single source and sink.

2.1. Definition $((s, t)\text{-flows (single commodity flows)})$. Consider a graph $G = (V, E, \kappa)$ with edge capacity $\kappa : E \rightarrow \mathbb{R}_{\geq 0}$. Let s, t be the source and sink and let \mathcal{P}_{st} be the set of all (s, t) -paths in G . An (s, t) -flow f is simply a function that assigns non-negative values to paths in \mathcal{P}_{st} . That is, $f : \mathcal{P}_{st} \rightarrow \mathbb{R}_{\geq 0}$.

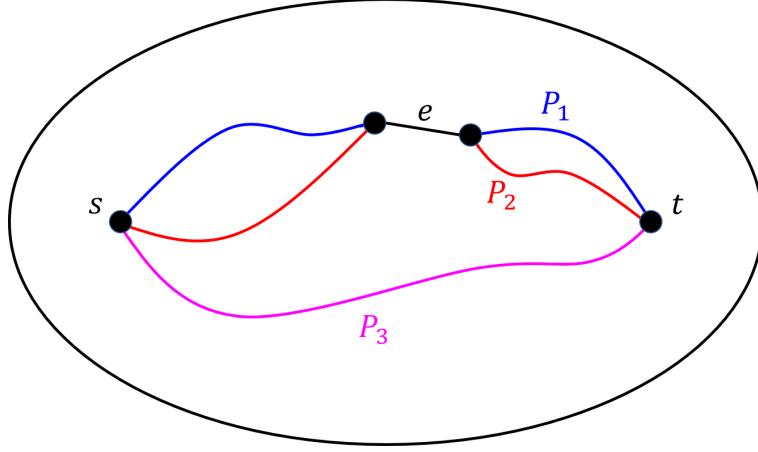
- The **value** of f is $\|f\| = \sum_P f(P)$, i.e. total amount of flow sent through (s, t) -paths.
- The **congestion** of flow f on edge $e = (u, v)$ is the total amount of flow on e relative to the capacity of e :

$$\text{cong}_f(e) = \frac{\sum_{P \ni e} f(P)}{\kappa(e)}.$$

- The **congestion of the flow** is the maximum congestion over all edges:

$$\text{cong}_f = \max_{e \in E} \text{cong}_f(e).$$

See figure 1 for an example.



1. ábra. In this example, let s and t denote the source and the sink respectively. There are three path P_1, P_2, P_3 connecting s and t . Let's say the (s, t) -flow f is defined by $f(P_1) = 10, f(P_2) = 3, f(P_3) = 2$. The value of f is $\|f\| = f(P_1) + f(P_2) + f(P_3) = 15$ and the congestion on edge e is $\text{cong}_f(e) = (f(P_1) + f(P_2))/\kappa(e) = \frac{13}{\kappa(e)}$.

Actually, the definition of single commodity flows is equivalent to the classical definition of single-source and single-sink flows. In the classical definition, an (s, t) -flow $f : E \rightarrow \mathbb{R}_{\geq 0}$ assigns value on **edges** (not paths) such that for every node $u \neq s, t$, the total flow coming in to u equals the total flow going out of u (i.e. $\sum_v f(v, u) = \sum_v f(u, v)$). However, given any (s, t) -flow in the sense of this definition, we can always decompose it into a collection of (s, t) -paths and cycles. Now, by removing all cycles that does not contribute anything to the value of the flow. We obtain precisely the flow in the definition we are working with here.

2.2. Definition (Concurrent flows (multi-commodity flows)). Consider two graphs $G = (V, E_G, \kappa_G)$ and $H = (V, E_H, \kappa_H)$ with nonnegative edge capacities. An **H -concurrent flow \mathcal{F}** is a collection of (s, t) -flows of value $\kappa_H(s, t)$ for all $(s, t) \in E(H)$:

$$\mathcal{F} = \{f_{(s,t)} \mid \|f_{(s,t)}\| = \kappa_H(s, t)\}_{(s,t) \in E(H)}.$$

\mathcal{F} is also called a **flow-embedding** and we say \mathcal{F} **embeds H into G** .

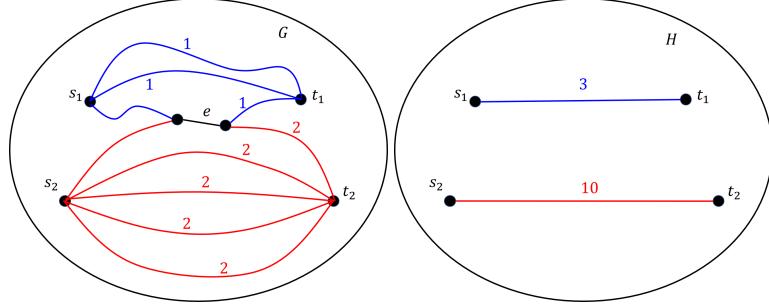
The **congestion of multi-commodity flow \mathcal{F} on e** is

$$\text{cong}_{\mathcal{F}}(e) = \sum_{(s,t) \in E(H)} \text{cong}_{f_{(s,t)}}(e).$$

and the **congestion of \mathcal{F}** is

$$\text{cong}_{\mathcal{F}} = \max_{e \in E} \text{cong}_{\mathcal{F}}(e).$$

If congestion $\text{cong}_{\mathcal{F}} \leq 1$, we say that \mathcal{F} has **no congestion** or \mathcal{F} is **feasible**.



2. ábra. In this example, let \mathcal{F} be the H -concurrent flow on G . We can see that \mathcal{F} embeds blue edge (s_1, t_1) in H into three blue paths in G and embeds red edge (s_2, t_2) in H into five red paths in G . The congestion of edge e in G is $\text{cong}_{\mathcal{F}}(e) = 1/\kappa(e) + 2/\kappa(e) = 3/\kappa(e)$.

Intuitively, H is the **demand graph** of \mathcal{F} because each edge in H represents a unique type of commodity with its own source, sink and demand. Note that concurrent flows are not the same as the multiple-source multiple-sink flows, where the latter essentially has only one type of commodity but multiple sources and sinks.

3. Partial Order based on Concurrent Flow

3.1. Definition (Partial order based on concurrent flow). Consider two undirected weighted graphs G and H .

We define $H \preceq^{\text{flow}} G$ if there is a feasible H -concurrent flow \mathcal{F} in G . In this case, we say that **the demand H is routable in G** or **H is embeddable into G** . Notice that this defines a partial order between graphs.

We also say that **the demand H is routable in G with congestion q** or **H is embeddable into G with congestion q** if \mathcal{F} has congestion q .

3.2. Fact. For any graph G and G' ,

- let $G'' = G + G'$ be such that $\kappa_{G''}(u, v) = \kappa_G(u, v) + \kappa_{G'}(u, v)$ for all u, v .
- let $G' = \alpha \cdot G$ be such that $\kappa_{G'}(u, v) = \alpha \cdot \kappa_G(u, v)$ for all u, v .

From the definition, H is embeddable into G with congestion q iff $H \preceq^{\text{flow}} q \cdot G$.

The partial order also has following properties.

- $G \preceq^{\text{flow}} G$
- If $G_1 \preceq^{\text{flow}} G_2$ and $G_2 \preceq^{\text{flow}} G_3$, then $G_1 \preceq^{\text{flow}} G_3$.
- $G_1 \preceq^{\text{flow}} G_2$ iff $G_1 + H \preceq^{\text{cut}} G_2 + H$.
- $G_1 \preceq^{\text{flow}} \alpha \cdot G_2$ iff $\frac{1}{\alpha} G_1 \preceq^{\text{flow}} G_2$.

4. The Maximum Concurrent Flow Problem

4.1. Definition (Maximum concurrent flow problem). Given a graph G and a demand graph H , we need to compute a feasible $\tau \cdot H$ -concurrent flow in G with maximum τ . We denote the optimal value τ^* by $\text{mcf}(G, H) = \max\{\tau \mid \tau \cdot H \leqslant^{\text{flow}} G\}$. In particular, we have $H \leqslant^{\text{flow}} G \iff \text{mcf}(G, H) \geq 1$.

Note that this problem is a generalization of the single-source single-sink maximum flow problem. Let H_{st} be the graph containing only one edge (s, t) with capacity 1. $\text{mcf}(G, H)$ is exactly the value of maximum (s, t) -flow and $\Phi(G, H)$ is the value of minimum (s, t) -cut. The classical max-flow min-cut theorem states the following.

4.2. Theorem (Max-flow min-cut theorem). *For any graph G where $s, t \in V(G)$, we have $\text{mcf}(G, H_{st}) = \Phi(G, H_{st})$.*

5. An Approximate Max-flow Min-cut theorem

5.1. Theorem (Approximate Maxflow Mincut Theorem). *Given undirected graphs G and H ,*

$$\text{mcf}(G, H) \leq \Phi(G, H) \leq O(\log n) \cdot \text{mcf}(G, H).$$

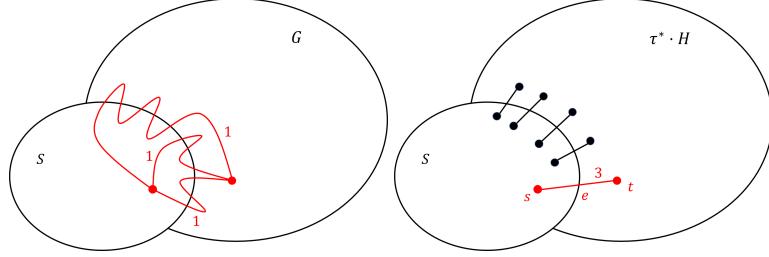
proof of one direction of theorem 5.1, $\text{mcf}(G, H) \leq \Phi(G, H)$.

- Let $\tau^* = \text{mcf}(G, H)$ be the optimal value of the concurrent flow. Let \mathcal{F} be the multi-commodity flow that embeds $\tau^* H$ to G .
- Consider any cut $S \subset V$.
- For each edge $(s, t) \in \partial_H S$, \mathcal{F} sends at least $\tau^* \cdot \kappa_H(s, t)$ unit of flow crossing the cut S .
- So the total unit of flow crossing the cut S is $\tau^* \cdot \kappa_H(\partial_H S)$.
- But the total capacity of the cut is $\kappa_G(\partial_G S)$. So

$$\tau^* \cdot \kappa_H(\partial_H S) \leq \kappa_G(\partial_G S).$$

Formally, let $\mathcal{F} := \{f_{s,t}\}_{(s,t) \in E(H)}$ be the feasible flow that embeds $\tau^* \cdot H$ into G . $\text{cong}_{\mathcal{F}} \leq 1$ implies that $\forall e \in E(G), \sum_{(s,t) \in H} \sum_{P \ni e} f_{s,t}(P) \leq \kappa_G(e)$. Since $LHS = \sum_{(s,t) \in \partial_H S} \sum_{P \ni e} f_{s,t}(P) \leq \sum_{(s,t) \in H} \sum_{P \ni e} f_{s,t}(P) \leq \kappa_G(e) = RHS$, summing over all edges in $\partial_G S$ for both sides, we have $LHS \geq \sum_{(s,t) \in \partial_H S} \sum_{P \cap \partial_G S \neq \emptyset} f_{s,t}(P) = \sum_{(s,t) \in \partial_H S} \sum_{P \in \mathcal{P}_{s,t}} f_{s,t}(P) = \sum_{(s,t) \in \partial_H S} \tau^* \cdot \kappa_H(s, t) = \tau^* \cdot \kappa_H(\partial_H S); RHS = \kappa_G(\partial_G S)$. $\implies \tau^* \cdot \kappa_H(\partial_H S) \leq \kappa_G(\partial_G S)$.

See figure 3 for an intuitive proof.



3. ábra. Let the left figure represent G and the right figure represent $\tau^* \cdot H$. The capacity of the cut S in $\tau^* \cdot H$ is $\tau^* \cdot \kappa_H(\partial_H S)$, represented by five edges in the right figure. For each edge in $\tau^* \cdot H$ crossing the cut, it is embedded into several paths in G and each path will cross the cut in G at least once. Take the red edge e with $\kappa_H(e) = 3$ as an example. It is embedded into three paths each of which has flow 1, which means that at least 3 units of capacity of the cut in G are owned by e . It turns out that $\tau^* \cdot \kappa_H(\partial_H S) \leq \kappa_G(\partial_G S)$.

- Therefore, we have

$$\text{mcf}(G, H) = \tau^* \leq \min_S \frac{\kappa_G(\partial_G S)}{\kappa_H(\partial_H S)} = \Phi(G, H)$$

□

We prove another direction of theorem 5.1 (i.e. $\Phi(G, H) \leq O(\log n) \cdot \text{mcf}(G, H)$) via LP duality. In fact, we will show that the optimal value of concurrent flow $\text{mcf}(G, H)$ is exactly the same as the optimal value from the Leighton-Rao relaxation $\text{LR}(G, H)$ of $\Phi(G, H)$ (see LP 1), as lemma 5.2 says. By lemma 5.2 and theorem 1.3, we immediately get

$$\Phi(G, H) \leq O(\log n) \cdot \text{LR}(G, H) = (\log n) \cdot \text{mcf}(G, H).$$

5.2. Lemma (Key lemma). $\text{mcf}(G, H) = \text{LR}(G, H)$.

proof of one direction of theorem 5.1. $\Phi(G, H) \leq O(\log n) \cdot \text{mcf}(G, H)$. Because $\text{mcf}(G, H) = \text{LR}(G, H)$ by lemma 5.2 and $\Phi(G, H) \leq O(\log n) \cdot \text{LR}(G, H)$ by theorem 1.3, we immediately get $\Phi(G, H) \leq O(\log n) \cdot \text{mcf}(G, H)$.

□

6. Proof of Lemma 5.2

We start by modeling the maximum concurrent flow problem as an LP.

- The problem is fully specified given **parameters** $\{C_{uv}\}_{(u,v) \in \binom{V}{2}}$ and $\{D_{uv}\}_{(u,v) \in \binom{V}{2}}$ defined as follows.
 - Let $C_{uv} = \kappa_G(u, v)$ be the capacity of edge (u, v) in G . If (u, v) is not in G , then $C_{uv} = 0$.
 - Let $D_{uv} = \kappa_H(u, v)$ be the demand of from u to v defined by H . If (u, v) is not in H , then $D_{uv} = 0$.

- Recall definition 4.1 of the maximum concurrent flow problem and definition 2.2 of concurrent flows. The **variables** of this LP are the scaling factor τ and flow assignments $f(P)$ to all (u, v) -paths for all u, v .

$$\begin{aligned}
& (\text{Primal } \mathbf{P}_{LP}) \\
\max \quad & \tau \\
\text{s.t.} \quad & D_{uv} \cdot \tau - \sum_{P \in \mathcal{P}_{u,v}} f(P) \leq 0 \quad \forall u, v \\
& \sum_{P \ni (u,v)} f(P) \leq C_{uv} \quad \forall u, v \\
& f(P) \geq 0 \quad \forall P \in \mathcal{P}_{uv} \text{ over all } u, v
\end{aligned}$$

We let $y(u, v)$ be variables in the dual corresponding to the first type of constraints in the primal LP and $x(u, v)$ be variables corresponding to the second type of constraints in the primal LP. We immediately get the dual LP as follows.

$$\begin{aligned}
& (\text{Dual } \mathbf{D}_{LP}) \\
\min \quad & \sum_{u,v} C_{uv} \cdot x(u, v) \\
\text{s.t.} \quad & \sum_{u,v} D_{uv} \cdot y(u, v) \geq 1 \\
& -y(u, v) + \sum_{(u',v') \in P} x(u', v') \geq 0 \quad \forall P \in \mathcal{P}_{uv} \text{ over all } u, v \\
& x(u, v), y(u, v) \geq 0 \quad \forall u, v
\end{aligned}$$

6.1. *Remark.* Deriving dual LPs is a very important skill for researching in TCS. To interpret the problem that \mathbf{D}_{LP} is solving, we can treat $x(u, v)$ as "the length of edge uv ", C_{uv} as "the capacity or width of the edge uv ", and $y(u, v)$ as the "distance between u and v " since it plays the role of a lower bound of the length of paths between u and v . The objective function is trying to minimize the "total volume" of the graph, yet still guarantee the distances between vertices are big enough.

By definition, $\text{mcf}(G, H)$ is exactly the optimal value of \mathbf{P}_{LP} , and by strong duality, we have $\text{OPT}(\mathbf{P}_{LP}) = \text{OPT}(\mathbf{D}_{LP})$, which means $\text{mcf}(G, H) = \text{OPT}(\mathbf{D}_{LP})$. In what follows, we will show that $\text{OPT}(\mathbf{D}_{LP}) = \text{OPT}(\text{LR})$, where LR is LP 1 in definition 1.2 (see below for an equivalent version).

$$\begin{aligned}
& (\text{LR}) \\
\min \quad & \sum_{u,v} C_{uv} \cdot d(u, v) \\
\text{s.t.} \quad & \sum_{u,v} D_{uv} \cdot d(u, v) \geq 1 \\
& d(u, v) \leq d(u, w) + d(w, v) \quad \forall u, v, w
\end{aligned}$$

$$d(u, v) \geq 0 \quad \forall u, v$$

By lemma 6.3 and definition 1.2, we immediately have

$$\text{mcf}(G, H) = \text{OPT}(\mathcal{D}_{LP}) = \text{OPT}(\text{LR}) = \text{LR}(G, H).$$

6.2. Fact. Let $\text{dist}_x(u, v)$ be the distance metric defined by x . We have $y(u, v) \leq \text{dist}_x(u, v) \leq x(u, v)$ for all u, v

6.3. Lemma. $\text{OPT}(\text{LR}) = \text{OPT}(\mathcal{D}_{LP})$.

Proof of $\text{OPT}(\text{LR}) \leq \text{OPT}(\mathcal{D}_{LP})$.

Let (x, y) be an optimal solution for \mathcal{D}_{LP} . Let $d(u, v) = \text{dist}_x(u, v)$ for all u, v . We have $d(u, v)$ also a feasible solution for LR because

- $\sum_{u,v} D_{uv} \cdot d(u, v) \geq \sum_{u,v} D_{uv} \cdot y(u, v) \geq 1$ and
- the triangle inequality is trivially satisfied as d is a metric.

Furthermore,

$$\text{OPT}(\text{LR}) \leq \sum_{u,v} C_{uv} \cdot d(u, v) \leq \sum_{u,v} C_{uv} \cdot x(u, v) = \text{OPT}(\mathcal{D}_{LP})$$

□

Proof of $\text{OPT}(\text{LR}) \geq \text{OPT}(\mathcal{D}_{LP})$.

For convenience, we write $d(P) = \sum_{(u',v') \in P} d(u', v')$, and the constraint in \mathcal{D}_{LP} can be rewritten as $x(P) \geq y(u, v) \forall P \in \mathcal{P}_{uv}$.

Obviously, for any function $d : V \times V \rightarrow \mathbb{R}_{\geq 0}$, if $d(u, v) \leq d(u, w) + d(w, v)$ for all u, v, w (i.e. d satisfies the triangle inequality), then for all u, v and $P \in \mathcal{P}_{uv}$, we have $d(u, v) \leq d(P)$.

Let d be an optimal solution for LR. Let $x(u, v) = y(u, v) = d(u, v)$ for all u, v . We have

- $\sum_{u,v} D_{uv} \cdot y(u, v) = \sum_{u,v} D_{uv} \cdot d(u, v) \geq 1$ and
- $y(u, v) = d(u, v) \leq d(P) = x(P)$ for all u, v and $P \in \mathcal{P}_{uv}$.

Furthermore,

$$\text{OPT}(\mathcal{D}_{LP}) \leq \sum_{u,v} C_{uv} \cdot x(u, v) = \sum_{u,v} C_{uv} \cdot d(u, v) = \text{OPT}(\text{LR})$$

□

7. The Flow-Cut Gap

Actually, the factor $O(\log n)$ in the approximate max-flow min-cut theorem is tight. This is called the **flow-cut gap**.

7.1. Theorem. There exist G and H where $\text{mcf}(G, H) \leq \frac{1}{O(\log n)} \cdot \Phi(G, H)$.

Sketch of proof. The proof is from problem 5 in homework 1.

Let K be a complete graph with n nodes where each edge has weight $1/n$. Let G be an $\Omega(1)$ -expander with constant maximum degree. We have $\Phi(G, K) \geq \Omega(1)$.

Because G is a graph with bounded degree, there will be at least $\Omega(n^2)$ pairs (u, v) of nodes where $\text{dist}_G(u, v) \geq \Omega(\log n)$. Let \mathcal{F} be a concurrent flow that embeds K into G and define the volume of \mathcal{F} as $\sum_{P \in \mathcal{F}} f(P) \cdot |P|$, where $f(P)$ is the flow value along the path P and $|P|$ is the number of edges in the path. The volume of \mathcal{F} will be at least $\Omega(n \log n)$. Therefore, $\text{mcf}(G, K) \leq O(1/\log n)$.

□

8. Partial Order based on Cut

In section 3, we define partial order based on concurrent flow. In this section, we define another partial order based on cut. We will see that these two partial orders are essentially equivalent.

8.1. Definition (Partial order based on cut). Consider two undirected weighted graphs $G = (V, E_G, \kappa_G)$ and $H = (V, E_H, \kappa_H)$.

We define $H \leqslant^{\text{cut}} G$ if for all $S \subset V$, we have $\kappa_H(\partial_H S) \leq \kappa_G(\partial_G S)$. In this case, we say that H is **cut-wise at most** G .

8.2. Fact. $H \leqslant^{\text{cut}} G \iff \Phi(G, H) \geq 1$.

For any graph G and G'

- let $G'' = G + G'$ be such that $\kappa_{G''}(u, v) = \kappa_G(u, v) + \kappa_{G'}(u, v)$ for all u, v and
- let $G' = \alpha \cdot G$ be such that $\kappa_{G'}(u, v) = \alpha \cdot \kappa_G(u, v)$.

Similar to the partial order based on concurrent flow, the partial order based on cut also has following properties.

- If $G_1 \leqslant^{\text{cut}} G_2$ and $G_2 \leqslant^{\text{cut}} G_3$, then $G_1 \leqslant^{\text{cut}} G_3$.
- $G_1 \leqslant^{\text{cut}} G_2$ iff $G_1 + H \leqslant^{\text{cut}} G_2 + H$.
- $G_1 \leqslant^{\text{cut}} \alpha \cdot G_2$ iff $\frac{1}{\alpha} G_1 \leqslant^{\text{cut}} G_2$.

We have seen two ways to compare graphs \leqslant^{cut} and \leqslant^{flow} . The two relations are basically the same because of the approximate max-flow min-cut theorem. Roughly speaking, for two graphs G and H ,

- If there exists a sparse cut in G w.r.t. H , then we cannot route the H -demand.
- If there is no sparse cut in G w.r.t. H , then we can route the H -demand with small congestion.

See lemma 8.3 for a formal description.

8.3. Lemma. Consider two weighted undirected graphs G and H .

- If $H \leqslant^{\text{flow}} G$, then $H \leqslant^{\text{cut}} G$.

- If $H \leqslant^{\text{cut}} G$, then $H \leqslant^{\text{flow}} O(\log n) \cdot G$.

Proof.

- Recall the approximate max-flow min-cut theorem.

$$\text{mcf}(G, H) \leq \Phi(G, H) \leq O(\log n) \cdot \text{mcf}(G, H).$$

- Recall that $H \leqslant^{\text{flow}} G \iff \text{mcf}(G, H) \geq 1$ and $H \leqslant^{\text{cut}} G \iff \Phi(G, H) \geq 1$.
- If $H \leqslant^{\text{flow}} G$, then $\Phi(G, H) \geq \text{mcf}(G, H) \geq 1$ and so $H \leqslant^{\text{cut}} G$.
- If $H \leqslant^{\text{cut}} G$, then $\Phi(G, H) \geq 1$. So $\text{mcf}(G, H) \geq \frac{1}{O(\log n)}$. So $H \leqslant^{\text{flow}} O(\log n) \cdot G$.

□