A Characterization of Irreducible Infeasible Subsystems in Flow Networks

Imke Joormann[*], James B. Orlin[**], and Marc E. Pfetsch[***]

[*]Institute for Mathematical Optimization, TU Braunschweig, Germany, i.joormann@tu-braunschweig.de
[**]Sloan School of Management, Massachusetts Institute of Technology, USA, jorlin@mit.edu
[***]Research Group Optimization, Dept. of Mathematics, TU Darmstadt, Germany, pfetsch@mathematik.tu-darmstadt.de

May 2016

Abstract

Infeasible network flow problems with supplies and demands can be characterized via violated cut-inequalities of the classical Gale-Hoffman theorem. Written as a linear program, irreducible infeasible subsystems (IISs) provide a different means of infeasibility characterization. In this article, we answer a question left open in the literature by showing a one-to-one correspondence between IISs and Gale-Hoffman-inequalities in which one side of the cut has to be weakly connected. We also show that a single max-flow computation allows one to compute an IIS. Moreover, we prove that finding an IIS of minimal cardinality in this special case of flow networks is strongly $\mathcal{NP}$-hard.

Keywords – flow network, irreducible infeasible subsystem, Gale-Hoffman theorem, infeasibility analysis, witness, max flow–min cut, $\mathcal{NP}$-hardness

1 Introduction

Sometimes a linear program (LP) turns out to be infeasible, e.g., because of modeling errors or structural reasons. In this case, one would like to find the cause for its infeasibility. One way is to study *irreducible infeasible subsystems* (IISs), i.e., infeasible subsystems such that each proper subsystem is feasible. IISs might help to identify the reason of infeasibility and are basic structures in infeasibility analysis.

In this article, we study the special case of a flow system for a simple, directed graph $G = (V, A)$ with upper flow bounds $u \in \mathbb{R}^A$, lower flow bounds $\ell \in \mathbb{R}^A$, and a supply vector $b \in \mathbb{R}^V$. Thus, we consider the system

$$\begin{aligned} x(\delta^+(v)) - x(\delta^-(v)) &= b(v) \qquad \forall\, v \in V, \\ \ell \le x &\le u, \end{aligned} \tag{F}$$

where, for $S \subseteq V$ and $\bar{S} \coloneqq V \setminus S$, we use the following standard notation: $\delta^+(S) \coloneqq \{(v, w) \in A \mid v \in S,\ w \in \bar{S}\}$, $\delta^-(S) \coloneqq \{(v, w) \in A \mid v \in \bar{S},\ w \in S\}$, and $\delta(S) \coloneqq \delta^+(S) \cup \delta^-(S)$; we also abbreviate $\delta^+(v) \coloneqq \delta^+(\{v\})$ and similarly for $\delta^-(v)$ and $\delta(v)$. Moreover, for some vector $y \in \mathbb{R}^I$ with finite index sets I and $I' \subseteq I$, we use $y(I') \coloneqq \sum_{i \in I'} y_i$ and often write $y(i) \coloneqq y(\{i\}) = y_i$. We assume throughout the article that $0 \le \ell \le u$ in order to avoid trivial infeasibilities.

A characterization of feasibility for this flow system is well known:

Theorem 1 (Gale and Hoffman [11, 18])**.** *The network flow system* (F) *is infeasible if and only if there exists $S \subseteq V$ such that*

$$b(S) > u(\delta^+(S)) - \ell(\delta^-(S)). \tag{1}$$

A natural question is how IISs in the flow case and the *Gale-Hoffman-inequalities* (1) (GH-inequalities) are related. In this article, we show that the IISs of (F) correspond to exactly those violated inequalities (1) for which the induced subgraph $G[S]$ is weakly connected, i.e., the undirected version of $G[S]$ is connected. This implies, for instance, that there can be exponentially many IISs; see Corollary 9. The corollary follows with a result by Wallace and Wets [28], who showed that a GH-inequality is nonredundant if and only if S *and* $\bar{S}$ are weakly connected. This was generalized for multicommodity flows by Zullo [29].

Further related work in the literature includes Greenberg [14], who discusses the analysis of infeasible flow systems (see also [13]). He presents several heuristics to "localize" the cause of infeasibility, i.e., he tries to isolate small sets S with violated GH-inequalities. In [15], Greenberg further gives an example of a violated GH-inequality that does not lead to an IIS and states that "there is presently no theory to construct an IIS from a violating cut, other than general methods [...]". The missing link is connectivity of one of the sides of the cut corresponding to a GH-inequality; see Section 2.

The problem of finding small sets S with violated GH-inequalities was investigated by Aggarwal et al. [1]. They call S a *witness* of infeasibility and show that the problem of finding a *minimum* witness (i.e., one of smallest cardinality) is strongly $\mathcal{NP}$-hard. They further design an efficient algorithm to find a *minimal* witness (w.r.t. inclusion) based on preflow-push algorithms.

IISs and witnesses in flow networks can both be used to reveal a smaller portion of the network "witnessing" the infeasibility. For witnesses, the number of nodes is relevant, while for IISs, the number of constraints corresponding to both nodes and arcs count. In Section 4, we further discuss the relation of IISs and witnesses. With respect to computational complexity, IISs have similar properties as witnesses: We show, in Section 5, that the *minimum IIS problem*, i.e., to find an IIS of smallest cardinality, is strongly $\mathcal{NP}$-hard, extending the result that this problem is strongly $\mathcal{NP}$-hard for linear inequality systems [3]. An IIS, however, can be computed in polynomial time using one maximum flow computation; see Section 3.

For the analysis of general infeasible inequality systems, many approaches have been developed – we refer to the book of Chinneck [9] for an overview and only mention selected references here. The term IIS was coined by van Loon [27] as a means to analyze infeasibilities in linear programs. Gleeson and Ryan [12] gave a characterization of IISs related to an alternative polyhedron. The analysis of infeasible linear programs was further discussed by Greenberg and Murphy [16], including the case of flow networks. Dravnieks and Chinneck [10] and Chinneck [7] developed heuristics to isolate IISs (see also [5] for an application to flow systems). An overview of these approaches appeared in [6]. Finally, Ryan [25] investigated combinatorial properties of IISs.

Related topics are methods to find maximum feasible subsystems (which are complementary to covers of IISs). Chinneck [4, 8] develops heuristics for this problem. An exact algorithm appears in [23], based on [3]. Moreover, McCormick [22] studies the related problem of finding least infeasible flows.

The remainder of this paper is structured as follows: In Section 2, we show the mentioned correspondence between IISs and GH-inequalities. Section 3 shows that an IIS can be computed in polynomial time, using a single max-flow computation. In Section 4, we discuss the relation of minimum/minimal witnesses and IISs. Section 5 shows that, even in the flow case, it is strongly $\mathcal{NP}$-hard to compute an IIS of minimum cardinality. We close with an outlook on future research in Section 6.

2 IISs Correspond to Connected Gale-Hoffman-Inequalities

To show the correspondence between IISs and GH-inequalities, we need the following definition, where a node set is called *weakly connected* if it is connected in the underlying undirected network.

Definition 2 (GH-cuts)**.** *For $S \subseteq V$, a cut $\delta(S)$ is called a* GH-cut *if $b(S) > u(\delta^+(S)) - \ell(\delta^-(S))$. A GH-cut is* connected *if $G[S] \coloneqq (S, A[S])$ is weakly connected, where $A[S] \coloneqq \{(u,v) \in A \mid u,v \in S\}$. For a GH-cut $\delta(S)$, we call the system*

$$\begin{aligned} x(\delta^+(v)) - x(\delta^-(v)) &= b(v) && \forall\, v \in S, \\ x_a &\geq \ell_a && \forall\, a \in \delta^-(S), \\ x_a &\leq u_a && \forall\, a \in \delta^+(S), \end{aligned} \tag{2}$$

a GH-subsystem*, denoted by $\mathcal{I}(S)$, and S the associated* GH-set*.*

The characterization of infeasible network problems in Theorem 1 could just as well have been given in terms of the complementary form of the violated GH-inequalities:

$$-b(S) > u(\delta^-(S)) - \ell(\delta^+(S)).$$

We call this the *demand form*, since it belongs to a subset with an unsatisfied demand, whereas the form stated in Theorem 1 corresponds to a subset with an unmet supply. It is obvious that if for a subset $S \subseteq V$ the supply form is violated, $\bar{S}$ violates the demand form, and vice versa. However, S might be connected, while $\bar{S}$ is not, and conversely. This means that we have to take both forms into account when determining IISs. To this end, the *GH-demand-subsystem* is defined analogously to Definition 2:

$$\begin{aligned} x(\delta^+(v)) - x(\delta^-(v)) &= b(v) && \forall\, v \in S, \\ x_a &\geq \ell_a && \forall\, a \in \delta^+(S), \\ x_a &\leq u_a && \forall\, a \in \delta^-(S). \end{aligned}$$

Obviously, every GH-subsystem is infeasible.

We need some more notation: Let σ_v refer to the flow conservation constraint for a node $v \in V$, μ_a to the constraint $x_a \leq u_a$, and λ_a to the constraint $x_a \geq \ell_a$ for $a \in A$. Moreover, for a subset $\mathcal{J}$ of constraints of (F), we define $S(\mathcal{J}) \coloneqq \{v \in V \mid \sigma_v \in \mathcal{J}\}$.

The following observation will be used in the proofs below: Every subset $\mathcal{J}$ of constraints of (F) itself defines a network problem, consisting of the nodes $v \in S(\mathcal{J})$ and incident arcs. For arcs with a missing endnode, we introduce an auxiliary node r; missing bounds are replaced by $\pm\infty$. Hence, this network is given by $G(\mathcal{J}) \coloneqq (V(\mathcal{J}), A(\mathcal{J}))$ with

$$\begin{aligned} V(\mathcal{J}) &\coloneqq S(\mathcal{J}) \cup \{r\}, \\ A(\mathcal{J}) &\coloneqq \{(v,w) \in A \mid v,w \in S(\mathcal{J})\} \cup \\ &\qquad \{(v,r) \mid (v,w) \in \delta^+(S(\mathcal{J}))\} \cup \{(r,w) \mid (v,w) \in \delta^-(S(\mathcal{J}))\} \cup \\ &\qquad \{(r,r)_a \mid \mu_a \in \mathcal{J},\ a \cap S(\mathcal{J}) = \emptyset\} \cup \{(r,r)_a \mid \lambda_a \in \mathcal{J},\ a \cap S(\mathcal{J}) = \emptyset\}. \end{aligned}$$

Here, $(r,r)_a$ refers to a loop indexed by arc $a \in A$. The construction implies that for each arc $a \in A(\mathcal{J})$ there exists a unique arc $\tilde{a}$ in A from which a originates (but not conversely). We then inherit the bounds for arcs represented in $\mathcal{J}$ and use $\pm\infty$ otherwise. Thus, we obtain for $a \in A(\mathcal{J})$ and corresponding $\tilde{a} \in A$:

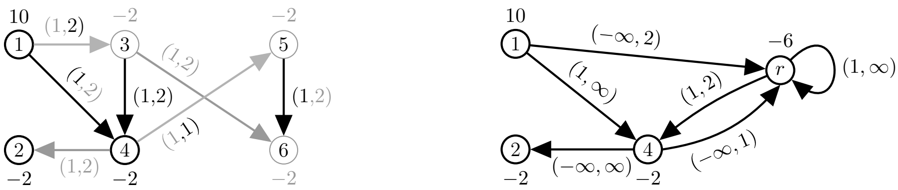
$$u(\mathcal{J})_a \coloneqq \begin{cases} u_{\tilde{a}}, & \mu_{\tilde{a}} \in \mathcal{J}, \\ \infty, & \mu_{\tilde{a}} \notin \mathcal{J}, \end{cases} \qquad \ell(\mathcal{J})_a \coloneqq \begin{cases} \ell_{\tilde{a}}, & \lambda_{\tilde{a}} \in \mathcal{J}, \\ -\infty, & \lambda_{\tilde{a}} \notin \mathcal{J}. \end{cases}$$

Finally, the (balanced) supply/demand for $v \in V(\mathcal{J})$ is defined as follows:

$$b(\mathcal{J})_v \coloneqq \begin{cases} b_v, & v \in S(\mathcal{J}), \\ -\sum_{v \in S(\mathcal{J})} b_v, & v = r. \end{cases}$$

Example 3. Figure 1 shows an example for the construction of $G(\mathcal{J})$, where $\mathcal{J}$ is

$$\begin{aligned} x_{13} + x_{14} &= 10, && x_{13} \leq 2, && x_{56} \geq 1, \\ -x_{42} &= -2, && x_{14} \geq 1, && x_{45} \leq 1, \\ -x_{14} - x_{34} + x_{42} + x_{45} &= -2, && 1 \leq x_{34} \leq 2. && \end{aligned}$$



(a) $\mathcal{J}$ marked in black (b) $G(\mathcal{J})$

Figure 1: Example for the construction of a network problem $(G; b, u, \ell)$ describing an infeasible subset of constraints $\mathcal{J}$ with $S(\mathcal{J}) = \{1, 2, 4\}$; node labels b, arc labels $[\ell, u]$.

The above construction allows us to apply Theorem 1 to a constraint subset $\mathcal{J}$. Note that a GH-inequality cannot involve infinite bound values, otherwise it would not be violated.

Lemma 4. *$\mathcal{J}$ is infeasible if and only if there exists a violated GH-inequality in $G(\mathcal{J})$.*

Proof. Ignoring infinite bounds, every constraint arising from the network problem for $(G(\mathcal{J}); b(\mathcal{J}), u(\mathcal{J}), \ell(\mathcal{J}))$ is also a constraint in $\mathcal{J}$, with the exception of the flow conservation constraint for r. Thus, $\mathcal{J}$ is infeasible if and only if the network problem for $(G(\mathcal{J}); b(\mathcal{J}), u(\mathcal{J}), \ell(\mathcal{J}))$ is infeasible. By Theorem 1, this holds if and only if there exists a violated GH-inequality. □

The following lemmas will show that connected GH-cut systems and IISs are equivalent (see Theorem 8 below).

Lemma 5. *Every IIS of (F) is a GH-subsystem.*

Proof. Let $\mathcal{J}$ be an IIS. Then $S(\mathcal{J}) \neq \emptyset$ (by the assumption $\ell \leq u$). Since $\mathcal{J}$ is infeasible, there exist at least two GH-sets S' and $\bar{S}'$ in $G(\mathcal{J})$ by Lemma 4. W.l.o.g. let $r \in \bar{S}'$, whence $S' \subseteq S(\mathcal{J})$. Because bounds not in $\mathcal{J}$ are infinite, $\mathcal{I}(S') \subseteq \mathcal{J}$, since otherwise the corresponding GH-inequality would automatically be satisfied. Then, if $\mathcal{I}(S') = \mathcal{J}$ (i.e., $S' = S(\mathcal{J})$ and all bounds on $A(\mathcal{J})[S(\mathcal{J})]$ are infinite), we have a GH-subsystem for $\mathcal{J}$. Otherwise the infeasible $\mathcal{I}(S')$ would be a proper subset of $\mathcal{J}$, and $\mathcal{J}$ would not be irreducible. □

We can now use the fact that every IIS has the form $\mathcal{I}(S)$ (see Definition 2). In the following, $\dot\cup$ denotes a *disjoint* union.

Lemma 6. *The GH-cut for an IIS is connected.*

Proof. Let $\mathcal{I}(S)$ be an IIS and suppose S is disconnected, i.e., there is a nontrivial partition $S_1 \,\dot\cup\, S_2 = S$, with $A[S_1] \,\dot\cup\, A[S_2] = A[S]$ and $\delta(S_1) \,\dot\cup\, \delta(S_2) = \delta(S)$. Consider the systems $\mathcal{I}_1 \coloneqq \{\sigma_v \mid v \in S_1\} \cup \{\lambda_a \in \mathcal{I} \mid a \in \delta(S_1)\} \cup \{\mu_a \in \mathcal{I} \mid a \in \delta(S_1)\}$ and $\mathcal{I}_2 \coloneqq \{\sigma_v \mid v \in S_2\} \cup \{\lambda_a \in \mathcal{I} \mid a \in \delta(S_2)\} \cup \{\mu_a \in \mathcal{I} \mid a \in \delta(S_2)\}$. Note that $\mathcal{I}_1$ and $\mathcal{I}_2$ are proper subsets of $\mathcal{I}(S)$. Thus, they must have feasible solutions x^1 and x^2, respectively. But then

$$x_a \coloneqq \begin{cases} x_a^1, & a \in A[S_1] \cup \delta(S_1), \\ x_a^2, & a \in A[S_2] \cup \delta(S_2) \end{cases}$$

gives a feasible solution for $\mathcal{I}(S)$, because $\delta(S_1) \cap \delta(S_2) = \emptyset$; a contradiction. □

Lemma 7. *Every connected GH-subsystem is an IIS.*

Proof. Let $\mathcal{I} \coloneqq \mathcal{I}(S)$ be a connected GH-subsystem (and hence infeasible by Definition 2), and suppose $\mathcal{I}$ is reducible. Then there exists an infeasible $\mathcal{I}' \subset \mathcal{I}$.

Suppose first that $\mathcal{I}'$ is obtained from $\mathcal{I}$ by dropping at least one node constraint σ_v, and consider the network problems on $G(\mathcal{I})$ and $G(\mathcal{I}')$. $G(\mathcal{I}')$ results from $G(\mathcal{I})$ by removing v and redirecting incident arcs to the root node r. By Lemma 4, the graph $G(\mathcal{I}')$ must contain a GH-set S' with $S' \subset S$. Since $\mathcal{I}$ is a GH-subsystem, $u_a = \infty$ and $\ell_a = -\infty$ for all $a \in A(\mathcal{I})[S]$.

Because S is connected, there exists $a \in \delta(S') \cap A(\mathcal{I})[S]$. But then $u(\delta^+(S')) - \ell(\delta^-(S')) = \infty > b(S')$, i.e., $\mathcal{I}'$ would be feasible.

Now suppose $\mathcal{I}'$ contains at least one fewer bound on some $a \in \delta(S)$ than $\mathcal{I}$, i.e., the graph $G(\mathcal{I}')$ is obtained from $G(\mathcal{I})$ by setting $u_a = \infty$ or $\ell_a = -\infty$, respectively. Again, $b(S) < \infty = u(\delta^+(S)) - \ell(\delta^-(S))$ and no GH-cut exists. As a consequence of Theorem 1, every subset of $\mathcal{I}$ is feasible, so $\mathcal{I}$ is indeed irreducible. □

Using Lemma 5, Lemma 6, and Lemma 7, we have proven the following theorem:

Theorem 8. *A subsystem $\mathcal{I}$ of the network flow system (F) is an IIS if and only if $\mathcal{I}$ is the GH-subsystem of a connected GH-cut.*

Corollary 9. *There can be exponentially many IISs of (F).*

Proof. As mentioned in the introduction, Wallace and Wets [28] proved that the GH-inequalities are nonredundant if and only if both sides of the cut are weakly connected. Moreover, they showed that there can be exponentially (in the number of nodes) many nonredundant GH-cuts. Hence, there can be exponentially many IISs of (F). □

3 Computing IISs in Flow Networks

An IIS of a linear program can be computed in weakly polynomial time by finding a vertex of the alternative polyhedron (Gleeson and Ryan [12]). For network flow problems, this can even be done in strongly polynomial time; see Tardos [26]. In the following, we show that performing a single max-flow computation suffices to find an IIS, yielding a strongly polynomial time combinatorial algorithm.

First suppose that a GH-cut $\delta(S)$ has been found. If S is weakly connected, we have found an IIS by Theorem 8; otherwise, we can extract an IIS in roughly the time needed to compute the connected components:

Proposition 10. *Let $\mathcal{I}(S)$ be a GH-subsystem, where S is* disconnected. *Then there exists a connected component $(S_1, A[S_1])$ with $S_1 \subset S$ such that $\mathcal{I}(S_1)$ is an IIS.*

Proof. Let $\mathcal{I}(S)$ be a GH-subsystem, and suppose there are $k \geq 2$ connected components $(S_1, A[S_1])$, …, $(S_k, A[S_k])$ with $S = S_1 \dot\cup \cdots \dot\cup S_k$. W.l.o.g., we assume

$$b(S_1) - u(\delta^+(S_1)) + \ell(\delta^-(S_1)) \;\geq\; b(S_i) - u(\delta^+(S_i)) + \ell(\delta^-(S_i)) \tag{3}$$

for all $i = 2, \dots, k$. Since by assumption $\delta(S_i) \cap \delta(S_j) = \emptyset$ for all $i \neq j$, we have

$$0 < b(S) - u(\delta^+(S)) + \ell(\delta^-(S)) = \sum_{i=1}^{k} \big(b(S_i) - u(\delta^+(S_i)) + \ell(\delta^-(S_i))\big)$$
$$\overset{(3)}{\leq} k\big(b(S_1) - u(\delta^+(S_1)) + \ell(\delta^-(S_1))\big).$$

Thus, we have $b(S_1) > u(\delta^+(S_1)) - \ell(\delta^-(S_1))$ and $\mathcal{I}(S_1)$ is an IIS by Theorem 8. For $\mathcal{I}(S)$ in demand form, the proof runs analogously. □

Therefore, to obtain an IIS, we essentially just need to find a violated GH-inequality. It is well known [2] that this can be done by computing a maximum (s-t)-flow for the following extended network $G' = (V', A')$ with source s, sink t, and capacities u': The nodes are $V' \coloneqq V \cup \{s, t\}$. For each $v \in V$, let

$$d_v \coloneqq b_v - \sum_{a \in \delta^+(v)} \ell_a + \sum_{a \in \delta^-(v)} \ell_a.$$

The arcs are

$$A' \coloneqq A \cup \{(s, v) \mid d_v > 0\} \cup \{(v, t) \mid d_v < 0\}.$$

For each arc of the form $a = (s, v)$, define $u'_a := d_v > 0$. For each arc $a = (v, t)$, define $u'_a := -d_v > 0$. For every arc $a \in A$, set $u'_a := u_a - \ell_a$. All lower bounds in G' are 0.

Let $D := \sum_{v \in V : d_v > 0} d_v$. If the maximum ($s$-$t$)-flow in G' has value $< D$, the original instance is infeasible. In fact, we show that every GH-set in G corresponds to a cut in G' with capacity $< D$. Aggarwal et al. [1, Lemma 1] proved a similar result for the witness problem. Our result is an extension to problems with lower bounds and holds w.r.t. IISs.

Lemma 11. *There is a one-to-one correspondence between GH-sets $S \subseteq V$ in G and sets $S' = S \cup \{s\} \subseteq V'$ with a cut value $u'(\delta^+(S')) < D$ in G'.*

Proof. In the following, we will use the notation $S^+ := \{v \in S \mid d_v > 0\}$ and $\bar{S}^+ := \{v \in V \setminus S \mid d_v > 0\}$, and analogously S^-, $\bar{S}^-$ for the demand nodes. Furthermore, $[V_1 : V_2] := \{(v_1, v_2) \in A \mid v_1 \in V_1, v_2 \in V_2\}$.

Let $S \subseteq V$ and $S' := S \cup \{s\}$. Then

$$\begin{aligned} u'(\delta^+(S')) &= u'(\delta^+(S)) + u'([s : \bar{S}^+]) \\ &= u(\delta^+(S)) - \ell(\delta^+(S)) - d(S^-) + d(\bar{S}^+). \end{aligned}$$

In the last line, all values are with respect to the graph G.

Note that for every $U \subseteq V$, we have $d(U) = b(U) - \ell(\delta^+(U)) + \ell(\delta^-(U))$. Moreover, $D = d(S^+) + d(\bar{S}^+)$, i.e., $d(\bar{S}^+) = D - d(S^+) = D - b(S^+) + \ell(\delta^+(S^+)) - \ell(\delta^-(S^+))$. Thus,

$$\begin{aligned} u'(\delta^+(S')) = u(\delta^+(S)) - \ell(\delta^+(S)) - b(S^-) + \ell(\delta^+(S^-)) - \ell(\delta^-(S^-)) \\ + D - b(S^+) + \ell(\delta^+(S^+)) - \ell(\delta^-(S^+)). \end{aligned}$$

We observe that $\ell(\delta^+(S^-)) - \ell(\delta^-(S^-)) + \ell(\delta^+(S^+)) - \ell(\delta^-(S^+)) = \ell(\delta^+(S)) - \ell(\delta^-(S))$. Since $b(S^+) + b(S^-) = b(S)$, this leads to

$$u'(\delta^+(S')) = u(\delta^+(S)) - \ell(\delta^-(S)) + D - b(S).$$

Therefore, $u'(\delta^+(S')) < D$ if and only if $u(\delta^+(S)) - \ell(\delta^-(S)) < b(S)$, which concludes the proof. □

We have thus shown that we can find GH-cuts with a max-flow algorithm:

Corollary 12. *An IIS for a network flow problem can be computed in the time needed to compute a maximum flow.*

In Proposition 10 we have seen that every (disconnected) GH-cut yields an IIS for *at least one* connected component. It turns out that *every* connected component yields an IIS if the GH-cut is computed using a max-flow algorithm.

Theorem 13. *Let x be a maximum (s-t)-flow in G' with value $< D$, and let G'_x be its residual graph. Define $S := \{v \in V \mid v \text{ reachable from } s \text{ in } G'_x\}$, and suppose that $G[S]$ has k weakly connected components $S_1, \ldots, S_k$. Then every $\mathcal{I}(S_1)$, ..., $\mathcal{I}(S_k)$ is an IIS.*

Proof. We first observe that the demand form of the GH-inequalities is irrelevant in this setting: By construction of G' and S_i, we have $d(S_i) > 0$. Thus, for every $i \in [k] := \{1, \ldots, k\}$,

$$0 < d(S_i) = b(S_i) - \ell(\delta^+(S_i)) + \ell(\delta^-(S_i)) \leq b(S_i) - \ell(\delta^+(S_i)) + u(\delta^-(S_i)),$$

i.e., the demand form (see Section 2) cannot be violated.

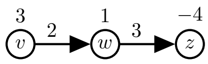
Consider the arcs $a = (v, w) \in \delta^+(S_i)$ in G. Since S_i is a connected component, $w \notin S$. Thus, $x_a = u_a$, i.e., $x(\delta^+(S_i)) = u(\delta^+(S_i))$. Similarly, for arcs $a = (v, w) \in \delta^-(S_i)$, $v \notin S$. If $a \in [(V \setminus S) : S_i]$, we have $x_a = 0$. Moreover, since S_i is reachable from s in G'_x, there must be unsaturated arcs (s, w) for $w \in S_i$, i.e., $x(\delta^-(S_i)) < d(S_i)$. We obtain

$$0 = x(\delta^-(S_i)) - x(\delta^+(S_i)) = x(\delta^-(S_i)) - u(\delta^+(S_i)) < d(S_i) - u(\delta^+(S_i)).$$

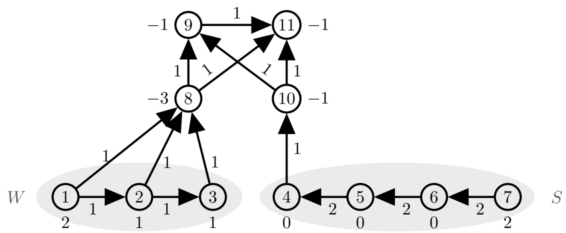
This yields

$$\begin{aligned} & 0 < b(S_i) - \ell(\delta^+(S_i)) + \ell(\delta^-(S_i)) - u(\delta^+(S_i)) \\ \Rightarrow \quad & u(\delta^+(S_i)) - \ell(\delta^-(S_i)) < b(S_i) - \ell(\delta^+(S_i)) \leq b(S_i). \end{aligned}$$

Hence, $\mathcal{I}(S_i)$ is infeasible, and, since S_i is connected, an IIS. □



(a) $\mathcal{I}(\{v,w\})$ is an IIS, but $\{v,w\}$ is not a minimal witness



(b) minimum IIS $\mathcal{I}(S)$ and minimum witness W of the same network problem, with $|S| = 4$, $|W| = 3$, $|\mathcal{I}(S)| = 5$, and $|\mathcal{I}(W)| = 6$

Figure 2: Examples for the difference between witnesses and IISs; node labels specify b, arc labels give u; all lower bounds ℓ are 0.

Note that Theorem 13 does not rely on a particular max-flow algorithm, but holds for any of them, since we only use arguments in the residual graph.

4 A Comparison of IISs and Witnesses

IISs in flow networks and the witness concept share certain similarities, especially since they are both intended to highlight a smaller portion of the network exposing the infeasibility, and rely on the GH-inequalities. The difference is that in the witness problem, one minimizes the number of nodes (in minimum or minimal meaning), while IISs minimize both nodes and arcs.

Note that a witness conforms with our notation of a GH-set.

Lemma 14. *Every minimal witness is connected, but not every connected GH-set is a minimal witness.*

Proof. The first part follows from the proof of Proposition 10. An example for the second statement is given in Figure 2(a). □

Thus, while connectedness of the GH-set is necessary and sufficient for IISs, it is necessary, but not sufficient for minimal witnesses. In the following proposition, we will summarize their connection.

Proposition 15.

1. *$\mathcal{I}(W)$ is infeasible for every witness W, but $S(\mathcal{I})$ is not necessarily a witness for every infeasible $\mathcal{I}$.*
2. *$\mathcal{I}(W)$ is an IIS for every minimal witness W, but $S(\mathcal{I})$ is not necessarily a minimal witness for an IIS $\mathcal{I}$.*
3. *For a minimum witness W, $\mathcal{I}(W)$ is not necessarily a minimum IIS and for a minimum IIS $\mathcal{I}$, $S(\mathcal{I})$ is not necessarily a minimum witness.*

Proof.

1. Since a witness W is a GH-set, $\mathcal{I}(W)$ is a GH-subsystem per definition and therefore infeasible. Conversely, an infeasible $\mathcal{I}$ can contain an infeasible GH-subsystem and arbitrarily more node constraints, whence $b(S(\mathcal{I})) \leq u(\delta^+(S(\mathcal{I}))) - \ell(\delta^-(S(\mathcal{I})))$ might hold.
2. Follows from Lemma 14.
3. An example in which the minimum IIS and the minimum witness are different is given in Figure 2(b). □

Algorithmically, IISs and minimal witnesses have a further difference: While an IIS can be computed with any max-flow algorithm (unmodified), the minimal witness computation is only known to work with preflow-push algorithms with certain adaptions in the labeling procedure (see [1] for details).

5 Minimum IIS in Flow Networks

Greenberg [14] pointed out that, in a practical application, the chance to understand the cause for the infeasibility is increased if an IIS is small. In our setting, it is thus interesting to ask for a minimum cardinality IIS (minIIS). For linear inequality systems, this problem was shown to be strongly $\mathcal{NP}$-hard in [3]. We will extend this result to also hold for the special case of network flow systems.

First, notice that minimum IISs for flow networks can be characterized as follows:

Corollary 16. *Every minimum IIS of the network flow system* (F) *is given by a GH-subsystem $\mathcal{I}(S)$, where S is an optimal solution of*

$$\min_{S\subseteq V} \; |S| + |\delta(S)| \tag{4}$$
$$\text{s.t. } b(S) > u(\delta^+(S)) - \ell(\delta^-(S)) \quad \text{or} \quad -b(S) > u(\delta^-(S)) - \ell(\delta^+(S)).$$

Proof. By Theorem 8, a minimum IIS is the GH-subsystem of a connected GH-cut with a minimum number of nodes plus arcs. Furthermore, any S optimal for (4) is necessarily connected: Suppose there exists a proper connected component $T \subset S$ such that $\delta(T) \cap \delta(S \setminus T) = \emptyset$, which, by Proposition 10, induces an IIS $\mathcal{I}(T)$. Then $|\mathcal{I}(T)| = |\delta(T)| + |T| < |\delta(S)| + |S| = |\mathcal{I}(S)|$, contradicting the optimality of S. □

We can use this characterization for a reduction from the maximum clique problem on regular graphs (more precisely, the respective decision problem). Aggarwal et al. [1] showed that the clique problem remains strongly $\mathcal{NP}$-hard when restricted to regular graphs, by the observation that it is equivalent to the independent set problem on the complement graph, and independent set is $\mathcal{NP}$-hard even on planar cubic graphs.

Theorem 17. *Given a network flow problem and a positive integer $\tilde{k}$, it is $\mathcal{NP}$-complete in the strong sense to decide whether an IIS of size at most $\tilde{k}$ exists.*

Proof. Note first that the problem is in $\mathcal{NP}$ by Theorem 8: We can check in polynomial time whether a given subsystem $\mathcal{I}$ has size at most $\tilde{k}$, whether $\mathcal{I}$ has the form of a GH-subsystem, and whether the induced graph of $S(\mathcal{I})$ is connected.

We reduce the strongly $\mathcal{NP}$-complete regular maximum clique problem: Given an r-regular, undirected graph $G' = (V', E')$, with $|V'| := n$, $|E'| := m$, and a positive integer k, does there exist a clique $C \subseteq V'$ (i.e., $G'[C]$ is a complete graph) such that $|C| \geq k$? For the reduction, we will construct a network flow problem instance $(V,\, A,\, b,\, u,\, \ell)$ that has an IIS $\mathcal{I}$ of size at most $\tilde{k} := k + 2k(r-k+1) + kn^2$ if and only if G' has a clique of size k. Note that $k \leq r+1 \leq n$.

First, we take all nodes in V' and assign a supply of $r-k+3$ to each, and we replace every edge in E' by a pair of oppositely directed arcs with an upper bound of 1 and a lower bound of 0. We need an additional n^4 copies of V', yielding a set of intermediate nodes V_a, each copy connected to the next one by forward arcs A_3. The first n^2 nodes in V_a are connected to the corresponding node in V' by arcs A_2. Moreover, we add a sink node t connected to the last copy of V'; see Figure 3 for an illustration of the construction.

The resulting directed graph $G = (V, A)$ is formally defined by its arc set $A := A_1 \cup A_2 \cup A_3$ and node set $V := V' \cup (V' \times [n^4]) \cup \{t\}$ ($[s] := \{1, \dots, s\}$), where we write $(v \times j)$ for the nodes in $V' \times [n^4]$, with

$$\begin{aligned}
A_1 &:= \{(v,w) \mid \{v,w\} \in E'\} \cup \{(w,v) \mid \{v,w\} \in E'\},\\
A_2 &:= \{\big(v,\, (v \times j)\big) \mid v \in V',\; j \in [n^2]\},\\
A_3 &:= \{\big((v \times j),\, (v \times j+1)\big) \mid v \in V',\; j \in [n^4-1]\} \cup \{\big((v \times n^4),\, t\big) \mid v \in V'\}.
\end{aligned}$$

Furthermore, let $\ell := 0$ and

$$b_v := \begin{cases} r-k+3, & v \in V',\\ -n(r-k+3), & v = t,\\ 0, & \text{otherwise,} \end{cases} \qquad u_a := \begin{cases} 1, & a \in A_1,\\ 1/n^2, & a \in A_2,\\ n(r-k+3), & a \in A_3. \end{cases}$$

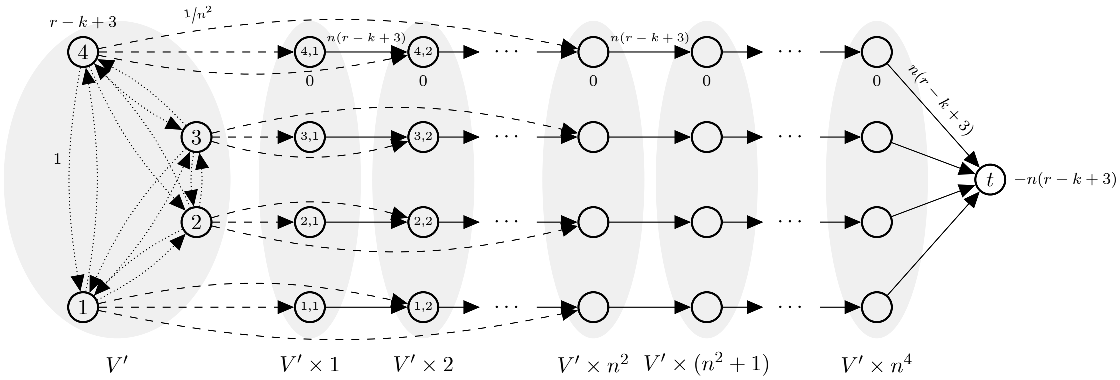


Figure 3: Sketch of the construction for the reduction; $\ell_a = 0$, $u_a = n(r-k+3)$ for solid, 1 for dotted, and $1/n^2$ for dashed arcs a.

A clique C with $|C| = k$ in an r-regular graph has $|\delta(C)| = k(r-k+1)$ (every clique node has r incident edges, from which $k-1$ are connected to nodes inside the clique). Hereby, the "if"-direction is easy: Consider a clique $C \subseteq V'$ in the constructed graph G. Then

$$b(C) = k(r-k+3) > k(r-k+1) + kn^2\frac{1}{n^2} - 0 = u(\delta^+(C)) - \ell(\delta^-(C)).$$

Thus, since $G[C]$ is connected, $\mathcal{I}(C)$ is an IIS by Theorem 8. It has size $\tilde{k} = k + 2k(r-k+1) + kn^2$, since we have k nodes, $k(r-k+1)$ out- and ingoing arcs, respectively, and for every node, there are n^2 arcs in A_2.

For the converse, suppose there exits an IIS $\mathcal{I}$ with $|\mathcal{I}| \leq \tilde{k}$. First, assume that $t \in S := S(\mathcal{I})$, which is the only possibility for the demand case. Then $V' \times \{j\} \subset S$ for all $j \in [n^4]$, since otherwise, there would exist $a \in \delta^-(S) \cap A_3$ with $u_a = n(r-k+3) \geq -b_t$, rendering the subsystem feasible (by Theorem 8, an IIS must yield a violated GH-inequality). Then $|\mathcal{I}| > n^5 + 1 \geq (n+1)^4/8 \geq \tilde{k}$, as can be easily verified (in particular, the last inequality becomes apparent from bounding $\tilde{k}$ from above using $r+1 \leq n$ and maximizing the resulting expression w.r.t. $k \leq n$); this is a contradiction. Consequently, $t \notin S$, and only the supply form can occur.

Moreover, if there exists a node $(v \times j) \in S$ for $v \in V'$ and $j \in [n^4]$, we would necessarily have $u(\delta^+(S)) \geq n(r-k+3) \geq b(S)$. It follows that $S \subseteq V'$.

Let $k' := |S|$. If $k' > k$, then

$$\begin{aligned}|\mathcal{I}| &> k'\,(n^2+1) = (k'-k)(n^2+1) + k(n^2+1) > n^2 + 1 + k(n^2+1)\\ &> 2k(r-k+1) + k(n^2+1) = \tilde{k},\end{aligned}$$

since $2k(r-k+1) \leq 2k(n-1-k+1) \leq n^2$, which can be seen by some easy calculations. This is again a contradiction, so we conclude that $k' \leq k$. Suppose that $k' < k$. Then,

$$u(\delta^+(S)) \geq k'\frac{n^2}{n^2} + k'(r-k'+1) = k'(r-k'+2) \geq k'(r-k+3) = b(S),$$

which contradicts the infeasibility of $\mathcal{I}$.

Eventually, consider $k' = k$, and suppose that no clique of size at least k exists in G'. Hence, there must be at least one more pair of oppositely directed arcs in $\delta(S)$ than for a clique. Consequently,

$$|\mathcal{I}| = k + |\delta(S)| \geq k + 2k(r-k+1) + 2 + kn^2 > \tilde{k},$$

again a contradiction. In conclusion, a clique of size at least k has to exist.

Finally note that the encoding length of the resulting instance is clearly polynomial in that of the graph G'. In fact, all numbers occurring in the instance are bounded by a polynomial in n, which proves strong $\mathcal{NP}$-completeness. □

Corollary 18. *Computing a minIIS in flow networks is strongly $\mathcal{NP}$-hard.*

Remark 19. Aggarwal et al. [1] used the regular clique problem to show $\mathcal{NP}$-hardness of the minimum witness problem; the above proof relies on their idea and uses the fact that the number of arcs leaving a clique is constant in a regular graph for each fixed clique size. Our auxiliary graph, however, needs to be much larger than theirs, since we cover the demand form, while the witness problem is conveniently defined only w.r.t. subsets of a GH-cut with a positive supply.

One can also consider a weighted version of minIIS, where we minimize the product of a weight-vector with the incidence vector of constraints in the IIS. An application of this problem is found in the computation of *maximum feasible subsystems* (maxFS), i.e., the largest number of constraints of the system with a solution. The maxFS problem on a system with r constraints can be formulated as

$$\begin{aligned} \max\ & \sum_{i=1}^{r} y_i \\ & \sum_{i\in\mathcal{I}} y_i \leq |\mathcal{I}| - 1 \quad \forall \mathcal{I} \in I \\ & y \in \{0,1\}^r, \end{aligned} \tag{5}$$

where I is the set of all IISs of the system. Amaldi et al. [3] showed that the separation over (5), also called *IIS-inequalities*, is $\mathcal{NP}$-hard in the general case of linear systems. The proof works via a reduction from minIIS, where an IIS of size at most k is linked to the separation of a vector $y = (1 - \frac{1}{k+1})\,\mathbb{1}$. If we replace the hardness result for minIIS with Theorem 17, this proof works without further changes for the flow case, whence we have the following result:

Corollary 20. *The separation problem for IIS-inequalities on flow networks is $\mathcal{NP}$-hard.*

6 Outlook

The results of this paper characterize IISs in terms of connected Gale-Hoffman-inequalities. This can possibly be used for general mathematical programs that contain flows as a substructure. In fact, one motivation for this article was the analysis of infeasible systems arising in stationary gas transportation, where the systems are nonlinear, nonconvex and can contain discrete variables [21]. For instance, in such general systems on networks, the components corresponding to IISs are necessarily connected [19]. Other generalizations will be the topic of future research; a first topic might be multicommodity flows using [29].

The infeasibility characterizations developed in this paper might also be relevant for network reliability [24] and survivable network design [17].

Another open issue is to obtain inapproximability results for determining a minimum IIS. For the general case, strong inapproximability results exist [3]. Moreover, the related problem to compute a cover of IISs of smallest cardinality is interesting as well [19, 20].

Acknowledgement

We thank the anonymous referees for their suggestions and the editor for his careful handling of this paper.

References

[1] C.C. Aggarwal, R.K. Ahuja, J. Hao, and J.B. Orlin, Diagnosing infeasibilities in network flow problems, Math. Program. 81 (1998), 263–280.

[2] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, Network flows: theory, algorithms, and applications, Prentice Hall, 1993.

[3] E. Amaldi, M.E. Pfetsch, and L.E. Trotter Jr, On the maximum feasible subsystem problem, IISs and IIS-hypergraphs, Math. Program. 95 (2003), 533–554.

[4] J.W. Chinneck, An effective polynomial-time heuristic for the minimum-cardinality IIS set-covering problem, Ann. Math. Artif. Intell. 17 (1996), 127–144.

[5] J.W. Chinneck, Localizing and diagnosing infeasibilities in networks, INFORMS J. Comput. 8 (1996), 55–62.

[6] J.W. Chinneck, "Feasibility and viability," Advances in sensitivity analysis and parametric programming, T. Gál and H.J. Greenberg (Editors), Kluwer Academic Publishers, Dordrecht, 1997, Vol. 6 of Int. Ser. Oper. Res. Manag. Sci., Chapter 14, pp. 1–41.

[7] J.W. Chinneck, Finding a useful subset of constraints for analysis in an infeasible linear program, INFORMS J. Comput. 9 (1997), 164–174.

[8] J.W. Chinneck, Fast heuristics for the maximum feasible subsystem problem, INFORMS J. Comput. 13 (2001), 210–223.

[9] J.W. Chinneck, Feasibility and infeasibility in optimization: algorithms and computational methods Vol. 118 of *International Series in Operations Research and Management Sciences*, Springer, 2008.

[10] J.W. Chinneck and E.W. Dravnieks, Locating minimal infeasible constraint sets in linear programs, ORSA J. Comput. 3 (1991), 157–168.

[11] D. Gale, A theorem on flows in networks., Pacific J. Math. 7 (1957), 1073–1082.

[12] J. Gleeson and J. Ryan, Identifying minimally infeasible subsystems of inequalities, ORSA J. Comput. 2 (1990), 61–63.

[13] H.J. Greenberg, Diagnosing infeasibility in min-cost network flow problems part I: Dual infeasibility, IMA J. Manag. Math. 1 (1987), 99–109.

[14] H.J. Greenberg, Diagnosing infeasibility in min-cost network flow problems part II: Primal infeasibility, IMA J. Manag. Math. 2 (1988), 39–50.

[15] H.J. Greenberg, Consistency, redundancy, and implied equalities in linear systems, Ann. Math. Artif. Intell. 17 (1996), 37–83.

[16] H.J. Greenberg and F.H. Murphy, Approaches to diagnosing infeasible linear programs, ORSA J. Comput. 3 (1991), 253–261.

[17] M. Grötschel, C.L. Monma, and M. Stoer, Polyhedral and computational investigations for designing communication networks with high survivability requirements, Oper. Res. 43 (1995), 1012–1024.

[18] A.J. Hoffman, Some recent applications of the theory of linear inequalities to extremal combinatorial analysis, Proc. Sympos. Appl. Math, Vol. 10, World Scientific 1960, pp. 113–127.

[19] I. Joormann, Analyzing infeasible flow networks, Dr. Hut Verlag, 2015, Ph.D. Thesis, TU Darmstadt.

[20] I. Joormann and M.E. Pfetsch, Complexity of minimum irreducible infeasible subsystem covers for flow networks, Technical report, Optimization Online, 2015, `http://www.optimization-online.org/DB_HTML/2015/03/4809.html`.

[21] T. Koch, B. Hiller, M.E. Pfetsch, and L. Schewe (Editors), Evaluating gas network capacities, SIAM-MOS series on Optimization, SIAM, 2015.

[22] S.T. McCormick, How to compute least infeasible flows, Math. Progam. 78 (1997), 179–194.

[23] M.E. Pfetsch, Branch-and-cut for the maximum feasible subsystem problem, SIAM J. Optim. 19 (2008), 21–38.

[24] J.S. Provan, Polyhedral combinatorics and network reliability, Math. Oper. Res. 11 (1986), 36–61.

[25] J. Ryan, IIS-hypergraphs, SIAM J. Discr Math. 9 (1996), 643–653.

[26] E. Tardos, A strongly polynomial algorithm to solve combinatorial linear programs, Oper. Res. 34 (1986), 250–256.

[27] J. Van Loon, Irreducibly inconsistent systems of linear inequalities, Eur. J. Oper. Res. 8 (1981), 283–288.

[28] S.W. Wallace and R.J.B. Wets, The facets of the polyhedral set determined by the Gale-Hoffman inequalities, Math. Program. 62 (1993), 215–222.

[29] H.S. Zullo, Feasible flows in multicommodity graphs, Ph.D. Thesis, University of Colorado at Denver, 1995.