

# The Complexity of Flow Expansion and Electrical Flow Expansion <sup>★</sup>

Dorothea Wagner<sup>[0000–0003–4728–7013]</sup> and Matthias Wolf<sup>[0000–0003–1411–6330]</sup>

Karlsruhe Institute of Technology, Karlsruhe, Germany  
{dorothea.wagner,matthias.wolf}@kit.edu

**Abstract.** FLOWEXPANSION is a network design problem, in which the input consists of a flow network and a set of candidate edges, which may be added to the network. Adding a candidate incurs given costs. The goal is to determine the cheapest set of candidate edges that, if added, allow the demands to be satisfied. FLOWEXPANSION is a variant of the MINIMUM-COST FLOW problem with non-linear edge costs.

We study FLOWEXPANSION for both graph-theoretical and electrical flow networks. In the latter case this problem is also known as the TRANSMISSION NETWORK EXPANSION PLANNING problem. We give a structured view over the complexity of the variants of FLOWEXPANSION that arise from restricting, e.g., the graph classes, the capacities, or the number of sources and sinks. Our goal is to determine which restrictions have a crucial impact on the computational complexity. The results in this paper range from polynomial-time algorithms for the more restricted variants over  $\mathcal{NP}$ -hardness proofs to proofs that certain variants are  $\mathcal{NP}$ -hard to approximate even within a logarithmic factor of the optimal solution.

**Keywords:** Flow Networks · Electrical Flows · Expansion Planning  
· Minimum Cost Flow

## 1 Introduction

Expanding flow networks is a challenging task with a wide range of applications such as deciding where to build new roads, which regions to connect by new rail lines, or where to build new power lines for transmitting electrical power. The latter problem is often called TRANSMISSION NETWORK EXPANSION PLANNING (TNEP or TEP), and a huge body of research on it exists in the electrical engineering community; see for example [9] for a recent survey. What distinguishes

---

<sup>★</sup> This work was funded (in part) by the Helmholtz Program Storage and Cross-linked Infrastructures, Topic 6 Superconductivity, Networks and System Integration and by the German Research Foundation (DFG) as part of the Research Training Group GRK 2153: Energy Status Data – Informatics Methods for its Collection, Analysis and Exploitation.

The final authenticated version is available online at [https://doi.org/10.1007/978-3-030-67731-2\\_32](https://doi.org/10.1007/978-3-030-67731-2_32).

TNEP from other flow expansion problems is the underlying flow model, which needs to capture additional physical laws such as Kirchhoff’s Voltage Law.

In this work we take a more theoretical view on flow expansion problems and TNEP in particular. An instance of FLOWEXPANSION consists of an undirected flow network and a set of *candidate edges* with given costs. The goal is to find the cheapest subset of candidate edges to add to the network such that the resulting network admits a flow satisfying all constraints. As outlined before, the underlying flow model plays a crucial role. We consider two types of flow in this work: (*graph-theoretic*) flows, which require flow conservation at the vertices and impose a maximum flow on each edge, and *electrical flows*, where additional electrical constraints need to be satisfied. We study the computational complexity of FLOWEXPANSION under various restrictions, e.g., restricting the networks to certain graph classes or restricting the number of sources and sinks. Our goal is to understand which restrictions have a crucial impact on the complexity.

The problem of expanding electrical networks (TNEP) is well-studied [9], but the focus in most works lies on having more realistic models of the real-world transmission system or finding more efficient solution methods. From a more theoretical point of view, TNEP is known to be  $\mathcal{NP}$ -hard [12]. The related MAXIMUM TRANSMISSION SWITCHING FLOW problem (MTSF), where edges may be removed instead of added, is  $\mathcal{NP}$ -hard as well even if the underlying graph is series-parallel [6,4]. On general graphs there is no PTAS for MTSF unless  $\mathcal{P} = \mathcal{NP}$  [8]. These results can easily be transferred to TNEP.

Expanding graph-theoretical (non-electrical) flow networks can be considered as a minimum-cost flow problem with non-linear cost functions. We make this connection more precise in Section 2. For the related problem of maximizing the flow subject to a budget constraint heuristics exist [3]. A generalization of FLOWEXPANSION is the FIXED CHARGE TRANSPORTATION PROBLEM, where the cost of each edge with non-zero flow is given by a fixed amount plus an amount proportional to the flow on the edge. It is  $\mathcal{NP}$ -hard as well, and there are several exact algorithms based on integer linear programming [13,10].

*Contribution and Outline.* We give a fine-grained view on the complexity of various variants of FLOWEXPANSION. To this end we classify the variants of FLOWEXPANSION according to the graph classes, the number of sources and sinks, variable vs. fixed production, unit vs. arbitrary capacities, unit vs. arbitrary candidate edge costs, graph-theoretic vs. electrical flow, and (for electrical flows) unit vs. arbitrary resistances. We combine existing complexity results with a variety of new results. These include proofs of the  $\mathcal{NP}$ -hardness of FLOWEXPANSION in trees (Lemma 2) and the  $\mathcal{NP}$ -hardness of approximating FLOWEXPANSION with electrical flows better than within a logarithmic factor of the optimal solution in general graphs (Theorem 4). For more restricted variants, e.g., cacti with fixed production (Lemma 8), we give polynomial-time algorithms.

We formally define FLOWEXPANSION and relate it to the MINIMUM-COST FLOW problem in Section 2. Section 3 contains the hardness proofs for FLOWEXPANSION with variable production. In contrast, having fixed productions makes the problem easier as shown in Section 4. In Section 5 we consider the variants of

FLOWEXPANSION with only one source and one sink. We conclude with a short summary of the results in Section 6.

## 2 Problem Definition

In this work we consider flows in undirected graphs. For notational convenience, we assume that each edge has been given an arbitrary direction. We use  $uv$  to denote the edge between  $u$  and  $v$  with direction towards  $v$ . Let  $G = (V, E)$  be a graph where the edges are directed as described above. Throughout this work we denote the number of vertices and edges by  $n$  and  $m$ , respectively. Let  $\text{cap}: E \rightarrow \mathbb{N}$  be a function that describes the edge capacities. Moreover, there are two disjoint vertex sets  $S$  and  $T$ , which represent sources and sinks in the network. Each sink  $t \in T$  has a given demand  $d(t) \in \mathbb{N}$ . For sources we distinguish two cases. Either each source  $s \in S$  has a fixed production  $p(s) \in \mathbb{N}$  or an upper bound  $\bar{p}(s) \in \mathbb{N}$  is given. We call the tuple  $(G, S, T, \text{cap}, p, d)$  a *flow network with fixed productions* or an *F-network* for short, and the tuple  $(G, S, T, \text{cap}, \bar{p}, d)$  a *flow network with variable production* or a *V-network*.

Note that in a flow network with fixed productions there is no flow if the sum of the productions does not equal the sum of the demands. In the remainder of this work we therefore only consider networks in which these two sums are equal. Similarly, there is no flow in a flow network with variable demands if the total maximum production is less than the total demand, and we only consider flow networks where the total maximum production is at least the total demand.

A (*network*) *flow* in an F- or V-network is a function  $f: E \rightarrow \mathbb{R}$  such that

$$|f(vw)| \leq \text{cap}(vw) \quad \forall vw \in E, \quad (1)$$

$$\sum_{u:uv \in E} f(uv) - \sum_{w:vw \in E} f(vw) = 0 \quad \forall v \in V \setminus (S \cup T), \quad (2)$$

$$\sum_{u:ut \in E} f(ut) - \sum_{w:tw \in E} f(tw) = d(t) \quad \forall t \in T, \quad (3)$$

and for F-networks, we have

$$\sum_{u:us \in E} f(us) - \sum_{w:sw \in E} f(sw) = -p(s) \quad \forall s \in S, \quad (4)$$

whereas for V-networks, we have

$$\sum_{u:us \in E} f(us) - \sum_{w:sw \in E} f(sw) \in [-\bar{p}(s), 0] \quad \forall s \in S. \quad (5)$$

In the context of electrical flows every edge  $e$  is equipped with a positive resistance  $r(e)$  in addition to its capacity. We then call a tuple  $(G, S, T, \text{cap}, r, p, d)$  an *electrical flow network with fixed productions* (an *EF-network*). Similarly, a tuple  $(G, S, T, \text{cap}, r, \bar{p}, d)$  is an *electrical flow network with variable productions* (an *EV-network*). A flow  $f$  is an *electrical flow* if there are vertex potentials

$\varphi: V \rightarrow \mathbb{R}$  such that for all  $uv \in E$  we have  $r(uv) \cdot f(uv) = \varphi(v) - \varphi(u)$ . In flow networks with fixed production there is at most one electrical flow, and if the capacities are sufficiently large, there is exactly one flow [1, Ch. II]. This distinguishes electrical flows from network flows, which are not unique in general.

In the expansion problem, the edge set  $E$  is partitioned into *existing edges*  $E_0$  and *candidate edges*  $E_1$ , and there is a cost  $c(e) \in \mathbb{R}_{\geq 0}$  associated with every candidate edge  $e \in E_1$ . For  $E' \subseteq E_1$  we call the flow network with underlying graph  $H = (V, E_0 \cup E')$  an *expansion* if there is a flow in  $H$ . The sum of all candidate edge costs in an expansion  $H$  is the *total cost* of  $H$  and denoted by  $\text{cost}(H)$ . The objective of FLOWEXPANSION is to find an expansion of minimum total cost. We specify the type of the flow (e.g., electrical flow with fixed productions) by prefixing FLOWEXPANSION with the abbreviation for the flow type (e.g., EF-FLOWEXPANSION). If we do not explicitly prefix FLOWEXPANSION in a statement, the statement is applicable to all four variants. Note that while we consider simple graphs only, parallel edges can be modeled by parallel paths.

FLOWEXPANSION for network flows can be interpreted as a minimum-cost flow problem with a step cost function. The cost for a candidate edge  $e \in E_1$  is 0 if the flow on  $e$  is 0, and  $c(e)$  otherwise. All existing edges have costs 0. Note that to any flow  $f$  in this resulting network there is a flow  $f'$  with only integral flow values costing at most as much as  $f$  since all edge capacities, productions, and demands are integral. The special case that all candidate edges have capacity 1 is therefore equivalent to a minimum-cost flow with linear edge costs. Hence, it can be solved in  $\tilde{O}(m\sqrt{n})$  time [7]. Note that in simple graph classes, e.g., trees, faster algorithms exist. However, our focus in this work lies on distinguishing  $\mathcal{NP}$ -hard cases from cases solvable in polynomial time.

**Lemma 1.** *Finding an optimal solution for F- or V-FLOWEXPANSION with unit capacities is equivalent to solving a minimum-cost flow with linear costs.*

### 3 Networks with Variable Production

First, we consider FLOWEXPANSION with variable productions. We show that it is  $\mathcal{NP}$ -hard even in the simple case that the input network is a tree. Moreover, for general graphs, we prove that EV-FLOWEXPANSION cannot be approximated within a logarithmic factor of the optimal solution unless  $\mathcal{P} = \mathcal{NP}$ .

#### 3.1 $\mathcal{NP}$ -Hardness on Trees

We reduce the problem SUBSETSUM to FLOWEXPANSION with variable production. An instance  $(A, k)$  of SUBSETSUM consists of a finite set  $A \subseteq \mathbb{N}$  and some  $k \in \mathbb{N}$ . The goal is to find a subset of  $A$  with sum  $k$ .

**Lemma 2.** *Finding an optimal solution to FLOWEXPANSION with variable production is  $\mathcal{NP}$ -hard even if the graph is a star.*

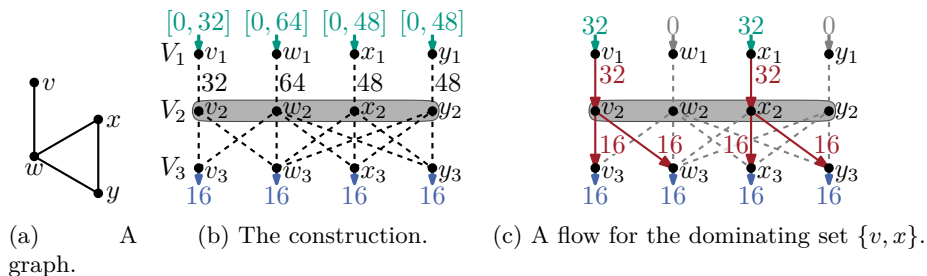


Fig. 1: An example of the network constructed in the reduction from MINIMUM DOMINATING SET in the proof of Theorem 4. The candidate edges are dashed, the intermediate layer is a clique, and  $n = 4$ .

*Proof.* Let  $(A, k)$  be an instance of SUBSETSUM. We construct a star graph  $K_{1,|A|}$ , where the center  $u$  is a sink with demand  $k$ . Each leaf  $v_a$  corresponds to an element  $a \in A$  and is a source with maximum production  $\bar{p}(v_a) = a$ . All edges are candidate edges, and the edge  $uv_a$  has capacity  $a$  and cost  $a$ .

As every leaf and consequently every edge corresponds to one element of  $A$ , there is a one-to-one correspondence between subsets of  $A$  and subsets of the edges. A subset  $E'$  of edges, which corresponds to  $A' \subseteq A$ , is the edge set of an expansion if and only if the sum of the edge capacities is at least  $k$ , which means  $\sum_{a' \in A'} a' \geq k$ , because then the demand of the source can be satisfied. As the cost of  $E'$  is  $\sum_{a' \in A'} a' \geq k$  as well, any expansion costs at least  $k$  and an expansion costing exactly  $k$  corresponds to a solution of the SUBSETSUM-instance. Hence, the SUBSETSUM-instance has a solution if and only if there is an expansion costing exactly  $k$ .  $\square$

### 3.2 Approximation Hardness on General Graphs

For electrical flows, the expansion problems are  $\mathcal{NP}$ -hard by a reduction from STEINERTREE [12, Prop. 1]. This reduction is applicable to all variants of FLOWEXPANSION. Thus, we can transfer the inapproximability of STEINERTREE [2].

**Lemma 3.** *FLOWEXPANSION is  $\mathcal{NP}$ -hard on planar graphs, and on general graphs it is  $\mathcal{NP}$ -hard to approximate FLOWEXPANSION within a factor of  $96/95$  even if there is only one source or only one sink.*

But for EV-flows we can give a logarithmic lower bound on the approximation factor by a reduction from MINIMUMDOMINATINGSET.

**Theorem 4.** *For every  $\varepsilon > 0$  EV-FLOWEXPANSION with unit resistances cannot be approximated within a factor of  $(1/2 - \varepsilon) \cdot \ln(|V|/3) + 1/2$  unless  $\mathcal{P} = \mathcal{NP}$ .*

*Proof.* An instance of the MINIMUMDOMINATINGSET problem consists of an undirected graph  $H = (V_H, E_H)$ . The goal is to find a minimum size vertex set  $D \subseteq V_H$  such that every vertex in  $H$  belongs to  $D$  or is adjacent to a vertex

in  $D$ . Any set with these properties is called a *dominating set*. Given  $H$  we build an instance  $G$  of EV-FLOWEXPANSION such that expansions correspond to dominating sets in  $H$ ; see Fig. 1 for an example. All edges in  $G$  have resistance 1. The graph  $G$  is organized in three layers: the *source layer* with vertex set  $V_1$ , the *intermediate layer* ( $V_2$ ), and the *sink layer* ( $V_3$ ). Each layer contains a copy of all vertices of  $H$ . For  $v \in V_H$  we denote its copy in the source, intermediate, and sink layers by  $v_1$ ,  $v_2$ , and  $v_3$ , respectively. The vertex  $v_1$  is a source with a maximum production of  $\bar{p}(v_1) = (\deg(v)+1) \cdot n^2$ , where  $\deg(v)$  denotes the degree of  $v$  in  $H$  and  $n = |V_H|$ . The vertex  $v_3$  is a sink with demand  $d(v_3) = n^2$ . The vertices in the intermediate layer form a clique of existing edges with capacity 1. Each vertex  $v_1 \in V_1$  is connected to the corresponding vertex  $v_2$  by a candidate edge with capacity  $(\deg(v)+1) \cdot n^2$  and cost  $n$ . Finally, for  $v, w \in V_H$  there is the candidate edge  $v_2w_3$  if and only if  $v = w$  or  $v$  and  $w$  are adjacent in  $H$ . The capacity of these edges is  $n^2$  and their cost is 1.

Consider any expansion of  $G$  and let  $\varphi$  be a potential function of an electrical flow in the expansion. Let  $v_2 \in V_2$  and  $w_3 \in V_3$ . We show that  $\varphi(v_2) < \varphi(w_3)$ . The vertex  $w_3$  is a sink with demand  $n^2$  and incident to at most  $n$  edges. Hence, one of these edges has a flow of at least  $n$ ; say the edge to  $x_2 \in V_2$ . This implies  $\varphi(w_3) - \varphi(x_2) \geq n$ . If  $v_2 = x_2$ , we are done. Otherwise, we note that the capacity of the edge  $x_2v_2$  is 1, and therefore, we have  $\varphi(x_2) - \varphi(v_2) \geq -1$ . Adding these two inequalities gives  $\varphi(w_3) - \varphi(v_2) \geq n - 1 > 0$ , since  $v_2 \neq x_2$  and thus  $n \geq 2$ . Hence, there is no positive flow from the sink layer to the intermediate layer.

Similarly, there is no positive flow from the intermediate layer to the source layer as there are no sinks in the source layer and all sources have degree 1. We claim that in every expansion every sink is connected to a source by a path of length 2. Let  $G' = (V, E')$  be an expansion of  $G$  and let  $v_3 \in V_3$  be a sink. As the demand of  $v_3$  is  $n^2$  and  $v_3$  has at most  $n$  neighbors, there is an edge  $w_2v_3$  with  $f(w_2v_3) \geq n$ . But there is no positive flow to  $w_2$  from the sink layer, and the flow from other intermediate layer vertices is bounded by

$$\sum_{x \in V_2 \setminus \{w_2\}} f(xw_2) \leq \sum_{x \in V_2 \setminus \{w_2\}} \text{cap}(xw_2) = n - 1 < n.$$

Hence, there must be positive flow from the source  $w_1$  to  $w_2$ . Thus, the expansion contains the path  $w_1w_2v_3$ . As there is such a path for every  $v \in V_H$ , we can transform any expansion  $G' = (V, E')$  to a dominating set  $D$  by setting  $D = \{u \in V_H \mid u_1u_2 \in E'\}$ . The existence of length-2 paths to all sinks guarantees that  $D$  dominates all vertices of  $H$ .

Conversely, from any dominating set  $D$  we can construct an expansion of  $G$  of cost  $(|D| + 1) \cdot n$  as follows. The additional edges are the edge  $u_1u_2$  for each  $u \in D$  and for every  $v \in V_H$  the edge  $w_2v_3$ , where  $w \in D$  is an arbitrarily chosen vertex that dominates  $v$ . To see that there is a flow in the expansion, we assign  $\varphi(v_2) = 0$  and  $\varphi(v_3) = n^2$  for all  $v \in V_H$ , i.e., vertices in the intermediate layer have a potential of 0 and sinks have a potential of  $n^2$ . The production  $p(v_1)$  of each source is set to  $n^2$  times the number of neighbors of  $v_2$  in the sink layer, and we set  $\varphi(v_1) = -p(v_1)$ . Note that for  $w \notin D$  we have  $p(w_1) = 0$ . It is easy to verify that this is indeed an expansion with cost  $(|D| + 1) \cdot n$ .

The transformation presented above can be performed in polynomial time. It remains to show a relation between approximate solutions of FLOWEXPANSION and the transformed dominating set.

*Claim.* Let  $c^*$  be the cost of a cost-minimal expansion, and let  $d^*$  be the size of a minimum dominating set. If  $\text{cost}(H) \leq (1 + \alpha)c^*$  for some  $\alpha > 0$ , then  $|D| \leq (1 + 2\alpha) \cdot d^*$ .

*Proof.* We first note that in any expansion each sink is incident to at least one edge, each costing 1. In total the edges between the intermediate and the sink layer cost at least  $n$ . Since each edge between the source and the intermediate layer costs  $n$  and corresponds to one vertex in the dominating set  $D$ , we have  $|D| \leq (\text{cost}(H) - n)/n = \text{cost}(H)/n - 1$ .

From any minimum dominating set of size  $d^*$  we can construct an expansion of cost  $(d^* + 1) \cdot n$ . By the inequality above this is a cost-minimal expansion as otherwise we could construct a smaller dominating set. Hence, we have  $d^* = c^*/n - 1$ . Together with  $d^* \geq 1$  and  $\text{cost}(H) \leq (1 + \alpha) \cdot c^*$ , we obtain

$$|D| \leq \frac{\text{cost}(H)}{n} - 1 \leq (1 + \alpha) \cdot \frac{c^*}{n} - 1 = (1 + \alpha) \cdot (d^* + 1) - 1 \leq (1 + 2\alpha) \cdot d^*. \triangleleft$$

Unless  $\mathcal{P} = \mathcal{NP}$ , MINIMUMDOMINATINGSET cannot be approximated within a factor of  $(1 - \delta) \ln|V_H|$  for  $\delta \in (0, 1)$  in polynomial time [11]. Hence, it follows from the claim that EV-FLOWEXPANSION cannot be approximated within a factor of  $(1/2 - \varepsilon) \ln|V_H| + 1/2 = (1/2 - \varepsilon) \ln(|V|/3) + 1/2$  for  $\varepsilon \in (0, 1/2)$ .  $\square$

## 4 Networks with Fixed Production

Consider an F- or EF-network  $\mathcal{N}$ . We construct an equivalent network  $\mathcal{N}'$  with variable production. Recall that we assume that the sum of all productions equals the sum of all demands. We use the production in  $\mathcal{N}$  as an upper bound for the production in  $\mathcal{N}'$ , i.e., we set  $\bar{p} = p$ . As all demands have to be satisfied, all possible production needs to be used. Hence, any flow in  $\mathcal{N}'$  is a flow in  $\mathcal{N}$ . Clearly, any flow in  $\mathcal{N}$  is a flow in  $\mathcal{N}'$  as well. Thus,  $\mathcal{N}$  and  $\mathcal{N}'$  are equivalent.

**Observation 5.** *For any F-network (EF-network), there is an equivalent V-network (EV-network) on the same graph.*

In the other direction, a V-network  $\mathcal{N}$  can be transformed into an equivalent F-network  $\mathcal{N}'$ . We add a super source  $s^*$  and connect it to all original sources. We set the capacity of each such edge  $s^*s$  to  $\bar{p}(s)$ . The capacities of the other edges stay the same. We define the new set of sources  $S' = \{s^*\}$  and set  $p(s^*) = \sum_{t \in T} d(t)$ . Restricting any flow  $f'$  in  $\mathcal{N}'$  to the original edges yields a flow in  $\mathcal{N}$ . Conversely, any flow in  $\mathcal{N}$  can be extended to a flow in  $\mathcal{N}'$ .

**Observation 6.** *For any V-network there is an equivalent F-network with one additional vertex that is connected to all sources in the original network.*

Note that adding the super source to a graph in a certain graph class  $\mathcal{C}$  may create a graph that does not belong to  $\mathcal{C}$  anymore, e.g., a planar graph may become non-planar. Further, the construction does not work for electrical flow networks because it is not clear how to set the resistances of the new edges.

Applying the construction that leads to Observation 6 to the one in Lemma 2, we obtain that F-FLOWEXPANSION is  $\mathcal{NP}$ -hard on parallel paths. By setting the resistances of the edges such that all paths induce the same angle difference when their edges are saturated, we further see that EF-FLOWEXPANSION is  $\mathcal{NP}$ -hard as well. By Observation 5 the same holds if the production is variable.

**Lemma 7.** *FLOWEXPANSION is  $\mathcal{NP}$ -hard on graphs formed by parallel paths between the only source and the only sink.*

In the remainder of this section, we make some observations that lead to an  $O(n^2)$  time algorithm for F- and EF-FLOWEXPANSION on cacti, which are graphs in which two simple cycles do not share any edges. This contrasts with V- and EV-FLOWEXPANSION, which are  $\mathcal{NP}$ -hard on trees by Lemma 2.

Suppose  $e$  is a candidate edge and a bridge, i.e.,  $G - e$  contains two components. If the generation equals the demand in both components, there is a cost-minimal expansion without  $e$ . Otherwise, any expansion contains  $e$ .

Let  $v$  be a cut-vertex, i.e., the graph  $G - v$  consists of  $k \geq 2$  connected components  $G_1, \dots, G_k$ . There is a unique way of distributing the demand or production of  $v$  to demands or productions of  $H_i := G_i + v$  such that the total demand in  $H_i$  equals the total production. The graphs  $H_1, \dots, H_k$  can then be considered independently. Repeating this process,  $G$  can be split into its blocks (maximal biconnected subgraphs). With these observations we obtain quadratic time algorithms for FLOWEXPANSION with fixed productions on cacti.

**Lemma 8.** *F- and EF-FLOWEXPANSION on cacti can be solved in  $O(n^2)$  time.*

*Proof.* All blocks can be handled independently, and blocks containing bridges can be handled as described above. All other blocks are cycles. Let  $C$  be a cycle with  $\ell$  edges. In a cost-minimal expansion either all edges are included or at least one of the edges is missing. We check for all  $\ell + 1$  cases whether there is a flow in the resulting subgraph of  $C$  and, if so, we compute the cost of the expansion ignoring the candidate edges with flow 0. The computation for  $C$  takes  $O(\ell^2)$  time, and hence  $O(n^2)$  time for the whole network.  $\square$

## 5 Single Source, Single Sink

In this section we consider the special case that there is only one source and one sink in the network. Since there is only one source, its production is determined by the demand. We may therefore consider the production as fixed.

FLOWEXPANSION is  $\mathcal{NP}$ -hard on graphs where the source and the sink are connected by parallel paths by Lemma 7. But for series-parallel graphs a straightforward dynamic programming on the graph structure yields a pseudo-polynomial time algorithm for F-FLOWEXPANSION, where one terminal of the graph is the only source and the other one is the only sink.



**Lemma 9.** *On series-parallel graphs with one source and one sink, the problem F-FLOWEXPANSION can be solved in  $O(n^3 \cdot \min\{C^2, U^2\})$  time, where  $C$  and  $U$  are upper bounds for the edge costs and capacities, respectively.*

The analogous problem for electrical flows, EF-FLOWEXPANSION, however, is  $\mathcal{NP}$ -hard even if all capacities are 1 and edge costs are ignored [4]. Hence, there is no pseudo-polynomial time algorithm for this problem unless  $\mathcal{P} = \mathcal{NP}$ .

But we prove that there is a pseudo-polynomial time algorithm for FLOWEXPANSION on a subclass of the series-parallel graph, namely those graphs that are formed by a series composition of blocks consisting of parallel paths. We call such graphs *sps-graphs* (for series-parallel-series).

**Lemma 10.** *In sps-graphs with one source and one sink, FLOWEXPANSION can be solved in pseudo-polynomial time.*

*Proof.* We note that in sps-graphs each block, which consists of parallel paths, can be considered independently. We reduce finding an optimal solution for FLOWEXPANSION on parallel paths to solving a linear number of MINKNAPSACK instances. An instance of MINKNAPSACK consists of a set of objects  $A$ , some  $k \in \mathbb{R}_{>0}$ , as well as a cost  $c(a)$  and a value  $v(a)$  for each  $a \in A$ . The goal is to find a subset  $A' \subseteq A$  with minimum cost such that the sum of the values is at least  $k$ . A straightforward dynamic programming approach yields a pseudo-polynomial time algorithm for MINKNAPSACK [5, Sec. 13.3.3].

We only present the reduction for EF-FLOWEXPANSION; similar approaches work for all other variants. Let  $P_1, \dots, P_k$  be the paths between the source  $s$  and the sink  $t$ , each associated with its costs  $c(P_i)$ , resistance  $r(P_i)$ , and capacity  $\text{cap}(P_i)$ . Let further be  $\Phi(P_i) = \text{cap}(P_i) \cdot r(P_i)$ , which describes the maximum potential difference induced by  $P_i$ . We assume that the paths are ordered such that  $\Phi(P_i) \leq \Phi(P_j)$  for  $i \leq j$ . We call the paths with candidate edges *candidate paths* and the others *existing paths*.

In any expansion there is a path  $P_j$  with minimum index that is completely contained in the expansion. This path  $P_j$  restricts the potential difference between  $s$  and  $t$  to (at most)  $\Phi(P_j)$ . For each choice of  $P_j$ , we reduce the search for a cost-minimal expansion to an instance of MINKNAPSACK. For each path  $P_i$  with  $i \geq j$  we define  $F(P_i) = \Phi(P_j)/r(P_i)$ . This describes the flow along  $P_i$  if the potential difference is exactly  $\Phi(P_j)$ . With the fixed potential difference of  $\Phi(P_j)$ , the remaining demand  $d'$  is the original demand  $d(t)$  minus the flows  $F(P_i)$  along all existing paths  $P_i$ . The goal is then to find a subset  $\Pi'$  of the candidate paths  $P_i$  with  $i \geq j$  such that  $\sum_{P \in \Pi'} F(P) \geq d'$  and  $\sum_{P \in \Pi'} c(P)$  is minimal. This is exactly an instance of the MINKNAPSACK problem, which can be solved in pseudo-polynomial time. The cost-minimal expansion then corresponds to the cheapest solution of one of the MINKNAPSACK instances for  $j \in \{1, \dots, k\}$ .  $\square$

## 6 Conclusion

FLOWEXPANSION is an optimization problem that can be seen as a variant of the MINIMUM-COST FLOW problem. It is  $\mathcal{NP}$ -hard even in very simple cases, e.g.,

if the network is a star (Lemma 2) or if there are only parallel paths between the only source and the only sink (Lemma 7). The more general variants even prove to be hard to approximate (Theorem 4). For some restricted cases, there are (pseudo-)polynomial time algorithms. Our results show that for all parameters (number of sources and sinks, capacities, edge costs, resistances, fixed vs. variable production, and the flow model) there are cases where the complexity differs. One of the most notable cases is the restriction of the capacities to 1, which reduces F- and V-FLOWEXPANSION to a standard minimum-cost flow.

This work studies the complexity of the main variants of FLOWEXPANSION, but the complexity of some special cases is still open. It may also be possible to extend the results for series-parallel graphs to graphs of bounded treewidth.

## References

1. Bollobás, B.: Modern graph theory. Graduate texts in mathematics; 184, Springer, New York (1998)
2. Chlebík, M., Chlebíková, J.: The Steiner tree problem on graphs: Inapproximability results. *Theoretical Computer Science* **406**(3), 207 – 214 (2008). <https://doi.org/10.1016/j.tcs.2008.06.046>
3. Eiselt, H.A., von Frajer, H.: On the budget-restricted max flow problem. *Operations-Research-Spektrum* **3**(4), 225–231 (1982). <https://doi.org/10.1007/BF01719791>
4. Grastien, A., Rutter, I., Wagner, D., Wegner, F., Wolf, M.: The maximum transmission switching flow problem. In: Proceedings of the Ninth International Conference on Future Energy Systems. pp. 340–360. e-Energy '18, ACM, New York, NY, USA (2018). <https://doi.org/10.1145/3208903.3208910>
5. Kellerer, H., Pferschy, U., Pisinger, D.: Knapsack problems. Springer, Berlin (2004)
6. Kocuk, B., Jeon, H., Dey, S.S., Linderoth, J., Luedtke, J., Sun, X.A.: A cycle-based formulation and valid inequalities for DC power transmission problems with switching. *Operations Research* **64**(4), 922–938 (2016). <https://doi.org/10.1287/opre.2015.1471>
7. Lee, Y.T., Sidford, A.: Path finding II : An  $\tilde{O}(m \sqrt{n})$  algorithm for the minimum cost flow problem. *CoRR abs/1312.6713* (2013), <https://arxiv.org/abs/1312.6713>
8. Lehmann, K., Grastien, A., Hentenryck, P.V.: The complexity of DC-switching problems. *CoRR abs/1411.4369* (2014), <https://arxiv.org/abs/1411.4369>
9. Mahdavi, M., Sabillon Antunez, C., Ajalli, M., Romero, R.: Transmission expansion planning: Literature review and classification. *IEEE Systems Journal* pp. 1–12 (2018). <https://doi.org/10.1109/JSYST.2018.2871793>
10. Mingozzi, A., Roberti, R.: An exact algorithm for the fixed charge transportation problem based on matching source and sink patterns. *Transportation Science* **52**(2), 229–238 (2018). <https://doi.org/10.1287/trsc.2017.0742>
11. Moshkovitz, D.: The projection games conjecture and the NP-hardness of  $\ln n$ -approximating set-cover. *Theory of Computing* **11**(7), 221–235 (2015). <https://doi.org/10.4086/toc.2015.v011a007>
12. Moulin, L.S., Poss, M., Sagastizábal, C.: Transmission expansion planning with re-design. *Energy Systems* **1**(2), 113–139 (2010). <https://doi.org/10.1007/s12667-010-0010-9>

13. Roberti, R., Bartolini, E., Mingozzi, A.: The fixed charge transportation problem: An exact algorithm based on a new integer programming formulation. *Management Science* **61**(6), 1275–1291 (2015). <https://doi.org/10.1287/mnsc.2014.1947>