Flow decompositions and directed graph minors

Andreas Grigorjew

29 August, 2024

Joint work with Manuel Cáceres, Massimo Cairo, Wanchote Jiamjitrak, Shahbaz Khan, Brendan Mumey, Romeo Rizzi, Alexandru I. Tomescu and Lucia Williams

UNIVERSITY OF HELSINKI

Flow graphs

Given is an s-t (multigraph) DAG $G = (V, E)$ and a flow $f : E \to \mathbb{N}$.
Conservation of flow: $\sum_{e \in \delta^+(v)} f(e) = \sum_{e \in \delta^-(v)} f(e) \quad \forall v \in V \setminus \{s, t\}$.

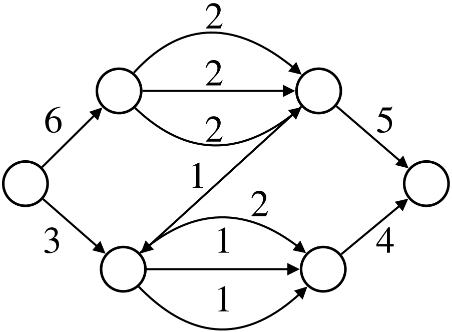


Figure: Simple flow graph.

Minimum Flow Decomposition (MFD)

Minimum Flow Decomposition (MFD) of (G, f): **minimum sized** set of s-t **paths and weights** $\{(P_1, w_1), \ldots, (P_k, w_k)\}$ $(w_i \in \mathbb{N})$ with

$$f = \sum_{i=1}^{k} w_i P_i.$$

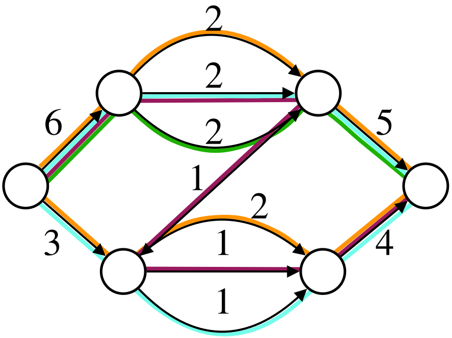


Figure: Flow decomposition in 6 s-t paths (weights omitted).

Solving MFD, state of the art

- MFD is NP-hard, even if the flow values come from $\{1, 2, 4\}$ [Hartman et al., 2012]
- MFD is APX-hard [Hartman et al., 2012]
- FPT in time $O(2^{k^2} \cdot (n + \log \|f\|))$ [Kloster et al., 2018] and ILP [Dias et al., 2022] solvers exist
- Greedy approximation factor $\Omega(m/\log m)$ [Cáceres et al., 2024]
- Greedy-weight commonly used in applications [Baaijens et al., 2020, Tomescu et al., 2013]

Outline

- Explore DAG structure with (minimal) flows
- $O(\log \|f\|)$-approximation for graphs with forbidden minors
- Implication: Quasi-polynomial algorithm for some class of instances
- Bridges gap between application and theory

Graph structure impact: digraph width

Definition

The *width* of an s-t DAG G is the **minimum sized set of s-t paths** to cover all edges.

Graph structure impact: digraph width

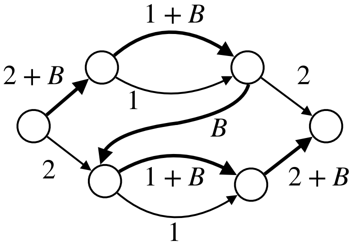
Definition

The *width* of an s-t DAG G is the **minimum sized set of** s-t **paths** to cover all edges.

We have $\text{width}(G) \leq \text{mfd}(G, f)$ for all $f > 0$.

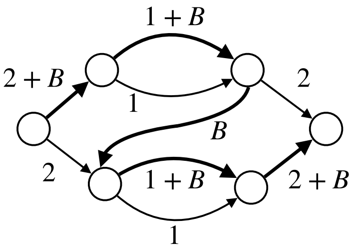
Width hinders greedy

Width hinders greedy

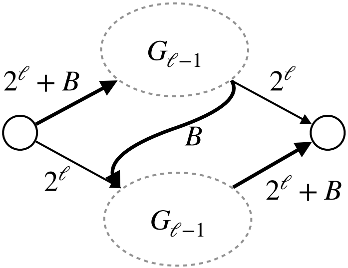


(a) The base case $(G_1, X_{1,B})$. Bold edges carry flow at least B.

Width hinders greedy

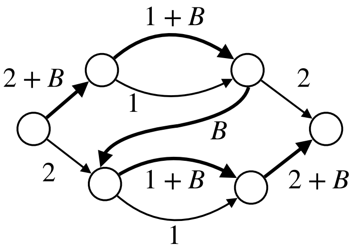


(a) The base case $(G_1, X_{1,B})$. Bold edges carry flow at least B.

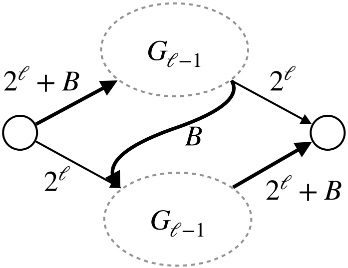


(b) Building $(G_\ell, X_{\ell,B})$ from two copies of $(G_{\ell-1}, X_{\ell-1,B})$ $(\ell > 1)$.

Width hinders greedy



(a) The base case $(G_1, X_{1,B})$. Bold edges carry flow at least B.



(b) Building $(G_\ell, X_{\ell,B})$ from two copies of $(G_{\ell-1}, X_{\ell-1,B})$ $(\ell > 1)$.

$(G_\ell, X_{\ell,B})$ can be decomposed into $\Theta(\ell)$ paths. Greedy-weight uses $\Theta(2^\ell)$ paths.

$\rightarrow$ Approximation ratio for greedy-weight on MFD is $\Omega(m/\log m)$.

Flow derived graph minors

Definition

H is a *butterfly contraction* of H' if H is obtained from contracting edges (u, v) in H' where $\deg^+(u) = 1$ or $\deg^-(v) = 1$.

Flow derived graph minors

Definition

H is a *butterfly contraction* of H' if H is obtained from contracting edges (u, v) in H' where $\deg^+(u) = 1$ or $\deg^-(v) = 1$.

Definition [Grigorjew et al., 2024]

We call H a *flow-minor* of an *s*-*t* DAG G, if there exists H' such that

- $E(H') = \text{supp}(f)$ for some flow $f : E(G) \to \mathbb{N}$ (write $H' = H_f$),
- H is a butterfly contraction of H'

Flow derived graph minors

Definition

H is a *butterfly contraction* of H' if H is obtained from contracting edges (u, v) in H' where $\deg^+(u) = 1$ or $\deg^-(v) = 1$.

Definition [Grigorjew et al., 2024]

We call H a *flow-minor* of an s-t DAG G, if there exists H' such that

- $E(H') = \text{supp}(f)$ for some flow $f : E(G) \to \mathbb{N}$ (write $H' = H_f$),
- H is a butterfly contraction of H'

From [Deligkas and Meir, 2017]: Delete edges (u, v), where $\deg^-(u) \geq 2$ and $\deg^+(v) \geq 2$, contract via butterfly contractions.

Flow derived graph minors

Definition [Deligkas and Meir, 2017]

The *parallel width* pw of an s-t DAG is the **largest minimal** s-t **cut-set**.

- Equivalently, the maximum width(G_f) throughout all $f \geq 0$.
- Equivalently, the maximum possible value of a minimal flow.

Flow derived graph minors

Definition [Deligkas and Meir, 2017]

The *parallel width* pw of an s-t DAG is the **largest minimal s-t cut-set**.

- Equivalently, the maximum width(G_f) throughout all $f \geq 0$.
- Equivalently, the maximum possible value of a minimal flow.

Let $GP(c)$ be the s-t DAG of c parallel edges (s, t).

Lemma [Deligkas and Meir, 2017]

$\mathrm{pw}(G) < c$ iff G is $GP(c)$-f-minor free.

Ch_k DAG

We call the following DAGs Ch_k:

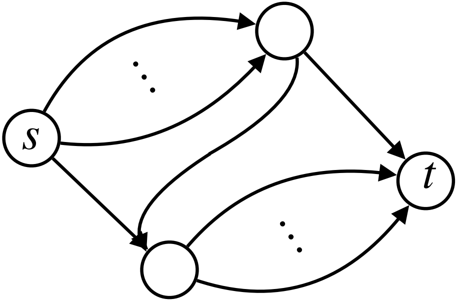


Figure: On each side, k parallel edges.

Lemma [Deligkas and Meir, 2017]

A DAG is series-parallel if and only if it is Ch_1-f-minor free.

Width-stable graphs

Definition ([Cáceres et al., 2024])

G is called *width-stable* if $\text{width}(G_f) \leq \text{width}(G_g)$ for all flows $f \leq g$ on G.

The following are equivalent [Cáceres et al., 2024]:

- G is width-stable,
- G is Ch_2-f-minor free.

$\rightarrow$ Greedy is a $O(\log \text{Val}(f))$-approximation on width-stable graphs [Cáceres et al., 2024].

Approximating MFD

Approximating MFD

Power of two–approach

1. "Remove" the odd part: Given flow f, define flow g of the **same parity** as f,
2. Choose a small g and decompose it to $\mathrm{Val}(g)$ paths,
3. Divide the remaining flow by 2.

Approximating MFD

Power of two–approach

1. "Remove" the odd part: Given flow f, define flow g of the **same parity** as f,
2. Choose a small g and decompose it to $\text{Val}(g)$ paths,
3. Divide the remaining flow by 2.

$\rightarrow$ We obtain a decomposition of size $\text{Val}(g_1) + \text{Val}(g_2) + \cdots + \text{Val}(g_{\log \|f\|})$.

How to choose g?

How to choose a small, parity fixing flow g?

$$\begin{aligned} &\min \mathrm{Val}(g), \text{s.t.} \\ &g \text{ is a flow on } G, \\ &0 \leq g \leq f, \\ &f - g \equiv_2 0. \end{aligned} \tag{1}$$

How to choose g?

How to choose a small, parity fixing flow g?

$$\begin{aligned} &\min \operatorname{Val}(g), \text{s.t.} \\ &g \text{ is a flow on } G, \\ &0 \leq g \leq f, \\ &f - g \equiv_2 0. \end{aligned} \tag{1}$$

Easier problem:

$$\begin{aligned} &\min \operatorname{Val}(h), \text{s.t.} \\ &h \text{ is a flow on } G, \\ &0 \leq h \leq f, \\ &0 < h(e) \quad \forall e \in E(G)\colon\ f(e) \text{ is odd.} \end{aligned} \tag{2}$$

How to choose g?

Lemma [Cáceres et al., 2024]

For every flow $f : E \to \mathbb{N}$ we can find a flow $\text{Unitary}(f) : E \to \{-1, 0, 1\}$ such that $f \equiv_2 \text{Unitary}(f)$, in time $O(m)$.

How to choose g?

Lemma [Cáceres et al., 2024]

For every flow $f : E \to \mathbb{N}$ we can find a flow $\text{Unitary}(f) : E \to \{-1, 0, 1\}$ such that $f \equiv_2 \text{Unitary}(f)$, in time $O(m)$.

Lemma [Grigorjew et al., 2024]

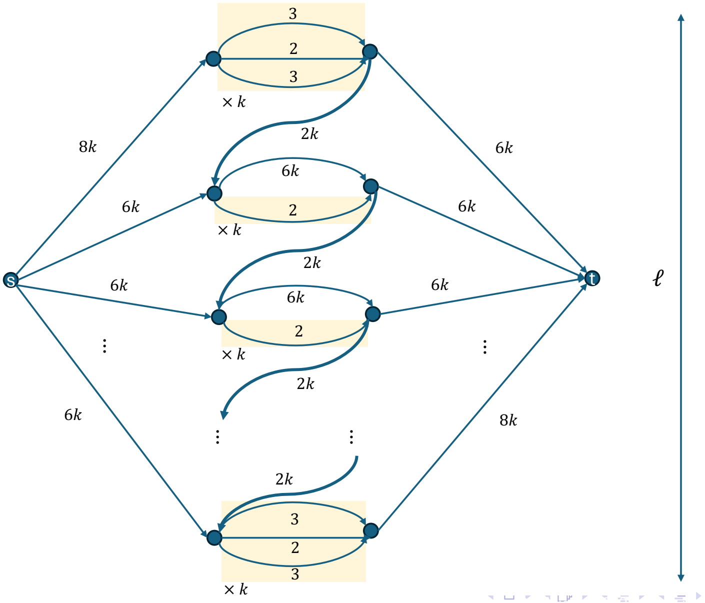
h is an optimal solution to (2)
$\iff g = h + \text{Unitary}(f + h)$ is an optimal solution to (1).

How large can the g get?

Lemma ([Grigorjew et al., 2024])

For all $c > 1$, there are instances with $pw(G)/mfd(G, f) > c$.

How large can the g get?



$O(\log \|f\|)$-approximation

Theorem [Grigorjew et al., 2024]

For all $c > 1$, MFD can be approximated with a factor of $O(\log \|f\|)$ in runtime $O(m \log ||f|| \cdot (\text{mfd}(G, f) + n))$ on Ch_2-f-minor free DAGs and on $GP(c)$-f-minor free DAGs.

$O(\log \|f\|)$-approximation

Theorem [Grigorjew et al., 2024]

For all $c > 1$, MFD can be approximated with a factor of $O(\log \|f\|)$ in runtime $O(m \log ||f|| \cdot (\text{mfd}(G, f) + n))$ on Ch_2-f-minor free DAGs and on $GP(c)$-f-minor free DAGs.

Theorem

For all $c > 1$, MFD can be approxmated with a factor of $O(\log \|f\|)$ on instances with $\text{pw}(G)/\text{mfd}(G, f) < c$.

Complexity result

Corollary [Grigorjew et al., 2024]

MFD can be solved in quasi-polynomial time on $GP(c)$-f-minor free DAGs, when the flow is coded in unary.

Proof:

- Approx. algorithm: $\mathrm{mfd}(G, f) \leq \mathrm{pw}(G) \cdot \log \|f\|$.
- MFD is in FPT [Kloster et al., 2018]: $O(2^{k^2} \cdot (n + \log \|f\|)) \leq O(\|f\|^{\log \|f\|} \cdot (n + \log \|f\|))$.

Further problems

- Good approximation for MFD, when $\mathrm{pw}(G)/\mathrm{mfd}(G,f) > c$.
- Is MFD in APX? What about $GP(c)$-f-minor free graphs?
- PTAS on $GP(c)$-f-minor free graphs?
- Parameterized algorithms for (generalized) max flow or other flow problems?
- Structural graph theorems using flows?

Further problems

- Good approximation for MFD, when $\text{pw}(G)/\text{mfd}(G, f) > c$.
- Is MFD in APX? What about $GP(c)$-f-minor free graphs?
- PTAS on $GP(c)$-f-minor free graphs?
- Parameterized algorithms for (generalized) max flow or other flow problems?
- Structural graph theorems using flows?

Thank you!

Baaijens, J. A., Stougie, L., and Schönhuth, A. (2020).
Strain-aware assembly of genomes from mixed samples using flow variation graphs.
In *International Conference on Research in Computational Molecular Biology*, pages 221–222. Springer.

Cáceres, M., Cairo, M., Grigorjew, A., Khan, S., Mumey, B., Rizzi, R., Tomescu, A. I., and Williams, L. (2024).
Width helps and hinders splitting flows.
ACM Transactions on Algorithms, 20(2):1–20.

Deligkas, A. and Meir, R. (2017).
Directed graph minors and serial-parallel width.
arXiv preprint arXiv:1711.01806.

Dias, F. H. C., Williams, L., Mumey, B., and Tomescu, A. I. (2022).
Fast, Flexible, and Exact Minimum Flow Decompositions via ILP.
arXiv, arXiv:2201.10923.

To appear in the Proceedings of RECOMB 2022 – 26th Annual International Conference on Research in Computational Molecular Biology.

Grigorjew, A., Wanchote, J., Mumey, B., and Tomescu, A. I. (2024).
To appear.

Hartman, T., Hassidim, A., Kaplan, H., Raz, D., and Segalov, M. (2012).
How to split a flow?
In *2012 Proceedings IEEE INFOCOM*, pages 828–836. IEEE.

Kloster, K., Kuinke, P., O'Brien, M. P., Reidl, F., Villaamil, F. S., Sullivan, B. D., and van der Poel, A. (2018).
A practical fpt algorithm for flow decomposition and transcript assembly.
In *2018 Proceedings of the Twentieth Workshop on Algorithm Engineering and Experiments (ALENEX)*, pages 75–86. SIAM.

Tomescu, A. I., Kuosmanen, A., Rizzi, R., and Mäkinen, V. (2013).

A novel min-cost flow method for estimating transcript expression with RNA-Seq.
In *BMC bioinformatics*, volume 14, pages S15:1–S15:10. Springer.