

Robust Metric Inequalities for the Γ -Robust Network Loading Problem

Grit Claßen^{*} Arie M.C.A. Koster^{*} Manuel Kutschka^{*} Issam Tahiri[†]

Abstract

In this paper, we consider the network loading problem under demand uncertainties with static routing, i.e., a single routing scheme based on the fraction of the demands has to be determined. We generalize the class of metric inequalities to the Γ -robust setting and show that they yield a formulation in the capacity space. We describe a polynomial time exact algorithm to separate violated robust metric inequalities as model constraints. Moreover, rounded and tight robust metric inequalities describing the convex hull of integer solutions are presented and separated in a cut-and-branch approach. Computational results using real-life telecommunication data demonstrate the major potential of (tight) robust metric inequalities by considering the integrality gaps at the root node and the overall optimality gaps. Speed-up factors between 2 and 5 for the compact flow and between 3 and 25 for the capacity formulation have been achieved by exploiting robust metric inequalities in the solving process.

Keywords: metric inequalities, network loading, robust optimization, capacity formulation

1 Introduction

In a variety of applications (telecommunications, railways, logistics), a network has to be designed by means of installing capacities on the edges/links of a potential network topology. A solution to the *Network Design Problem* (NDP) is feasible if the constructed network is actually capable of (simultaneously) transporting the entire traffic requirements (demands) between all node pairs. The goal is to select among all feasible solutions one that minimizes the *network cost*: the sum of link capacity installation costs. The special case of the network design problem when only integer capacities are available is called the *Network Loading Problem* (NLP). This problem has been investigated intensely in the literature, see [2, 3, 4, 5, 12, 14, 15] to name just some references. Different specifications of the NLP such as static (a single traffic matrix is considered) or dynamic (several traffic matrices are considered) routing, splittable or unsplittable flows can be considered. In this paper, we focus on a variant of static routing and splittable flows.

The NLP is NP-hard which raised the question of finding valid inequalities to improve the bounds. Hence, a branch-and-cut approach with problem specific valid inequalities is essential to achieve optimal solutions. Therefore, many types of valid inequalities have been investigated. The first well-known valid inequalities are the cut inequalities; they ensure that every cut is covered by sufficient

^{*}RWTH Aachen University, Lehrstuhl II für Mathematik, Wüllnerstr. 5b, D-52062 Aachen, Germany, {classen, koster, kutschka}@math2.rwth-aachen.de

[†]INRIA - 2004, route des Lucioles - B.P. 93, F-06902 Sophia.A, France, issam.tahiri@inria.fr

capacity. Multicut inequalities have been presented by Barahona [5] and aim at ensuring the connectivity of the network. Arc-residual capacity inequalities and 3-partition inequalities have been introduced by Magnanti et al. [14]. Extending this work, Agarwal [3] studied the 4-partition inequalities; he concluded that the development of efficient algorithms/heuristics to separate these valid inequalities has to be investigated to fully exploit their potential. A further extension of partition inequalities are the so-called mixed partition inequalities introduced by Günlük [8]. Finally, an important class of valid inequalities are the *metric inequalities* since they do not only generalize other valid inequalities such as cut inequalities and partition inequalities but they also describe the network design polytope in the capacity space completely [4] providing an alternative formulation for the NLP; we will present them briefly in Section 2.

As traffic (bandwidth) requirements are non-deterministic in real world applications, the consideration of a single traffic matrix is in general too restrictive or leads to problems such as delays during the runtime of the network. One possibility to take uncertain traffic requirements into account is to examine dynamic routing allowing for several traffic matrices. This approach can be regarded as a two stage concept where the capacities are computed in the first stage and the routing is determined in a second stage based on the traffic matrix, see e.g. Mattia [15]. On the other hand, a static routing, one stage concept such as we are considering in this paper, which takes several traffic matrices into account, is more practical. A NLP that considers uncertainties is also known as the *Robust Network Loading Problem* (RNLP). Recently, Koster et al. [12] have studied the RNLP deriving cutset-based valid inequalities.

Our contribution In this paper, we consider the Γ -robust NLP with static routing and splittable flows. We first generalize the class of metric inequalities to *robust metric inequalities* (RMIs). Analogue to the deterministic case, we show that these inequalities provide a formulation for the Γ -robust network loading problem in the capacity space. Further, we prove that the tight RMIs together with nonnegativity completely describe the convex hull of all integer capacity vectors. We present three separation routines: A polynomial time exact algorithm to separate RMIs as model constraints, an exact integer program for rounded RMIs, and a heuristic algorithm to separate tight RMIs. Using test instances based on real-life telecommunication data, we analyze the performance of RMIs in both the compact flow and the capacity formulation. To this end, rounded and tight RMIs as well as robust cutset inequalities are separated at the root node. Our computational results demonstrate that the integrality gap can be closed significantly yielding an overall speed-up, in particular, for the capacity formulation.

The remainder of this paper is organized as follows. In Section 2, we present the model of the classical NLP and give the state of the art of metric inequalities. In Section 3, we incorporate traffic demand uncertainties via the Γ -robustness approach to obtain the RNLP. Further, we introduce the RMIs and show that every non-dominated valid inequality of the robust network loading polytope and its relaxation is a RMI (or a nonnegativity constraint). Afterwards, we investigate the separation of RMIs as model inequalities and the separation of rounded and of tight RMIs in a cut-and-branch approach in Section 4. In Section 5, we study the performance of the RMIs and compare the capacity formulation to the compact flow formulation presented in Koster et al. [12] by computational results obtained for realistic communication network topologies. We conclude this paper by Section 6.

2 The Network Loading Problem

Let $G = (V, E)$ be a simple undirected graph. A set of commodities (traffic requests) K have to be routed between the nodes of G . Every commodity consists of an ordered source-target pair (s^k, t^k) and a demand value $d^k \in \mathbb{R}_+$. We denote the set of simple paths in G between s^k and t^k by P^k and their union by P , i.e., $P = \cup_{k \in K} P^k$.

The NDP consists of installing capacities on edges of G and finding a feasible routing for all demands respecting those capacities. In the NLP, only integer capacities can be installed on the edges with a unit cost κ_e for $e \in E$. The number of capacity units installed at edge $e \in E$ is variable and denoted by $x_e \in \mathbb{Z}_+$.

Given G and K , a routing can be described by means of fractional flow variables. For a commodity $k \in K$ and for a path $p \in P^k$, $f^k(p)$ represents the fraction of demand value d^k routed along path p , i.e., $f^k : P^k \rightarrow [0, 1]$. The set of these variables constitutes a multi(-commodity) flow. A formulation of the classical NLP is the following.

$$\min \sum_{e \in E} \kappa_e x_e \tag{1a}$$

$$\text{s.t. } \sum_{p \in P^k} f^k(p) \geq 1 \quad \forall k \in K \tag{1b}$$

$$\sum_{k \in K} d^k \left(\sum_{p \in P^k: e \in p} f^k(p) \right) \leq x_e \quad \forall e \in E \tag{1c}$$

$$f^k(p) \in [0, 1] \quad \forall k \in K, p \in P^k \tag{1d}$$

$$x_e \in \mathbb{Z}_+ \quad \forall e \in E \tag{1e}$$

We refer to the convex hull of a set X by $\text{conv}\{X\}$. Let

$$\tilde{P}(G, d) = \text{conv} \left\{ (f, x) \in [0, 1]^{|P|} \times \mathbb{Z}_+^{|E|} : (f, x) \text{ is feasible for (1b) and (1c)} \right\}$$

be the *Network Loading Polytope in the flow and capacity space* and

$$P(G, d) = \text{conv} \left\{ x \in \mathbb{Z}_+^{|E|} : \exists f \in [0, 1]^{|P|} \text{ so that } (f, x) \in \tilde{P}(G, d) \right\}$$

be the *Network Loading Polytope in the capacity space*.

Classical metric inequalities Let the function $\ell : E \rightarrow \mathbb{R}_+$ define (artificial) lengths on the edges. For simplicity, we denote the length of a path $p \in P$ as $\ell(p) = \sum_{e \in p} \ell(e)$. Further, for every $u, v \in V$ we denote the length of the shortest path in G regarding ℓ by $\ell(u, v)$. Notice that $\ell(s^k, t^k) = \min_{p \in P^k} \ell(p)$. Function ℓ defines a *metric* if $\ell(uv) \leq \ell(u, v)$ for every edge $uv \in E$. We use $\text{Met}(G)$ to refer to the cone defined by all metrics on G .

In Iri [10], Lomonosov [13], Onaga and Kakusho [17], the feasibility of multiflows was studied stating equivalent conditions based on (metric) length functions. In fact, considering metric length functions is sufficient. This results in a reformulation of the NLP in the capacity space by metric inequalities and integrality constraints only. In the following, we state these key results more formally.

Theorem 2.1 (Iri [10], Lomonosov [13], Onaga and Kakusho [17]). *The vector $x \in \mathbb{Z}_+^{|E|}$ is in $P(G, d)$ if and only if for all length functions $\ell : E \rightarrow \mathbb{R}_+$ holds*

$$\sum_{e \in E} \ell(e) x_e \geq \sum_{k \in K} d^k \ell(s^k, t^k). \quad (2)$$

The necessary condition for the existence of a feasible multiflow can be restricted to the cone of metric length functions.

Theorem 2.2 (Iri [10], Lomonosov [13], Onaga and Kakusho [17]). *The vector $x \in \mathbb{Z}_+^{|E|}$ is in $P(G, d)$ if and only if for all metric functions $\ell_M : E \rightarrow \mathbb{R}_+$ in $\text{Met}(G)$ holds*

$$\sum_{e \in E} \ell_M(e) x_e \geq \sum_{k \in K} d^k \ell_M(s^k, t^k). \quad (3)$$

Note, we may assume that length functions have only fractional values, or even integer values. By Theorem 2.2, the NLP can be formulated in the capacity space by using only the metric inequalities and integrality constraints.

Corollary 2.3 (Avella et al. [4]).

$$P(G, d) = \text{conv} \left\{ x \in \mathbb{Z}_+^{|E|} : x \text{ satisfies (3) for all } \ell_M \in \text{Met}(G) \right\}$$

3 The Γ -Robust Network Loading Problem

In practice, the demand values considered in the NLP (1) are fluctuating over time. To tackle these uncertainties, we apply the robust optimization approach by Bertsimas and Sim [6], see also [12]. Hence, the demand values are now modelled as symmetric and bounded random variables taking values in $[\bar{d}^k - \hat{d}^k, \bar{d}^k + \hat{d}^k]$, where \bar{d}^k denotes the nominal value and \hat{d}^k its highest deviation. The peak demand is thus denoted by $\bar{d}^k + \hat{d}^k$. We introduce a robustness parameter $\Gamma \in \{0, \dots, |K|\}$ which limits the number of commodities whose demands deviate from their nominal demands simultaneously. For the formulation of the Γ -robust NLP based on [6], the capacity constraints (1c) are replaced by

$$\sum_{k \in K} \bar{d}^k \left(\sum_{p \in P^k: e \in p} f^k(p) \right) + \max_{Q \subseteq K, |Q| \leq \Gamma} \sum_{k \in Q} \hat{d}^k \left(\sum_{p \in P^k: e \in p} f^k(p) \right) \leq x_e \quad \forall e \in E. \quad (4)$$

Due to the maximum, these constraints are non-linear. An intuitive linearization is to compute all possible subsets $Q \subseteq K$. However, there exist exponentially many of such subsets. Computations have shown that the exponential-sized formulation is outperformed by the following approach, see [12]. We reformulate the maximum as an ILP. Since the resulting matrix is totally unimodular, we can exploit

LP duality introducing dual variables π_e and ρ_e^k . Hence, the Γ -robust NLP can be described by the following compact ILP formulation.

$$\min \sum_{e \in E} \kappa_e x_e \quad (5a)$$

$$\text{s.t. } \sum_{p \in P^k} f^k(p) \geq 1 \quad \forall k \in K \quad (5b)$$

$$\sum_{k \in K} \bar{d}^k \left(\sum_{p \in P^k: e \in p} f^k(p) \right) + \Gamma \pi_e + \sum_{k \in K} \rho_e^k \leq x_e \quad \forall e \in E \quad (5c)$$

$$\rho_e^k + \pi_e \geq \hat{d}^k \sum_{p \in P^k: e \in p} f^k(p) \quad \forall e \in E, k \in K \quad (5d)$$

$$\pi_e, \rho_e^k \geq 0 \quad \forall e \in E, k \in K \quad (5e)$$

$$f^k(p) \in [0, 1] \quad \forall k \in K, p \in P^k \quad (5f)$$

$$x_e \in \mathbb{Z}_+ \quad \forall e \in E \quad (5g)$$

Constraints (5b)-(5d) together with (5e) and (5f) define a feasible robust multiflow.

Similar to $P(G, d)$ and $\tilde{P}(G, d)$, let

$$\tilde{P}_\Gamma(G, \bar{d}, \hat{d}) = \text{conv} \left\{ (f, x, \pi, \rho) \in [0, 1]^{|P|} \times \mathbb{Z}_+^{|E|} \times \mathbb{R}_+^{|E|} \times \mathbb{R}_+^{|E||K|} : (f, x, \pi, \rho) \text{ is feasible for (5b)-(5d)} \right\}$$

be the *Robust Network Loading Polytope in the flow, capacity and dual variables* and let

$$P_\Gamma(G, \bar{d}, \hat{d}) = \text{conv} \left\{ x \in \mathbb{Z}_+^{|E|} : \exists (f, \pi, \rho) \in [0, 1]^{|P|} \times \mathbb{R}_+^{|E|} \times \mathbb{R}_+^{|E||K|} \text{ so that } (f, x, \pi, \rho) \in \tilde{P}_\Gamma(G, \bar{d}, \hat{d}) \right\}$$

be the *Robust Network Loading Polytope in the capacity space*.

In contrast to Corollary 2.3, $P_\Gamma(G, \bar{d}, \hat{d})$ cannot be completely described by a straightforward generalization of the metric inequalities. First, consider the complete graph G with nodes $V := \{1, 2, 3\}$, edges $E := \{12, 13, 23\}$, commodities K (with $|K| = 3$), demand values $\bar{d}^{12} = \bar{d}^{13} = \bar{d}^{23} = 2$ and $\hat{d}^{12} = 1$, $\hat{d}^{13} = 4$, $\hat{d}^{23} = 2$, and $\Gamma = 1$. The dominating metric inequalities for the classical NLP (considering only nominal demands) on the complete graph with three nodes are

$$x_{12} + x_{13} \geq 4 \quad (6a)$$

$$x_{13} + x_{23} \geq 4 \quad (6b)$$

$$x_{12} + x_{23} \geq 4 \quad (6c)$$

$$x_{12} + x_{13} + x_{23} \geq 6, \quad (6d)$$

describing $P(G, d)$ together with the nonnegativity constraints completely.

Based on the robust counterpart of the capacity constraints (4), an intuitive way to extend these metric inequalities in the robust case is to increase the right hand side as follows

$$\sum_{e \in E} \ell_M(e) x_e \geq \sum_{k \in K} \bar{d}^k \ell_M(s^k, t^k) + \max_{Q \subseteq K, |Q| \leq \Gamma} \sum_{k \in Q} \hat{d}^k \ell_M(s^k, t^k).$$

Then, we get the right hand sides 8, 8, 6, and 10. However, for the corresponding RNLP the complete description is

$$x_{12} + x_{13} \geq 8 \quad (7a)$$

$$x_{13} + x_{23} \geq 8 \quad (7b)$$

$$x_{12} + x_{23} \geq 6 \quad (7c)$$

$$x_{12} + x_{13} + x_{23} \geq 13 \quad (7d)$$

$$2x_{12} + 2x_{13} + 3x_{23} \geq 28 \quad (7e)$$

$$4x_{12} + 3x_{13} + 3x_{23} \geq 42 \quad (7f)$$

$$6x_{12} + 4x_{13} + 5x_{23} \geq 62. \quad (7g)$$

Thus, we see that not only more inequalities are needed than for the classical NLP but, in particular, that the straightforward way of generalizing them yields inequalities with too weak right hand sides (13 vs. 10 in (7d), 28 vs. 22 in (7e), 42 vs. 32 in (7f), and 62 vs. 46 in (7g)). In fact, the right hand side of a RMI can be determined by the following lemma.

Lemma 3.1. *Given $\tilde{x} \in \mathbb{Z}_+^{|E|}$, there exists a flow satisfying (5b)–(5f) if and only if for all lengths functions ℓ*

$$\sum_{e \in E} \tilde{x}_e \ell(e) \geq b_\ell \quad (8)$$

holds, where b_ℓ is defined by the following LP.

$$b_\ell := \max \sum_{k \in K} b^k \quad (9a)$$

$$s.t. \ b^k - \sum_{e \in p} \hat{d}^k m^k(e) \leq \bar{d}^k \ell(p) \quad \forall k \in K, p \in P^k \quad (9b)$$

$$\sum_{k \in K} m^k(e) \leq \Gamma \ell(e) \quad \forall e \in E \quad (9c)$$

$$m^k(e) \leq \ell(e) \quad \forall k \in K, e \in E \quad (9d)$$

$$b^k, m^k(e) \geq 0 \quad \forall k \in K, e \in E \quad (9e)$$

We call (8) the Robust Length Inequality (RLI), and if ℓ is metric, (8) is called a RMI.

Proof. By setting $x = \tilde{x}$ in (5), the problem is turned into an LP. We introduce dual variables $\{b^k\}_{k \in K}$, $\{\ell(e)\}_{e \in E}$, $\{m^k(e)\}_{e \in E, k \in K}$ for constraints (5b), (5c) and (5d), respectively, and apply Farkas' lemma yielding (9b)–(9e) and $\ell(e) \geq 0$. Thus, a solution to (5b)–(5f) exists if and only if

$$\sum_{e \in E} \tilde{x}_e \ell(e) \geq \sum_{k \in K} b^k \quad (10)$$

is valid for all b^k , $\ell(e)$, $m^k(e) \geq 0$ satisfying (9b)–(9d). Therefore, for a given length function $\{\tilde{\ell}(e)\}_{e \in E}$ inequality (10) is valid for the linear relaxation of $P_\Gamma(G, \bar{d}, \hat{d})$ if constraints (9b)–(9e) are fulfilled. To determine the strongest valid inequality for $\tilde{\ell}$, we maximize the sum over b^k for $k \in K$. This can be formulated as LP (9), where $\ell = \tilde{\ell}$, completing the proof. \square

The LP (9) to compute the right hand side of a RLI is comparable to the problem (subsep_{\max}) of Mattia [15]. However, we specified the general constraint $Ad \leq b$ according to the compact

formulation of the Γ -robust NLP, which implies further constraints. Compared to (2), b^k is the length of the shortest path with respect to ℓ adjusted by values $m^k(e)$ for which additional constraints (9c) and (9d) hold.

By Lemma 3.1, $P_\Gamma(G, \bar{d}, \hat{d})$ can be described by linear constraints:

$$P_\Gamma(G, \bar{d}, \hat{d}) = \text{conv} \left\{ x \in \mathbb{Z}_+^{|E|} : \sum_{e \in E} \ell(e)x_e \geq b_\ell, \forall \ell \in \mathbb{R}^{|E|} \right\}.$$

Further, let $Q_\Gamma(G, \bar{d}, \hat{d})$ be the linear relaxation of $P_\Gamma(G, \bar{d}, \hat{d})$, i.e.,

$$Q_\Gamma(G, \bar{d}, \hat{d}) = \left\{ x \in \mathbb{R}_+^{|E|} : \sum_{e \in E} \ell(e)x_e \geq b_\ell, \forall \ell \in \mathbb{R}^{|E|} \right\}.$$

In the following theorem, we show that the left hand side of any valid inequality for $Q_\Gamma(G, d)$, which is not dominated by any other valid inequality, defines a metric. We present a constructive proof which is specific to the Γ -robust NLP with static routing and splittable flows. If we proved the theorem analogously to [4], we could not state anything on the right hand side b .

Theorem 3.2. *Let $\bar{\ell}x \geq b_{\bar{\ell}}$ be any valid inequality for $Q_\Gamma(G, d)$. Then there exists a metric $\ell_M \in \text{Met}(G)$ with*

- $\ell_M x \geq b_{\bar{\ell}}$ is valid,
- $\ell_M(ij) \leq \bar{\ell}(ij)$.

Proof. Let $\bar{\ell}$ be a non-metric length function and $b_{\bar{\ell}}$ the corresponding optimal solution of (9). The length function $\bar{\ell}$ is transformed to a metric by setting $\ell_M(ij) := \bar{\ell}(i, j)$ for all $ij \in E$. Since only values are changed that are not already the shortest path value, all changes can be done simultaneously and the resulting function is a metric. To prove that the right hand sides $b_{\bar{\ell}}$ and b_{ℓ_M} are the same, we however assume that the lengths are adapted one at a time. Hence, without loss of generality we assume that there exists only one edge uv with $\bar{\ell}(uv) > \bar{\ell}(u, v)$.

Let $p_1 = (u, v)$ be the direct path from u to v . Since $\bar{\ell}(uv) > \bar{\ell}(u, v)$, there must exist a path $p_2 \neq p_1$ such that $\bar{\ell}(p_2) = \bar{\ell}(u, v)$ and

$$\ell_M(p_1) = \ell_M(p_2) = \ell_M(uv) = \bar{\ell}(u, v).$$

Given a length function ℓ , the dual of (9) reads as follows.

$$\min \sum_{k \in K} \bar{d}^k \sum_{p \in P^k} \ell(p) \mu^k(p) + \Gamma \sum_{e \in E} \nu(e) \ell(e) + \sum_{k \in K} \sum_{e \in E} z^k(e) \ell(e) \quad (11a)$$

$$\text{s.t.} \quad \sum_{p \in P^k} \mu^k(p) \geq 1 \quad \forall k \in K \quad (11b)$$

$$\nu(e) + z^k(e) - \bar{d}^k \ell(e) \sum_{p \in P^k: e \in p} \mu^k(p) \geq 0 \quad \forall k \in K, e \in E \quad (11c)$$

$$\mu^k(p), \nu(e), z^k(e) \geq 0 \quad \forall k \in K, p \in P^k, e \in E, \quad (11d)$$

where $\{\mu^k(p)\}_{k \in K, p \in P^k}$, $\{\nu(e)\}_{e \in E}$ and $\{z^k(e)\}_{k \in K, e \in E}$ are the dual variables for (9b), (9c), and (9d), respectively. Let $(\tilde{\mu}, \tilde{\nu}, \tilde{z})$ be the optimal solution of this dual LP (11) with $\ell = \ell_M$. Hence, $\ell_M x \geq b_{\ell_M}$ is a valid inequality. If we show $b_{\ell_M} \geq b_{\bar{\ell}}$, then $\ell_M x \geq b_{\bar{\ell}}$ is valid and the proof is completed.

Therefore, we construct a feasible as well as optimal solution (μ, ν, z) of the dual LP (11) with $\ell = \ell_M$, prove that (μ, ν, z) is also feasible for the dual LP (11) with $\ell = \bar{\ell}$ and that the objective values coincide. Given $(\tilde{\mu}, \tilde{\nu}, \tilde{z})$, define

$$\nu(e) := \begin{cases} \tilde{\nu}(e) + \tilde{\nu}(uv) & e \in p_2 \\ 0 & e = uv \\ \tilde{\nu}(e) & \text{otherwise,} \end{cases}$$

$$z^k(e) := \begin{cases} \tilde{z}^k(e) + \tilde{z}^k(uv) & e \in p_2 \\ 0 & e = uv \\ \tilde{z}^k(e) & \text{otherwise.} \end{cases}$$

For the definition of μ , we first define two specific subsets of paths $P_{uv}^k := \{p \in P^k : uv \in p\} \subseteq P^k$ and $P_{p_2}^k := \{p \in P^k : p = p_{uv} \setminus \{uv\} \cup p_2, p_{uv} \in P_{uv}^k\} \subseteq P^k$, where $P_{p_2}^k$ is the set of paths which use p_2 instead of the edge uv . Hence, for each $p \in P_{p_2}^k$ exists exactly one path $p_{uv} \in P_{uv}^k$. We can now define μ as follows.

$$\mu^k(p) := \begin{cases} \tilde{\mu}^k(p) + \tilde{\mu}^k(p_{uv}) & p \in P_{p_2}^k \\ 0 & p \in P_{uv}^k \\ \tilde{\mu}^k(p) & \text{otherwise.} \end{cases}$$

Thus, we shift the value $\tilde{\mu}^k(p_{uv})$ for every path using edge uv to the corresponding path using p_2 instead of uv . It is easy to see that the objective values for $(\tilde{\mu}, \tilde{\nu}, \tilde{z})$ and (μ, ν, z) are equal regarding ℓ_M . Furthermore, (11b) is fulfilled for all $k \in K$. If $e = uv$ then (11c) is valid since

$$\nu(uv) + z^k(uv) = 0 + 0 \geq \hat{d}^k \ell_M(uv) \sum_{p \in P_{uv}^k} \mu^k(p) = 0.$$

If $e \in p$ with $p \in P_{p_2}^k$, then (11c) is equivalent to

$$\begin{aligned} & \nu(e) + z^k(e) - \hat{d}^k \ell_M(e) \mu^k(p) \\ &= \tilde{\nu}(e) + \tilde{\nu}(uv) + \tilde{z}^k(e) + \tilde{z}^k(uv) - \hat{d}^k \ell_M(e) \left(\tilde{\mu}^k(p_{uv}) + \tilde{\mu}^k(p) \right) \\ &= \underbrace{\tilde{\nu}(e) + \tilde{z}^k(e) - \hat{d}^k \ell_M(e) \tilde{\mu}^k(p)}_{\geq 0} + \tilde{\nu}(uv) + \tilde{z}^k(uv) - \hat{d}^k \underbrace{\ell_M(e)}_{\leq \ell_M(p_2) = \ell_M(uv)} \tilde{\mu}^k(p_{uv}) \\ &\geq 0 + \tilde{\nu}(uv) + \tilde{z}^k(uv) - \hat{d}^k \ell_M(uv) \tilde{\mu}^k(p_{uv}) \\ &\geq 0, \end{aligned}$$

thus, also fulfilled. So (μ, ν, z) is feasible (and also optimal) for (11) with $\ell = \ell_M$. Moreover, $\mu^k(p) = 0$ for all $p \in P^k$ with $uv \in p$. Therefore, the objective value does not change for $\bar{\ell}$ and (μ, ν, z) is also feasible for (11) with $\ell = \bar{\ell}$ since $\bar{\ell}(e) = \ell_M(e)$ for all $e \in E \setminus \{uv\}$. Thus, $b_{\ell_M} \geq b_{\bar{\ell}}$ and the proof is completed. Note that we even have $b_{\ell_M} = b_{\bar{\ell}}$ since $b_{\ell_M} \leq b_{\bar{\ell}}$ holds as $\ell_M(e) \leq \bar{\ell}(e) \forall e \in E$. \square

Corollary 3.3. *RMIs together with the nonnegativity completely describe $Q_\Gamma(G, \bar{d}, \hat{d})$, i.e.,*

$$Q_\Gamma(G, \bar{d}, \hat{d}) = \left\{ x \in \mathbb{R}_+^{|E|} : \sum_{e \in E} \ell(e)x_e \geq b_\ell, \forall \ell \in \text{Met}(G) \right\}.$$

Based on the preceding result, we can now introduce a complete description of the RNLP in the capacity space.

$$\begin{aligned} \min \quad & (5a) \\ \text{s.t.} \quad & \sum_{e \in E} \ell_M(e)x_e \geq b & \forall \ell_M \in \text{Met}(G) & (12a) \\ & x_e \geq 0 & \forall e \in E & (12b) \end{aligned}$$

Due to the fact that Farkas' lemma can only be applied to LPs, Theorem 3.2 cannot be directly transferred to the integer case since we do not always have a good characterization of the right hand side b and $P_\Gamma(G, \bar{d}, \hat{d})$ itself. Therefore, the following theorem extends Theorem 3.2 to integer capacity variables. (Here, b_ℓ is not necessarily defined by (9).)

Theorem 3.4. *Let $\ell x \geq b$ be any valid inequality for $P_\Gamma(G, \bar{d}, \hat{d})$. Then there exists a metric $\ell_M \in \text{Met}(G)$ with*

- $\ell_M x \geq b$ valid,
- $\ell_M(ij) \leq \ell(ij)$.

Proof. Let ℓ be a non-metric length function and b the corresponding right hand side of the valid inequality. We define the metric ℓ_M as $\ell_M(ij) := \ell(i, j)$. Now assume there exists an $\bar{x} \in P_\Gamma(G, \bar{d}, \hat{d})$ with $\ell \bar{x} \geq b$ but $\ell_M \bar{x} < b$. Thus, there must exist an edge $ij \in E$ with $\ell_M(ij) < \ell(ij)$ and $\bar{x}_{ij} > 0$.

Let b_ℓ be the optimal solution of (9) regarding ℓ . Then $\ell x \geq b_\ell$ is a valid inequality for $Q_\Gamma(G, \bar{d}, \hat{d})$ and by Theorem 3.2, $\ell_M x \geq b_\ell$ is also valid for $Q_\Gamma(G, \bar{d}, \hat{d})$. We determine a feasible solution \bar{y} of $Q_\Gamma(G, \bar{d}, \hat{d})$ with $\bar{y}_e \leq \bar{x}_e \forall e \in E$ based on a feasible flow $\bar{f} \in [0, 1]^{|P|}$ computed as follows.

By definition of $Q_\Gamma(G, \bar{d}, \hat{d})$, there exist $(\bar{f}, \bar{\pi}, \bar{\rho})$ such that $(\bar{x}, \bar{f}, \bar{\pi}, \bar{\rho})$ satisfies (5b)–(5f). We set for all $e \in E$

$$\bar{y}_e := \sum_{k \in K} \bar{d}^k \bar{f}^k(e) + \Gamma \bar{\pi}_e + \sum_{k \in K} \bar{\rho}_e^k,$$

the left hand side of the capacity constraint (5c). Hence, $\bar{y} \in Q_\Gamma(G, \bar{d}, \hat{d})$ and $\ell \bar{y} \geq b_\ell$ is valid. Now replace the objective of (5) by the all zero function and exploit LP duality. Compared to (9), the objective now reads

$$\max - \sum_{e \in E} \bar{x}_e \ell(e) + \sum_{k \in K} b_k.$$

This value constitutes the violation of the inequality $\ell \bar{x} \geq b_\ell$ and, by the strong duality, is equal to 0. Hence, $\ell \bar{x} = b_\ell$. Since $\ell \bar{y} \geq b_\ell$ and $\bar{x} \geq \bar{y}$, it follows $\ell \bar{y} = b_\ell$. Since $\ell_M \bar{x} \geq b_\ell$ is valid for all $x \in Q_\Gamma(G, \bar{d}, \hat{d})$, it follows that $\bar{y}_{ij} = 0$ for all $ij \in E$ with $\ell_M(ij) < \ell(ij)$. Set $\tilde{x}_e = \lceil \bar{y}_e \rceil \forall e \in E$. Clearly, $\tilde{x} \in P_\Gamma(G, \bar{d}, \hat{d})$ but $\ell \tilde{x} = \ell_M \tilde{x} \leq \ell_M \bar{x} < b$, which implies that $\ell x \geq b$ is not valid for $\tilde{x} \in P_\Gamma(G, \bar{d}, \hat{d})$, a contradiction. \square

It is possible to prove Theorem 3.4 analogously to [4] or [15]. However, we have chosen this proof since it is constructive and makes an explicit use of the compact formulation of the Γ -robust NLP with static routing and splittable flows. A further way to prove Theorem 3.4 is by Chvátal-Gomory derivations as follows. By Theorem 3.2, $Q_\Gamma(G, \bar{d}, \hat{d})$ can be completely described by RMIs. Additionally, $P_\Gamma(G, \bar{d}, \hat{d}) \subseteq Q_\Gamma(G, \bar{d}, \hat{d})$ holds. All facet-defining inequalities defining $P_\Gamma(G, \bar{d}, \hat{d})$ can be derived by a sequence of Chvátal-Gomory derivations from the facet-defining inequalities of $Q_\Gamma(G, \bar{d}, \hat{d})$. Since $\text{Met}(G)$ is a cone, these derived inequalities are also RMIs. Thus, $P_\Gamma(G, \bar{d}, \hat{d})$ can be completely described by RMIs.

The right hand side of valid inequalities $\ell_M x \geq b$ in Theorem 3.4 cannot be derived by (9) or a similar procedure. A first step can be done by Chvátal-Gomory rounding. For a RMI (12a), we can assume $\ell_M \in \mathbb{Z}_+^{|E|}$. (If $\ell_M(e) \notin \mathbb{Z}_+$ for $e \in E$, the metric can be scaled such that all lengths are integer.) Since $x \in \mathbb{Z}_+$, the *rounded robust metric inequality* (rounded RMI)

$$\sum_{e \in E} \ell_M(e) x_e \geq \lceil b_{\ell_M} \rceil \quad (13)$$

is also a valid inequality for $P_\Gamma(G, \bar{d}, \hat{d})$.

The strongest RMI is defined as follows.

Definition 3.5. Let $\ell_M \in \text{Met}(G)$ and let $\rho(\ell_M, G, \bar{d}, \hat{d})$ be the optimal solution of the Γ -robust NLP where the cost vector κ in the objective function (5a) is replaced by ℓ_M . Thus, $\rho(\ell_M, G, \bar{d}, \hat{d}) = \min\{\ell_M x : x \in \mathbb{Z}_+, x \in P_\Gamma(G, \bar{d}, \hat{d})\}$. Obviously, any valid RMI $\ell_M x \geq b$ for $P_\Gamma(G, \bar{d}, \hat{d})$ is dominated by $\ell_M x \geq \rho(\ell_M, G, \bar{d}, \hat{d})$. Hence, any inequality of the form $\ell_M x \geq \rho(\ell_M, G, \bar{d}, \hat{d})$ is denoted as *tight robust metric inequality* (tight RMI).

Corollary 3.6. *Tight RMIs $\ell_M x \geq \rho(\ell_M, G, \bar{d}, \hat{d})$ with $\ell_M \in \text{Met}(G)$ completely describe $P_\Gamma(G, \bar{d}, \hat{d})$.*

The right hand side value $\rho(\ell_M, G, \bar{d}, \hat{d})$ of a tight RMI can be determined easily without solving a minimization problem over $P_\Gamma(G, \bar{d}, \hat{d})$ if the metric ℓ_M is an extreme ray of the metric cone. This result is known for the classical NLP [4] and can be generalized to the RNLP with static routing and splittable flows as follows.

Theorem 3.7. *If $\ell_M : E \rightarrow \mathbb{Z}_+$ is an extreme ray of the metric cone $\text{Met}(G)$ such that the greatest common divisor of ℓ_M is 1, then $\rho(\ell_M, G, \bar{d}, \hat{d}) = \lceil b_{\ell_M} \rceil$.*

Proof. We recall that $Q_\Gamma(G, \bar{d}, \hat{d})$ can be completely described by RMIs due to Theorem 3.2 and the subset correlation $P_\Gamma(G, \bar{d}, \hat{d}) \subseteq Q_\Gamma(G, \bar{d}, \hat{d})$. All valid inequalities defining $P_\Gamma(G, \bar{d}, \hat{d})$ can be derived by a sequence of Chvátal-Gomory derivations from the facet-defining inequalities of $Q_\Gamma(G, \bar{d}, \hat{d})$. Consider a metric length function ℓ_M that satisfies the conditions of the theorem and suppose that, for all $x \in P_\Gamma(G, \bar{d}, \hat{d})$, ℓ_M satisfies the inequality $\ell_M x \geq \alpha$ with $\alpha > \lceil b_{\ell_M} \rceil$. This inequality cannot be derived by a Chvátal-Gomory procedure from the single RMI $\ell_M x \geq b_{\ell_M}$. Thus, it can only be derived by combining two or more inequalities in at least one step of the sequence of Chvátal-Gomory derivations. But this contradicts the assumption that ℓ_M is an extreme ray of $\text{Met}(G)$. \square

Note, following a two-stage approach to solve the RNLP, an analogous proof has been done by Mattia [15] in the context of the RNLP with dynamic routing.

4 Separation

In this section, we investigate the separation of RMIs. First, we present a polynomial time exact separation algorithm to separate RMIs as model inequalities. Second, we present an exact separation procedure to determine violated rounded RMIs in a cut-and-branch approach for solving the RNLP. Third, we address tight RMIs pointing out the differences to the non-robust setting.

The separation algorithms presented in this section separate violated RLIs in a first step. If the length function ℓ is not metric, a metric ℓ_M and hence a corresponding violated metric inequality can be constructed by $\ell_M(e) := \min\{\ell(e), \ell(u, v)\}$ for all $e = uv \in E$ (cf. Theorem 3.2). Note, the right hand side of the RLI does not change and the violation of the RMI is at least the violation of the previous RLI.

Separation of robust metric inequalities. In the following, we describe the exact separation of violated RMIs analogously to the non-robust case considered by Avella et al. [4] and analogously to the RNLP with dynamic routing by Mattia [15]. Therefore, we define $f_e^k := \sum_{p \in P^k: e \in p} f^k(p)$ as the flow on edge $e \in E$ for commodity $k \in K$. Given a capacity vector $\tilde{x} \in \mathbb{R}_+^{|E|}$, the feasibility problem can be formulated as

$$\max \alpha \tag{14a}$$

$$\text{s. t. } \sum_{uv \in E} (\tilde{f}_{uv}^k - \tilde{f}_{vu}^k) \geq \begin{cases} \alpha & u = s^k \\ -\alpha & u = t^k \\ 0 & \text{otherwise} \end{cases} \quad \forall u \in V, k \in K \tag{14b}$$

$$\sum_{k \in K} \tilde{d}^k \tilde{f}_e^k + \Gamma \pi_e + \sum_{k \in K} \rho_e^k \leq \tilde{x}_e \quad \forall e \in E \tag{14c}$$

$$\hat{d}^k \tilde{f}_e^k - \pi_e - \rho_e^k \leq 0 \quad \forall e \in E, k \in K \tag{14d}$$

$$f_{uv}^k, \pi_e, \rho_e^k, \alpha \geq 0 \quad \forall e = uv \in E, k \in K. \tag{14e}$$

The value of an optimal solution $(\alpha^*, \pi^*, \rho^*)$ of (14) is at least 1 if and only if the point $(\tilde{x}, \tilde{f}, \pi^*, \rho^*)$ is feasible for $Q_\Gamma(G, \tilde{d}, \hat{d})$. The dual of (14) reads

$$\min \sum_{e \in E} \tilde{x}_e \ell(e) \tag{15a}$$

$$\text{s. t. } \beta_u^k - \beta_v^k \leq \tilde{d}^k \ell(e) + \hat{d}^k m^k(e) \quad \forall e = uv \in E, k \in K \tag{15b}$$

$$\sum_{k \in K} m^k(e) \leq \Gamma \ell(e) \quad \forall e \in E \tag{15c}$$

$$m^k(e) \leq \ell(e) \quad \forall e \in E, k \in K \tag{15d}$$

$$\sum_{k \in K} (\beta_{s^k}^k - \beta_{t^k}^k) \geq 1 \tag{15e}$$

$$\beta_u^k, \ell(e), m^k(e) \geq 0 \quad \forall u \in V, e \in E, k \in K. \tag{15f}$$

By strong duality, \tilde{x} is feasible for $Q_\Gamma(G, \tilde{d}, \hat{d})$ if and only if the optimal value of (15) is at least 1. If the objective value of a solution (β^*, ℓ^*, m^*) is strictly less than 1, constraint (15e) implies the RLI

$$\sum_{e \in E} \ell^*(e) x_e \geq \sum_{k \in K} (\beta_{s^k}^{*k} - \beta_{t^k}^{*k}) \tag{16}$$

to be violated for \tilde{x} .

Corollary 4.1. *Violated RMIs can be separated in polynomial time.*

Proof. Since (15) is a pure linear program with polynomial size and ℓ_M is constructible from ℓ in polynomial time, violated RMIs can be exactly separated in polynomial time. \square

Despite this result, experiments have shown that the separation of these inequalities can be very time consuming in practice; cf. Section 5.3.

Separation of rounded robust metric inequalities. To cut-off a fractional solution of the LP relaxation of (5), the right hand side of a corresponding RLI (with integer lengths) has to be rounded up to yield a violation; cf. derivation of (13). Hence, formulation (15) has to be modified to take the integrality of ℓ and the rounding into account. The exact separation of rounded RLIs can be formulated as the following ILP.

$$\max R - \sum_{e \in E} \tilde{x}_e \ell(e) \quad (17a)$$

$$\text{s. t. (15b) - (15d)}$$

$$\sum_{k \in K} (\beta_{t^k}^k - \beta_{s^k}^k) + R \leq 1 - \varepsilon \quad (17b)$$

$$R \in \mathbb{Z} \quad (17c)$$

$$\ell(e) \in \mathbb{Z}_+ \quad \forall e \in E \quad (17d)$$

$$\beta_u^k, m^k(e) \geq 0 \quad \forall u \in V, e \in E, k \in K, \quad (17e)$$

where, given a small constant $\varepsilon > 0$, constraint (17b) determines the rounded up right hand side value R of the resulting RLI. The objective value of (17) equals the violation of the rounded RLI. Hence, a non-positive objective value gives a proof that no such inequality exists.

Given a violated rounded RLI, a corresponding violated rounded RMI (13) can be constructed as described above.

Separation of tight robust metric inequalities. The strongest RMIs are those with a tight right hand side; cf. Corollary 3.6. By definition, a minimization problem over the Γ -robust network loading polytope has to be solved to determine the best right hand side $\rho(\ell_M, G, \bar{d}, \hat{d})$ to obtain a tight RMI. Clearly, this is as hard as solving the original Γ -robust NLP. In this paragraph, we present a preprocessing method to speed up the computation of $\rho(\ell_M, G, \bar{d}, \hat{d})$ in practice.

Avella et al. [4] suggest a shrinking heuristic to reduce the network size before determining the best right hand side $\rho(\ell_M, G, \bar{d}, \hat{d})$ given a ℓ_M in the non-robust setting. In the following, we describe a version of this shrinking that can be used in the robust setting.

Theorem 4.2. *Let $\ell_M \in \text{Met}(G)$ and $\{i, j\} \in E$ with $\ell_M(ij) = 0$. Define the shrunken graph $G^h = (V^h, E^h)$ by*

$$V^h := V \setminus \{i, j\} \cup \{h\}$$

$$E^h := E \setminus \{uv \in E \mid u \in \{i, j\} \wedge v \in \{i, j\}\} \cup \{uh \mid ui \in E \wedge uj \in E\}.$$

Then $\rho(\ell_M^h, G^h, \bar{d}^h, \hat{d}^h) = \rho(\ell_M, G, \bar{d}, \hat{d})$.

Proof. The proof is analog to the proof of Theorem 3.3 in [4].

Suppose $\rho(\ell_M, G, \bar{d}, \hat{d}) < \rho(\ell_M^h, G^h, \bar{d}^h, \hat{d}^h)$. Let x satisfy $\ell_M x = \rho(\ell_M, G, \bar{d}, \hat{d})$ and x^h be the mapping of x on G^h . Then $\rho(\ell_M^h, G^h, \bar{d}^h, \hat{d}^h) \leq \ell_M^h x^h = \ell_M x = \rho(\ell_M, G, \bar{d}, \hat{d}) < \rho(\ell_M^h, G^h, \bar{d}^h, \hat{d}^h)$ holds; a contradiction.

Suppose $\rho(\ell_M, G, \bar{d}, \hat{d}) > \rho(\ell_M^h, G^h, \bar{d}^h, \hat{d}^h)$. Let x^h satisfy $\ell_M^h x^h = \rho(\ell_M^h, G^h, \bar{d}^h, \hat{d}^h)$. Define x as follows.

$$x_e := \left[\sum_{k \in K} \bar{d}^k f_e^k + \max_{Q \subseteq K, |Q| \leq \Gamma} \hat{d}^k f_e^k \right],$$

$x_{uh}^h = x_{ui} + x_{uj}$ for all $u \in V \setminus \{i, j\}$, and $x_{uv}^h := x_{uv}$ for all $uv \in E, u, v \in V \setminus \{i, j\}$. Then x is feasible for $P_\Gamma(G, \bar{d}, \hat{d})$ and $\ell_M x = \ell_M^h x^h$ holds by construction. Further, it holds $\ell_M x = \ell_M^h x^h = \rho(\ell_M^h, G^h, \bar{d}^h, \hat{d}^h) < \rho(\ell_M, G, \bar{d}, \hat{d})$; a contradiction to the minimality of $\rho(\ell_M, G, \bar{d}, \hat{d})$. \square

In contrast to the non-robust setting [4], commodities cannot be aggregated in the shrinking procedure for the robust setting. Instead, the shrunken graph may have ‘‘parallel’’ commodities with same source and destination nodes corresponding to different nodes in the original graph. The following example illustrates the difference to the non-robust setting.

Consider the complete graph G with nodes $V := \{n_1, n_2, n_3\}$, edges $E := \{n_1 n_2, n_1 n_3, n_2 n_3\}$, commodities K (with $|K| = 2$), and $\Gamma = 1$. Further, let $\ell_M(n_1 n_2) = 0, \ell_M(n_1 n_3) = \ell_M(n_2 n_3) = 1$ be a metric. In the robust setting, Γ many commodities may deviate on each edge in this network. Applying the shrinking heuristic of [4] results in the reduced graph G^h with nodes $V^h := \{h, n_3\}$, edges $E := \{h n_3\}$, and new metric $\ell_M^h(h n_3) = 1$. In contrast to the non-robust setting, the commodities cannot be aggregated since the deviations of a single aggregated commodity from h to n_3 cannot reflect the independent demand deviations on the two edges $n_1 n_3$ and $n_2 n_3$. Furthermore, this cannot be overcome by increasing the Γ for G^h as the resulting solution might violate the limit of $\Gamma = 1$ on the original edges $n_1 n_3$ and $n_2 n_3$.

Selected subclasses of robust metric inequalities Some well-known classes of valid inequalities for the RNLP are generalized by RMIs. In the following we point out selected subclasses of RMIs that are considered in the computations in Section 5.

We define (*L*-)bounded RMIs as the subclass of RMIs consisting of all rounded RMIs obtained from metrics with integer link lengths bounded by $L \in \mathbb{Z}_+$, i. e., $\text{Met}_L(G) := \{\ell_M \in \text{Met}(G) : \ell_M \in \{0, 1, \dots, L\}^{|E|}\}$. Bounded RMIs can be separated by solving formulation (17), where constraint (17d) is replaced by

$$\ell(e) \in \{0, 1, \dots, L\}, \quad (17d')$$

and strengthening the obtained RLI to a RMI as before.

The class of 1-bounded RMIs consists of all robust partition inequalities, i.e., the robust counterparts of partition inequalities (see Agarwal [3] and the references therein). A well-known subclass of robust partition inequalities are the robust 2-partition inequalities, the so-called robust cutset inequalities. Koster et al. [12] have considered the separation of violated robust cutset inequalities presenting exact as well as efficient heuristic separation approaches.

5 Computations

In this section, we present the results of two computational studies. First, we investigate the strengthening of the robust network loading formulation by RMIs. Second, we evaluate the speed-up by sep-

Network	<i>Abilene</i>	<i>GÉANT</i>	<i>germany17</i>	<i>germany50</i>
# nodes	12	22	17	50
# links	15	36	26	89
# demands	66	231	136	1044
available traffic period	6 months	4 months	1 day	1 day
traffic granularity	5 min	15 min	5 min	5 min
# traffic matrices used	2×2016	2×672	288	288
instances	<i>abilene1</i> <i>abilene2</i>	<i>geant1</i> <i>geant2</i>	<i>germany17</i>	<i>germany50</i>

Table 1: Network and traffic properties of considered data sets; cf. Koster et al. [12]

arating RMIs in a cut-and-branch approach. In addition, we compare the capacity formulation (12) of the RNLP with the compact flow formulation presented in Koster et al. [12] (objective (5a) with constraints (14b)–(14e), (5g) with $\alpha = 1$) computationally. All our computational experiments are carried out using the same instances from real-life traffic measurements as in [12].

Instances. Since the considered problem instances are described in detail by Koster et al. [12] and are publicly available at SNDlib [18], we just give a brief summary here. The live traffic data, given as a set of measured traffic matrices, is taken from the *Abilene* network [1], the *GÉANT* network, and the network operated by the German DFN-Verein [7] which is mapped on the network *germany17* [16], and in addition mapped on a larger network (*germany50*) [18]. The granularity of the traffic matrices is either 5 (*Abilene*, *germany17*, *germany50*) or 15 minutes (*GÉANT*). By considering two time periods of one week for *Abilene* and *GÉANT* and one day for *germany17* as well as for *germany50*, we create six problem instances which are summarized in Table 1. For each, we consider $\Gamma \in \{0, \dots, 10\}$ which yields 66 instances in total.

For each data set, let T denote the considered time period and let $d_{(t)}^k$ be the demand value for commodity $k \in K$ at time step $t \in T$. First, we scale the traffic data such that the sum of all peak demands $\max_{t \in T}(d_{(t)}^k)$ over all commodities $k \in K$ amounts to 1 Tbps. To determine the nominal value \bar{d}^k and the peak value $\bar{d}^k + \hat{d}^k$, respectively, we calculate the arithmetic mean and 95%-percentile of each commodity $k \in K$ using the scaled measurements. Thus, we set $\bar{d}^k := 1/|T| \sum_{t \in T} d_{(t)}^k$ and $\bar{d}^k + \hat{d}^k$ corresponds to the largest deviation from the nominal value in period T'_k , where T'_k is obtained from T by removing the 5% largest demands (arguments for this setting can be found in [11]).

General settings [12]. We implemented the formulation (5) of the Γ -robust NLP in C++ using IBM ILOG CPLEX 12.4 [9] as branch-and-cut framework. For the separation methods, we used the CONCERT framework of CPLEX and callbacks. We ran the computations single-threaded on a Linux machine with a 3.40 GHz Intel Core i7-2600 CPU and 16 GB RAM. A time limit of 12 hours was set for solving each problem instance. Default settings were used if not stated differently.

5.1 Strengthening the compact flow formulation

In our first computational study, we investigate the strengthening of the linear relaxation of the compact flow formulation of RNLP (objective (5a) with constraints (14b)–(14e), (5g) with $\alpha = 1$) at the

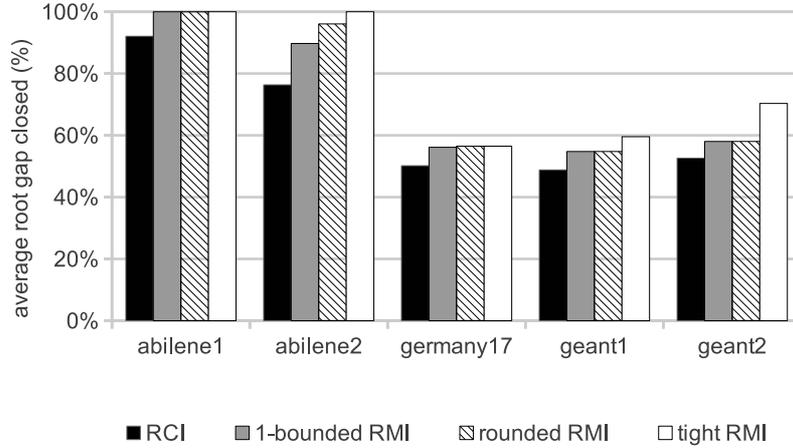


Figure 1: Average gap closed at the root node by RCIs, 1-bounded RMIs, rounded RMIs, and tight RMIs

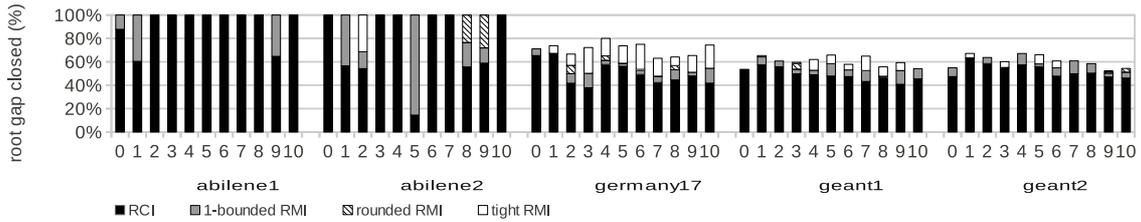


Figure 2: Gap closed at the root node by RCIs, 1-bounded RMIs, RMIs, and tight RMIs, for $\Gamma = 0, \dots, 10$

root node by different subclasses of RMIs. We consider the exact separation of the following four subclasses of RMIs: robust cutset inequalities, 1-bounded RMIs (i.e. robust partition inequalities), rounded RMIs, and tight RMIs.

Using the callback functionality of CPLEX, we add an exact separation algorithm solving (17) including the option to separate bounded RMIs or tighten the right hand side. Furthermore, we use the ILP-based separation algorithm presented in [12] to separate robust cutset inequalities exactly. In this study, only the root node is solved, all CPLEX cuts are turned off, and the gap closed at the end of the root node is evaluated. By this the added value of the RMIs for the compact flow formulation can be determined.

The gap closed is the ratio $(DB_{\text{root}} - DB_{\text{LP}}) / (PB_{\text{best}} - DB_{\text{LP}})$ where DB_{root} denotes the dual bound after solving the root node before branching, DB_{LP} the objective value of the LP relaxation at the end of the root node, and PB_{best} the (overall) best known primal bound. A gap closed of 100% is observed if the instance could be solved to optimality at the root node.

Figures 1 and 2 show the (average) gap closed for each network (except *germany50*), value of Γ , and considered subclasses of RMIs. For *germany50*, the root node could not be solved within the time limit in any setting. This is due to the fact that the standard cuts of CPLEX were turned off in

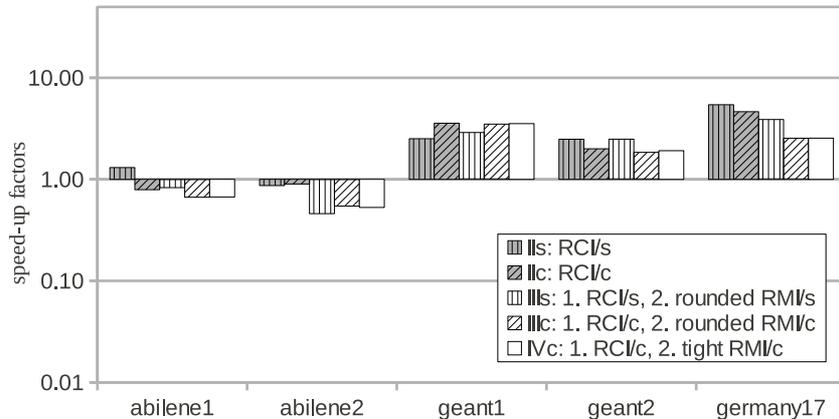


Figure 3: Speed-up factors of algorithms IIs, IIIs, IIc–IVc in a cut-and-branch approach to solve the compact flow formulation of the RNLP. All factors are normalized to solution times of CPLEX as stand-alone solver.

this study. A further evaluation of *germany50* in a full cut-and-branch approach is presented later (see Section 5.3).

Figure 1 shows the average gap closed by separating robust cutset inequalities (RCIs), 1-bounded RMIs, rounded RMIs, or tight RMIs exactly. We observe that already RCIs close the gap on average at least by 48.8 % (*geant1*). This concurs with the conclusion of the effectiveness of RCIs in [12]. When considering more general RMIs, the average gap can be closed by additional 7.7 % by 1-bounded RMIs, further 1.3 % by rounded RMIs, and even further 4.2 % by tight RMIs. In this way, the gap can be closed in total by additional 13.3 % by tight RMIs compared to RCIs. Furthermore, the total gap closed by tight RMIs is 100 % for both *Abilene* instances and $\Gamma = 0, \dots, 10$, solving these instances optimally at the root node.

Figure 2 provides a break down of the gap closed by individual subclasses of RMIs per instance and value of Γ . Here we also observe that RCI contribute most to the gap closed. Nevertheless, more general RMIs as (1-bounded) RMIs are needed to completely close the gap at the root node for all *Abilene* instances. In addition, tight RMIs are needed for *abilene2* and $\Gamma = 2$ to achieve 100 % gap closed.

The results of our first computational study show that separating violated RMIs closes the gap significantly at the root node.

5.2 Speeding-up the compact flow formulation

In our second computational study, we investigate the speed-up by integrating the separation of violated RMIs in a cut-and-branch approach to solve the compact flow formulation of the RNLP, i.e., objective (5a) with constraints (14b)–(14e), (5g) with $\alpha = 1$. We consider several heuristic algorithms summarized in Table 2. Note that algorithm IIs corresponds to the cut-and-branch using RCIs proposed in [12]. Figure 3 shows the average speed-up factors obtained by the individual algorithms. The factors are normalized to the solution time of CPLEX as stand-alone solver (algorithm I). For example, a solution time of 60 seconds compared to a corresponding solution time of 120 seconds of CPLEX yields a speed-up factor of 2. For the *Abilene* network, we observe that all average speed-up

	Algorithm	Description
I	CPLEX	CPLEX as stand-alone solver; no additional separation of RMIs.
IIs	RCI/s	Heuristic separation of violated RCIs using the shrinking heuristic described in [12]. The network is shrunken with respect to the slack values of inequalities (5c) and (5d). If no violated cut is found heuristically, an ILP-based exact separation algorithm is run.
IIC	RCI/c	Same as algorithm IIs, but the network is shrunken with respect to the LP value of the capacity variables x .
IIIs	1. RCI/s, 2. rounded RMI/s	First, algorithm IIs is run. Second, if no cut has been found, the network is shrunken with respect to the slack values of (5c) and (5d). Then violated RMIs are separated by solving the ILP (17) for the shrunken network.
IIIC	1. RCI/c, 2. rounded RMI/c	Same as algorithm IIIs, but all network shrinkings are done with respect to the LP value of the capacity variables x .
IIIC	1. RCI/c, 2. tight RMI/c	Same as algorithm IIIC. If a violated RMI is found, its best right hand side $\rho(\ell_M, G, \bar{d}, \hat{d})$ is determined by solving an ILP. Only the resulting tight RMI is separated.

Table 2: Summary of algorithms considered in Sections 5.2 and 5.3

factors are less than 1.0 except for algorithm IIs and *abilene1*. In fact, the solution times of *abilene1* and *abilene2* instances are in the range of only a few seconds. Therefore, the overhead introduced by all separation algorithms results in a slow-down on average. For all larger instances, there is a significant average speed-up factor of at least 1.85 (algorithm IIIC, *geant2*). Separating only violated RCIs (algorithms IIs and IIC) yields good speed-up factors for *geant1*, *geant2*, and *germany17*. Most of the times these are better than the corresponding speed-up factors obtained by algorithms IIIC and IIIC where additionally tight RMI are separated heuristically. This can be explained by the additional computational effort to solve the ILP (17) to separate violated RMIs. Although the network has been shrunken, the remaining integer problem may still be computationally hard reducing the speed-up factor compared to the RCI-only algorithm IIC. A detailed table of all speed-up factors for all algorithms, considered networks, and values of Γ can be found in Table A.1 in the appendix.

In our study, we have observed a substantial speed-up in solving the compact flow formulation of the RNLP on mid-sized instances (e.g., *geant1*, *geant2*, and *germany17*) by separating RMIs in a cut-and-branch approach. Furthermore, the highest speed-up factors are achieved by separating RCIs.

5.3 Performance of the capacity formulation

In our last computational study, we compare the compact flow formulation (objective (5a) with constraints (14b)–(14e), (5g) with $\alpha = 1$) to the capacity formulation (12) in a cut-and-branch approach. The exponentially many robust metric model inequalities of the capacity formulation are handled im-

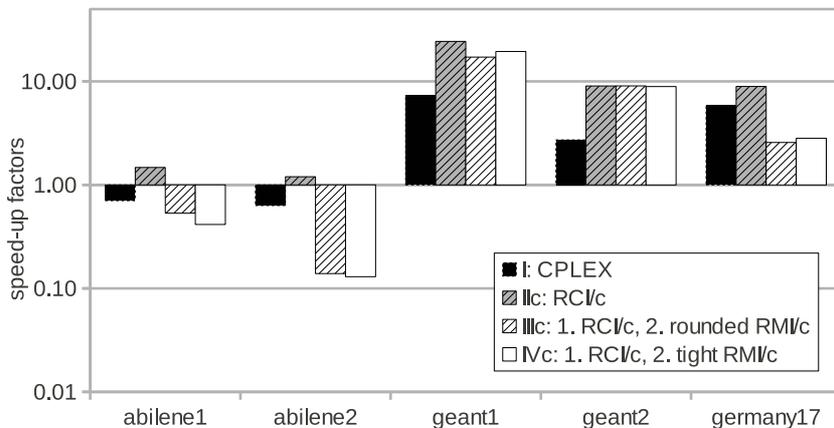


Figure 4: Speed-up factors of algorithms IIc–IVc in a cut-and-branch approach to solve the capacity formulation of the RNLP. Algorithm I: CPLEX corresponds to solving the capacity formulation with CPLEX as a stand-alone solver. All factors are normalized to solution times of CPLEX solving the compact flow formulation as stand-alone-solver.

plicity: violated RMIs (16) separated on-the-fly as lazy constraints using the callback capabilities of CPLEX whenever an integer solution is found. We start without any RMI.

We consider algorithms I, IIc, and IVc introduced in Section 5.2 and summarized in Table 2. Algorithms II and III are not applicable to the incomplete capacity formulation because the slack values of inequalities (5c) and (5d) cannot be evaluated as these constraints do not exist.

Figure 4 shows the average speed-up factors of algorithms IIc–IVc normalized to the solution times of CPLEX solving the compact flow formulation as stand-alone solver. Because the solution times of CPLEX as stand-alone solver solving the compact flow formulation and the capacity formulation differ, we report also on the normalized solution times of CPLEX for the capacity formulation. Similarly to the compact flow formulation, we observe that on the small-sized *abilene1* and *abilene2* instances no speed-up can be achieved. The computational effort to separate RMI slows down the instances compared to their fast solution times. For *geant1*, *geant2*, and *germany17*, we notice significant average speed-up factors of 2.7 (*geant1*), 5.9 (*germany17*), and 7.3 (*geant2*) obtained by CPLEX solving the capacity formulation (algorithm I). Introducing the separation of RMIs yields even higher speed-up factors on average whereas the best are achieved by algorithm IIc: 1.5 (*abilene1*), 1.2 (*abilene2*), 24.4 (*geant1*), 9.0 (*geant2*), and 8.9 (*germany17*). For each network, the best average speed-up factor for the capacity formulation is higher than the corresponding one for the compact flow formulation (cf. Figure 3). In particular the algorithms II (only compact flow formulation) and IIc perform well achieving the highest average speed-up factors in most cases. In summary, we conclude that the capacity formulation is computationally more tractable than the compact flow formulation for mid-sized instances. A significant speed-up could be observed in our computational studies. Details of all speed-up factors for all algorithms, considered networks, and values of Γ can be found in Table A.1 in the appendix.

To evaluate the impact of RMIs in a cut-and-branch approach to larger instances such as *germany50* which cannot be solved within the time limit of 12 hours, we consider the optimality gap after 12 hours. In particular, we evaluate the gap reduction factor compared to the optimality gap left

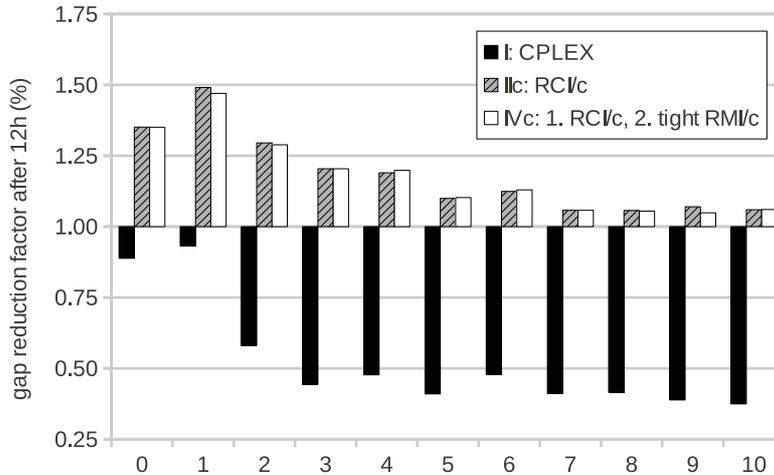


Figure 5: Gap reduction factors of algorithms I, IIc, and IVc in a cut-and-branch approach to solve the capacity formulation of the RNLP for the *germany50* instance. Algorithm I: CPLEX corresponds to solving the capacity formulation with CPLEX as a stand-alone solver. All factors are normalized to the optimality gap left by CPLEX as a stand-alone solver at the time limit of 12 hours when solving the compact flow formulation.

by CPLEX as stand-alone solver when solving the compact flow formulation to the time limit. For example, let CPLEX solve the *germany50* instance to 60 % of optimality and let another algorithm obtain a optimality gap of 40 %. Then the gap reduction factor of this other algorithm is $0.6/0.4 = 1.5$ compared to CPLEX. Therefore, a gap reduction factor less than 1.0 is given if the optimality gap after 12 hours is larger than the corresponding one by CPLEX and the compact flow formulation. The cuts obtained by algorithm IIc may be strengthened to tight RMIs and hence are dominated by those obtained by algorithm IVc. Thus, we do not consider algorithm IIc in this study.

Figure 5 shows the gap reduction factor for *germany50* compared to CPLEX as stand-alone solver solving the compact flow formulation. First, we notice that algorithm I (CPLEX) for the capacity formulation cannot close the optimality gap better than CPLEX for the compact flow formulation. In fact, in most cases algorithm I finishes with optimality gaps at least twice as high. This bad performance can be explained by the fact that during the solution process of algorithm I temporary solutions are obtained which are in fact infeasible to the complete capacity formulation but not the current incomplete one. To lower the computational effort, these solutions are only separated if they are integer. In contrast, algorithms IIc and IVc separate violated rounded RMIs for fractional solutions in addition to the model constraints. So both finish with smaller optimality gaps compared to CPLEX and the compact flow formulation. The gap reduction factor ranges from 1.1 ($\Gamma = 10$) to 1.5 ($\Gamma = 1$). For $\Gamma \leq 4$ the gap reduction factor is at least 1.2. For larger values of Γ it decreases. The gap reduction factors do not differ much between the algorithms IIc and IVc except for $\Gamma = 1$ and $\Gamma = 9$ where algorithm IIc clearly outperforms the others. In summary, we observe that a cut-and-branch approach with rounded RMIs on the capacity formulation of the RNLP significantly closes the optimality gap of the large *germany50* instance in our study.

In conclusion, in our studies it turns out that the capacity formulation of the RNLP is computationally more tractable. When used in a cut-and-branch approach, it offers higher speed-up factors on mid-sized instances and lowers the optimality gap left at the time limit for larger instances.

6 Conclusions

In this paper, following the requirements of (telecommunication) network operators we have investigated the Γ -robust NLP with static routing and splittable flows focusing on RMIs, the generalization of metric inequalities in the classical NLP to the robust setting. In particular, we have shown constructively that the robust network loading polytope in the capacity space is completely described by RMIs and nonnegativity constraints.

Furthermore, we have developed (i) a polynomial time exact separation algorithm to determine violated RMIs as model inequalities and (ii) exact ILP-based separation algorithms to separate violated rounded RMIs, bounded RMIs, and tight RMIs in a cut-and-branch approach.

To investigate the performance of the RMIs, we have carried out several tests on networks with demands based on realistic traffic data. The computations have shown that the integrality gap at the root node can be closed significantly by heuristic separation of tight RMIs. On mid-sized instances, a cut-and-branch approach including the separation of rounded RMIs yields an average speed-up factor between 2 and 5 (for the compact flow formulation) and between 3 and 25 for the capacity formulation, respectively. Furthermore, for the larger *germany50* instance the optimality gap left at the time limit could be closed by a factor up to 1.5 using a cut-and-branch approach to solve the capacity formulation of the RNLP. In conclusion, our computational results demonstrate the major potential of (tight) RMIs.

Acknowledgement

This work was supported by the German Federal Ministry of Education and Research (BMBF grant 03MS616A, project ROBUKOM, www.robukom.de), the UMIC Research Centre at RWTH Aachen University, the DFG research grant KO2311/3-1, SCHM2643/5-1, and the PROCOPE program of the German Academic Exchange Service (DAAD).

References

- [1] Abilene Internet2 Network. www.internet2.edu/network.
- [2] Y. K. Agarwal. k-Partition-based facets of the network design problem. *Networks*, 47(3):123–139, 2006.
- [3] Y. K. Agarwal. Polyhedral structure of the 4-node network design problem. *Networks*, 54: 139–149, October 2009. doi: 10.1002/net.v54:3.
- [4] P. Avella, S. Mattia, and A. Sassano. Metric inequalities and the network loading problem. *Discrete Optimization*, 4:103–114, 2006.
- [5] F. Barahona. Network design using cut inequalities. *SIAM J. on Optimization*, 6:823–837, March 1996. doi: 10.1137/S1052623494279134.
- [6] D. Bertsimas and M. Sim. The Price of Robustness. *Operations Research*, 52(1):35–53, 2004.
- [7] Deutsche Forschungsnetz (DFN). <http://www.dfn.de>.
- [8] O. Günlük. A branch and cut algorithm for capacitated network design problems. *Mathematical Programming*, 86:17–39, 1999.

- [9] IBM – ILOG. CPLEX version 12.4, 2011. <http://www.ilog.com/products/cplex>.
- [10] M. Iri. On an Extension of the Maximum-flow Minimum-cut Theorem to Multicommodity Flows. *Journal of the Operations Research Society of Japan*, 13(3):129–135, 1971.
- [11] A. M.C.A. Koster, M. Kutschka, and C. Raack. On the robustness of optimal network designs. In *Proceedings of IEEE International Conference on Communications, ICC 2011*, pages 1–5. IEEE Xplore, 2011.
- [12] A. M.C.A. Koster, M. Kutschka, and C. Raack. Robust Network Design: Formulations, Valid Inequalities, and Computations. *Networks*, to appear.
- [13] M. Lomonosov. On feasibility conditions of multicommodity flows in networks. *Discrete Mathematics*, 1982.
- [14] T. L. Magnanti, P. Mirchandani, and R. Vachani. The convex hull of two core capacitated network design problems. *Mathematical Programming*, 60:233–250, 1993.
- [15] S. Mattia. The robust network loading problem with dynamic routing. *Computational Optimization and Applications*, August 2012. doi: 0.1007/s10589-012-9500-0.
- [16] NOBEL. NOBEL – Next generation Optical networks for Broadband European Leadership , 2004–2007. <http://www.ist-nobel.org>.
- [17] K. Onaga and O. Kakusho. On feasibility conditions of multicommodity flows in networks. *Circuit Theory, IEEE Transactions on*, 18(4):425 – 429, July 1971. doi: 10.1109/TCT.1971.1083312.
- [18] S. Orłowski, R. Wessäly, M. Pióro, and A. Tomaszewski. SNDlib 1.0–Survivable Network Design Library. *Networks*, published online, 2009. <http://sndlib.zib.de>.

A Appendix

	compact flow formulation							capacity formulation				
	Γ	solution time		speed-up factors					I	IIc	IIIc	IVc
		I: CPLEX [s]	IIs	IIc	IIIs	IIIc	IVc					
<i>abilene1</i>	0	0.9	5.00	0.63	0.56	0.50	0.50	0.56	2.50	1.25	0.83	
	1	2.1	0.44	0.44	0.05	0.12	0.13	0.38	1.14	0.15	0.09	
	2	2.4	0.56	0.71	0.56	0.71	0.77	0.42	1.25	0.08	0.08	
	3	2.4	0.86	0.63	0.86	0.63	0.63	0.50	0.80	0.32	0.23	
	4	3.1	0.79	1.00	0.79	1.00	1.00	0.35	0.65	0.14	0.13	
	5	1.8	0.26	0.26	0.11	0.09	0.08	0.67	0.55	0.11	0.10	
	6	2.3	0.68	0.50	0.68	0.50	0.50	0.65	1.07	0.05	0.05	
	7	3.2	1.04	1.33	0.76	1.27	1.33	0.88	2.00	1.17	1.04	
	8	2.8	1.88	0.97	1.88	0.63	0.63	1.07	2.31	1.30	0.88	
	9	2.5	2.13	1.55	2.13	1.55	1.55	1.36	2.83	1.03	0.83	
10	2.1	0.76	0.68	0.76	0.35	0.26	0.90	1.12	0.27	0.31		
<i>abilene2</i>	0	0.9	1.50	0.30	0.60	0.19	0.19	0.33	0.75	0.60	0.50	
	1	1.7	0.57	0.80	0.55	0.20	0.20	0.94	1.60	0.24	0.21	
	2	2.7	0.49	0.83	0.17	0.19	0.19	0.70	1.58	0.14	0.13	
	3	2.8	0.57	0.59	0.54	0.57	0.39	0.46	0.93	0.13	0.16	
	4	1.7	0.39	0.35	0.39	0.33	0.25	0.41	0.70	0.03	0.01	
	5	3.1	0.20	0.25	0.09	0.15	0.29	0.26	0.25	0.01	0.02	
	6	3.3	0.47	0.70	0.41	0.70	0.67	0.42	0.78	0.08	0.05	
	7	3.2	1.33	1.60	1.33	1.60	1.60	0.25	0.29	0.01	0.02	
	8	3.5	1.50	1.80	0.35	0.28	0.28	1.29	2.14	0.14	0.21	
	9	5.2	0.97	1.43	0.05	0.32	0.30	0.58	1.43	0.07	0.05	
10	3.4	1.57	1.26	0.57	1.47	1.47	1.29	2.75	0.07	0.06		
<i>geant1</i>	0	1640.8	6.59	8.39	6.54	8.33	8.47	5.73	57.26	57.33	47.89	
	1	853.9	1.00	1.15	1.55	1.23	1.18	3.83	3.83	5.93	8.07	
	2	1053.3	3.06	2.48	3.02	2.74	2.78	19.91	68.53	27.15	49.31	
	3	3517.5	4.62	1.88	2.86	4.45	4.46	8.85	25.20	23.06	22.21	
	4	2141.3	1.18	1.67	1.09	1.53	1.54	3.99	13.80	9.94	13.43	
	5	1757.7	2.02	4.12	2.36	3.36	3.21	2.98	16.05	4.35	9.96	
	6	4587.1	1.33	1.84	2.03	2.32	2.41	6.84	19.13	15.36	18.98	
	7	3352.6	1.30	2.72	2.21	2.16	2.10	4.96	14.92	13.88	13.42	
	8	2067.1	3.44	3.49	2.86	2.94	2.88	7.06	15.19	9.13	12.38	
	9	1837.3	0.97	3.22	3.41	2.83	2.84	4.84	9.93	7.21	5.72	
10	1443.9	2.16	8.33	3.92	6.59	7.05	11.36	24.01	15.56	12.53		
<i>geant2</i>	0	2244.6	3.17	1.37	3.29	2.35	2.40	19.25	68.10	73.08	68.98	
	1	2571.1	1.95	3.28	1.96	3.12	3.19	3.29	9.35	6.47	9.26	
	2	11821.2	1.52	1.74	0.80	1.19	1.21	0.47	0.90	0.89	0.88	
	3	13299.4	0.62	0.49	0.60	0.49	0.51	0.47	0.77	0.73	0.59	
	4	6084.4	1.42	3.17	2.39	3.27	3.34	0.48	1.85	2.32	0.85	
	5	6699.6	2.43	3.56	1.32	1.56	1.54	0.58	1.07	0.79	0.69	
	6	16436.9	1.80	1.81	2.28	1.82	1.86	0.45	1.42	1.43	1.47	
	7	8962	1.57	1.44	1.69	1.52	1.54	0.44	1.53	0.96	0.97	
	8	3792.4	1.27	0.52	1.27	0.50	0.52	1.31	2.23	1.53	1.98	
	9	18969.8	5.97	3.27	6.21	3.15	3.65	1.94	6.90	6.75	7.33	
10	13286.4	5.61	1.34	5.55	1.32	1.33	1.26	5.26	4.54	5.05		
<i>germany17</i>	0	6.7	22.83	4.70	18.39	4.11	4.14	9.88	9.19	1.76	1.65	
	1	25.1	6.15	3.96	6.12	3.44	3.48	5.29	28.26	11.35	10.71	
	2	32.2	0.93	2.25	0.65	0.53	0.52	1.99	2.95	0.22	0.30	
	3	58.6	5.52	10.38	3.67	4.54	4.59	8.16	9.32	3.56	3.59	
	4	51.2	2.86	4.32	0.81	1.58	1.57	2.38	2.04	0.60	0.48	
	5	42.7	3.84	3.62	1.62	1.49	2.53	5.70	1.82	0.66	1.19	
	6	53.5	4.10	5.94	2.01	3.80	2.86	3.56	7.20	1.15	1.29	
	7	91.8	4.86	2.30	4.40	2.76	2.78	5.26	10.87	2.47	3.55	
	8	45.8	2.06	7.31	1.24	1.57	1.50	5.14	5.19	1.53	1.23	
	9	33.6	3.13	1.36	1.13	1.00	0.98	6.23	5.17	1.57	1.70	
10	53.2	3.54	4.90	2.68	3.12	3.03	10.85	16.26	3.40	5.34		

Table A.1: Speed-up factors of algorithms I, IIs, IIIs, IIc-IVc compared to the solution times of CPLEX and the compact flow formulation of the robust network loading problem.