MODELING NETWORK CONGESTION UNDER DEMAND UNCERTAINTY USING WARDROP PRINCIPLES

YASMINE BECK, FRANCESCA GIANCOLA, IVANA LJUBIĆ, AND SARA MATTIA

Abstract. Motivated by the need for reliable traffic management under fluctuating travel demand, we study the problem of determining the worst-case congestion in a multi-commodity traffic network subject to demand uncertainty. To this end, we stress-test a given network by identifying demand realizations and corresponding travelers' route choices that maximize congestion. The users of the traffic network are assumed to act according to one of the two Wardrop principles—the user equilibrium or the system optimum—so that the resulting congestion models can be seen as bilevel problems with a single leader and multiple followers. To address uncertain travel demand, we consider different models such as ellipsoidal or budgeted uncertainty sets and the hose polyhedron. We present single-level mixed-integer nonlinear reformulations of the congestion models that exploit binary variables and big-M constants, prove the existence of optimal solutions, derive valid big-Ms, and propose several enhancement techniques to further strengthen the formulations. An extensive computational study on instances of the Sioux Falls network and instances from the SNDlib demonstrates the computational effectiveness of the proposed techniques and provides insight into the impact of different congestion measures and uncertainty models on the resulting worst-case congestion.

1. Introduction

Congestion is one of the central challenges in urban transportation, leading to increased travel times, fuel consumption, and emissions. To mitigate these effects, effective traffic management requires anticipating how travelers respond to changing network conditions, especially variations in travel demand arising from, e.g., daily fluctuations, seasonal trends, or unforeseen events. Because such variations are inevitable and often difficult to predict accurately, understanding the potential worst-case congestion in a traffic network is crucial.

In this paper, we study the problem of determining the worst-case congestion in a multi-commodity traffic network subject to demand uncertainty. Effectively, we stress-test a given network by identifying demand realizations and corresponding travelers' route choices that maximize congestion. To this end, we propose a novel bilevel formulation in which a traffic planner acts as the leader and the users of the traffic network act as the followers. In the leader's problem, we explicitly model the uncertain travel demand against which the traffic planner wants to hedge. For this purpose, we exploit ideas from robust optimization (Ben-Tal et al. 2009; Bertsimas et al. 2011; Soyster 1973) so that demand realizations are drawn from a predefined uncertainty set. We use both ellipsoidal and polyhedral uncertainty sets, including the hose polyhedron (Duffield et al. 1999; Fingerhut et al. 1997) and the budgeted uncertainty set (Bertsimas and Sim 2003, 2004; Sim 2004). To model travelers' route choices, we consider the two Wardrop principles—the user equilibrium and

Date: March 13, 2026.

2020 *Mathematics Subject Classification.* 90B20,90C33,90C35,90C70.

Key words and phrases. Traffic networks, Wardrop principles, Bilevel optimization, Mixed-integer nonlinear optimization.

the system optimum (Wardrop 1952; Wardrop and Whitehead 1952). In a user equilibrium (UE), travelers select routes that minimize their individual travel costs so that no traveler can improve their travel time by unilaterally changing routes. In contrast, under the system optimum (SO), a central planner coordinates or assigns traffic such as to minimize the total travel time across the network. Overall, we thus consider a single-leader multi-follower bilevel problem in which Wardrop principles govern the followers' decisions. An accessible overview of the two Wardrop principles can, e.g., be found in Correa and Stier-Moses (2011). Moreover, we refer to Bard (1998), Dempe (2002), and Dempe et al. (2015), and the surveys in Colson et al. (2005, 2007) and Kleinert et al. (2021) for a general overview of bilevel optimization. Let us further mention that bilevel problems are notoriously hard to solve. Hansen et al. (1992) have shown that even linear bilevel problems, i.e., those with continuous variables, linear objective functions, and linear constraints, are strongly NP-hard in general.

Building on the modeling choices described above and acknowledging the computational challenges of bilevel optimization, the congestion models studied in this paper are reformulated as mixed-integer nonlinear problems that exploit binary variables and big-M constants, which can be tackled using state-of-the-art general-purpose solvers. We prove the existence of optimal solutions to these problems, derive valid big-Ms, and propose several enhancement techniques that exploit the network topology and structural properties of optimal flows to further strengthen the formulations. To measure congestion, we consider three different functions: the maximum utilization function, which focuses on the most congested arc; the total utilization function, which accounts for congestion on all arcs of the network; and the Bureau of Public Roads (BPR) function, which penalizes arcs with very high congestion levels. To compare the different variants of the congestion model and to assess the effectiveness of the proposed enhancement techniques, we conduct an extensive computational study on instances of the Sioux Falls network (LeBlanc et al. 1975) and instances from the SNDlib (Orlowski et al. 2007, 2010). We analyze both user equilibrium and system optimum formulations in terms of congestion levels and computational performance. Moreover, we compare different problem variants arising from the choice of congestion measures and uncertainty models, as well as from different travel cost functions, including linear, quadratic, and BPR cost functions. Our computational study reveals several key insights. First, the proposed enhancement techniques significantly improve computational performance, enabling the solution of larger instances that would otherwise remain unsolved. Second, the choice of the travel cost function has a greater impact on the computational tractability of the congestion models than the choice of the congestion measure. Third, the UE formulation can be solved considerably faster than the SO formulation, while congestion levels under the UE are typically close to those under the SO. Fourth and finally, we observe a trade-off between worst-case congestion and computational cost across different uncertainty models, as congestion estimates over larger uncertainty sets are typically obtained at increased computational effort.

Related Literature. Many applications of bilevel optimization arise in the context of transportation; see, e.g., Migdalas (1995) for a general overview. In particular, bilevel models have been studied in network design (Ben-Ayed et al. 1988; Fontaine and Minner 2014; LeBlanc and Boyce 1986; Marcotte 1986; Rey and Levin 2024), network pricing (Beck et al. 2024; Brotcorne et al. 2001; Dempe and Zemkoho 2012; Dewez et al. 2008; Kalashnikov et al. 2020; Labbé et al. 1998), or pricing with routing (Cerulli et al. 2024). Among these contributions, only a few works consider bilevel formulations in which the lower level is modeled using Wardrop principles. To the best of our knowledge, Beck et al. (2024), Dempe and Zemkoho (2012),

and Rey and Levin (2024) are the closest related studies in this context. In Rey and Levin (2024), the authors study the discrete network design problem in which travelers' act according to the Wardrop user equilibrium principle. To solve the resulting bilevel model, the authors present a branch-and-price-and-cut approach. In Dempe and Zemkoho (2012), the authors consider a bilevel toll-setting problem in which travelers again follow the Wardrop user equilibrium principle. In their work, the authors focus primarily on the theoretical properties of the model, but a closely related toll-setting problem is studied in Beck et al. (2024) from a computational point of view. In addition, Beck et al. (2024) incorporate robust Wardrop equilibria to model travelers' route choices subject to uncertain travel costs. Overall, our work thus differs from the above contributions in two main aspects. First, while Rey and Levin (2024) address network design and Beck et al. (2024) and Dempe and Zemkoho (2012) focus on network pricing, we study the problem of determining the worst-case congestion that can arise in a given traffic network. Second, we explicitly account for uncertainty in the travel demand, whereas Beck et al. (2024) consider uncertainty in the travel costs and Dempe and Zemkoho (2012) and Rey and Levin (2024) study deterministic settings. To the best of our knowledge, no existing work considers the problem of stress-testing a traffic network under uncertain travel demand using a bilevel formulation with Wardrop principles governing the travelers' behavior. Although we focus on existing traffic networks, let us mention that the insights provided by our models may also be useful in network design contexts, for example to identify which parts of a network are most susceptible to congestion under uncertain travel demand.

Outline. The remainder of this paper is organized as follows. In Section 2, we define the congestion models under the user equilibrium and the system optimum. In Section 3, we present mixed-integer nonlinear reformulations of these problems that exploit binary variables and big-M constants and prove the existence of valid big-Ms and optimal solutions. In Section 4, we propose several enhancement techniques to strengthen the formulations of the congestion models. In Section 5, we discuss further practical modeling techniques and elaborate on our specific choice of congestion measures, travel cost functions, and uncertainty models used in our computational study. In Section 6, we describe the experimental setup and the design of the computational study. The results of this study are discussed in Section 7. Finally, we derive conclusions in Section 8.

2. Problem Statement

We study the problem of determining the worst-case congestion in a traffic network under uncertain travel demand. To this end, we first introduce the network model and discuss how to account for congestion and demand uncertainty in Section 2.1. Afterward, in Section 2.2, we elaborate on the behavior of the network users, who are assumed to act according to one of the two Wardrop principles—the user equilibrium or the system optimum.

2.1. Congestion Model.

To model the multi-commodity traffic network, we consider a directed graph $G = (V, A)$ with node set V and arc set $A \subseteq V \times V$. Moreover, node subsets $S \subseteq V$ and $T \subseteq V$ denote the sets of origin and destination nodes, respectively. The set of all commodities to be routed through the network is given by $K \subseteq S \times T$ and each commodity $k \in K$ has a fixed demand $d_k \in \mathbb{R}_{\geq 0}$ to travel from its origin to its destination. For the ease of presentation, we consider a single commodity for each origin-destination (OD) pair and all other pairs of nodes are

assumed to have zero demand. For the remainder of this paper, we make the following connectivity assumption, which is standard in the transportation literature; see, e.g., Assumption 2.A in Patriksson (2015).

Assumption 1. *For every commodity $k = (s_k, t_k) \in K$, there exists at least one dipath that connects s_k and t_k.*

Let $f = (f_a)_{a \in A} \in \mathbb{R}^{|A|}$ denote the vector of all arc flows used to model traffic. To assess the network under adverse conditions, we consider the worst-possible realization of the travel demand and the corresponding travelers' route choices. This leads us to considering the problem

$$\max_{d,f} \quad L(f) \quad \text{s.t.} \quad d \in \mathcal{U},\ f \in F(d). \tag{1}$$

Here, we assume that the travel demand $d = (d_k)_{k \in K}$ is not known exactly but that it takes values within a given uncertainty set $\mathcal{U}$. Moreover, we use $F(d)$ to denote the set of traffic flows corresponding to the optimal travelers' responses for a given demand realization d. For the uncertainty set, we impose the following throughout the remainder of this paper.

Assumption 2. *The set of origin-destination pairs K is fixed across all realizations of the uncertain demand. Moreover, the uncertainty set $\mathcal{U}$ is non-empty, convex, compact, and contained in the nonnegative orthant $\mathbb{R}^{|K|}_{\geq 0}$.*

We assume that the uncertainty set $\mathcal{U}$ is non-empty and compact to ensure the existence of optimal solutions to the congestion models. Moreover, we assume convexity to obtain formulations that are more computationally tractable and amenable to state-of-the-art general-purpose solvers. In Assumption 2, we further require that $\mathcal{U}$ is contained in the nonnegative orthant and that the set of commodities is fixed, which guarantees the existence of optimal flows for all demand realizations. Note that, if the travel demand is perfectly known, i.e., if the uncertainty set $\mathcal{U}$ is a singleton, Problem (1) reduces to a deterministic traffic assignment problem aimed at determining the worst-case congestion. Throughout this paper, we refer to this setting as the *deterministic case* and use it as a baseline for assessing the impact of demand uncertainty on network congestion.

For the latency function $L(f)$ used to measure congestion, we make the following assumption.

Assumption 3. *The latency function $L : \mathbb{R}^{|A|}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $f \mapsto L(f)$ is continuous.*

Next, we elaborate on the Wardrop principles used to model traffic flow.

2.2. **Wardrop Principles.** For computational tractability, we adopt a fully aggregated node-arc formulation in which all commodities with the same origin are treated as a unique commodity with a single source and multiple destinations. In the transportation literature, it is well known that such an aggregation preserves the set of optimal flows, provided there are no commodity-dependent costs; see, e.g., Section 5.2.3 in Boyles et al. (2025), Bienstock et al. (1998), and Chouman et al. (2017) in which similar settings are studied. In what follows, we model aggregated commodity flows using variables $x = (x^s)_{s \in S}$ with $x^s = (x^s_a)_{a \in A} \in \mathbb{R}^{|A|}_{\geq 0}$ for all $s \in S$. The overall arc flows are then given by

$$f = \sum_{s \in S} x^s. \tag{2}$$

Moreover, for every source $s \in S$, we define the set of commodities that originate in s as

$$K_s := \{k = (s_k, t_k) \in K \colon s_k = s\}.$$

In what follows, we use $\delta^{\text{in}}(v)$ and $\delta^{\text{out}}(v)$ to denote the sets of in- and outgoing arcs of node $v \in V$, respectively. Flow conservation can then be stated as

$$\sum_{a \in \delta^{\text{out}}(v)} x_a^s - \sum_{a \in \delta^{\text{in}}(v)} x_a^s = d_v^s, \quad v \in V,\, s \in S, \tag{3}$$

with

$$d_v^s = \begin{cases} +\sum_{k \in K_s} d_k, & v = s, \\ -d_k, & k = (s, v) \in K_s, \\ 0, & \text{otherwise}, \end{cases}$$

for a given demand realization $d \in \mathcal{U}$. With the constraints in (2) and (3) at hand, which characterize feasible flows in the traffic network, we now extend the model to incorporate travel costs. These costs determine how users choose their routes and, consequently, how flows distribute across the network. For each arc $a \in A$, we consider a travel cost function

$$c_a : \mathbb{R}_{\geq 0}^{|A|} \to \mathbb{R}_{>0}, \quad f \mapsto c_a(f_a),$$

which captures the time required to traverse the arc as well as additional expenses such as fuel consumption or congestion-induced delays. In line with standard terminology, we use the terms travel time and travel cost interchangeably. Moreover, the following will be a standing assumption for the remainder of this paper.

Assumption 4. *For every arc $a \in A$, the function $c_a : \mathbb{R}_{\geq 0}^{|A|} \to \mathbb{R}_{>0}$, $f \mapsto c_a(f_a)$ is convex, non-decreasing, and continuously differentiable in f.*

In particular, we assume that travel costs are separable, i.e., the cost of traversing an arc $a \in A$ only depends on the flow on that arc but not on the flows on other arcs of the network. Assumption 4 is satisfied by many travel cost functions studied in the literature such as, e.g., the BPR function; see also Section 1.5 in Patriksson (2015) for further discussions. Let us now formalize the behavior of users in the traffic network. To this end, we follow the Wardrop principles (Wardrop 1952; Wardrop and Whitehead 1952) and distinguish between the user equilibrium and the system optimum. For both settings, we assume that travelers have complete knowledge of available paths and that traffic flows remain stable over time.

User Equilibrium. In the user equilibrium, travelers act selfishly by selecting their routes such as to minimize their own travel times. This behavior leads to a stable state in which no commodity can reduce their travel costs by unilaterally changing routes. Consequently, in a Wardrop user equilibrium, all paths used by a given commodity have equal travel costs. For separable travel cost functions, the UE can be obtained as an optimal solution to the problem

$$\min_{f, x \geq 0} \quad \sum_{a \in A} \int_0^{f_a} c_a(\xi)\, \mathrm{d}\xi \quad \text{s.t.} \quad (2),\, (3). \tag{4}$$

Hence, under the user equilibrium, the set $F(d)$ corresponds to the set of optimal solutions to Problem (4). A similar problem is, e.g., considered in Section 3.6.2 in Ferris and Pang (1997).

System Optimum. The system optimum seeks to minimize the total travel time across the network by coordinating the flows of all travelers. This corresponds to a situation in which the marginal travel times are the same on all routes used by a given commodity. Otherwise, shifting flow from a higher- to a lower-cost route

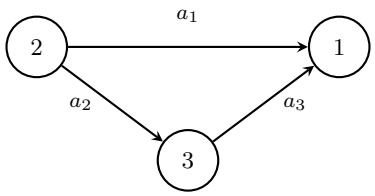


FIGURE 1. The traffic network considered in Example 1.

would reduce the overall travel time. The SO can be obtained as an optimal solution to the problem

$$\min_{f,x\geq 0} \quad \sum_{a\in A} c_a(f_a)f_a \quad \text{s.t.} \quad (2),\ (3). \tag{5}$$

Hence, under the system optimum, the set $F(d)$ corresponds to the set of optimal solutions to Problem (5). For further discussion of the SO, see also Section 2.4 in Patriksson (2015).

To sum up, for each demand realization $d \in \mathcal{U}$, both the UE and the SO can be obtained by solving appropriately chosen nonlinear optimization problems (NLPs). The two models share the same linear feasibility constraints, i.e., the only nonlinearities arise in the respective objective functions. As a result, the Abadie constraint qualification is satisfied for all feasible points. Moreover, under Assumption 4, Problems (4) and (5) are convex NLPs, so that the Karush–Kuhn–Tucker (KKT) conditions are both necessary and sufficient optimality conditions. Finally, let us mention that the main difference between the UE and SO models lies in their behavioral assumptions. The UE reflects decentralized, selfish route choices by individual travelers, whereas the SO represents a centralized assignment of traffic flows that minimizes total travel time.

2.3. **Congestion Ratio.** To compare the congestion levels under the user equilibrium and the system optimum, we introduce the so-called congestion ratio (CR), which we define as

$$\mathrm{CR} = \frac{L(f^{\mathrm{UE}})}{L(f^{\mathrm{SO}})}.$$

Here, f^{UE} and f^{SO} denote optimal arc flows in the UE and the SO formulation, respectively. A congestion ratio of 1 indicates that the user equilibrium solution achieves the same overall congestion as the system optimum. Values greater than 1 indicate higher congestion under the UE, whereas values lower than 1 indicate higher congestion under the SO.

Note that the congestion ratio differs from the price of anarchy (Koutsoupias and Papadimitriou 1999), a widely used metric in the transportation literature for comparing UE and SO solutions. The price of anarchy measures the efficiency loss caused by selfish routing by comparing aggregated travel costs under a given demand realization and is, by definition, always at least 1. In contrast, the CR focuses on congestion rather than total travel cost and may compare UE and SO flows corresponding to different demand realizations in our framework. As a result, the CR captures worst-case congestion levels and, unlike the price of anarchy, can take values below 1. We illustrate such a behavior with the following example.

Example 1 ($\mathrm{CR} < 1$)**.** *We consider the network depicted in Figure 1 with a single commodity $k = (2, 1)$ whose demand is uncertain and takes values in the interval $[30, 50]$. For all arcs $a \in A = \{a_1, a_2, a_3\}$, we use the travel cost function*

$c_a(f_a) = 1 + 0.15(f_a/20)^4$ and measure congestion as $L(f) = \sum_{a\in A} f_a$, which captures the overall traffic load. In this network, there are two paths from the origin node 2 to the destination node 1, namely $2 \to 1$ and $2 \to 3 \to 1$.

Under the UE formulation, the worst-case demand is $d^{UE} = 50$ with optimal flows $f^{UE}_{a_1} = 33.2$ and $f^{UE}_{a_2} = f^{UE}_{a_3} = 16.8$, which results in a congestion value of $L(f^{UE}) = 3.34$. Under the SO formulation, the worst-case demand is still $d^{SO} = 50$ but the optimal flows are $f^{SO}_{a_1} = 28.4$ and $f^{SO}_{a_2} = f^{SO}_{a_3} = 21.6$, which yields $L(f^{SO}) = 3.58$. Hence, the congestion ratio is

$$CR = \frac{L(f^{UE})}{L(f^{SO})} = \frac{3.34}{3.58} = 0.93 < 1,$$

which shows that the UE solution can achieve lower worst-case congestion than the SO solution, a behavior that cannot occur with the price of anarchy. Note that, in this example, both formulations yield the same worst-case demand realization, but this does not have to be the case in general.

3. Single-Level Reformulations

Because of its nested structure, the congestion model (1) is intrinsically hard to solve. Motivated by common approaches in bilevel optimization, we thus reformulate it as a single-level problem, which can then be tackled using state-of-the-art general-purpose solvers. Under Assumption 4, the lower-level problems (4) and (5) are convex d-parameterized NLPs, which allows to use their KKT conditions to obtain equivalent single-level reformulations. We present the single-level reformulations of the congestion models under the user equilibrium and the system optimum in Sections 3.1 and 3.2, respectively. As the resulting problems are mathematical problems with equilibrium constraints (MPECs)—see, e.g., Luo et al. (1996) for a general overview—we further reformulate them as mixed-integer nonlinear problems (MINLPs), following the same ideas as in Fortuny-Amat and McCarl (1981). To this end, we introduce auxiliary binary variables and sufficiently large big-M constants, whose existence we derive in Section 3.3. Finally, we prove the existence of optimal solutions to the congestion models in Section 3.4.

3.1. Congestion Model under the User Equilibrium.

For a given demand realization $d \in \mathcal{U}$, the KKT conditions of Problem (4) can be stated as

$$f = \sum_{s\in S} x^s, \tag{6a}$$

$$\sum_{a\in\delta^{\text{out}}(v)} x^s_a - \sum_{a\in\delta^{\text{in}}(v)} x^s_a = d^s_v, \qquad v \in V,\ s \in S, \tag{6b}$$

$$0 \le c_a(f_a) + \tau^s_i - \tau^s_j \perp x^s_a \ge 0, \qquad a = (i,j) \in A,\ s \in S. \tag{6c}$$

Here, the variables $\tau = (\tau^s)_{s\in S}$ with $\tau^s = (\tau^s_v)_{v\in V} \in \mathbb{R}^{|V|}$ correspond to the Lagrangian multipliers associated with the flow conservation constraints in (3). Because the KKT conditions are necessary and sufficient for Problem (4), we can replace Problem (4) with System (6) to obtain the single-level reformulation of the congestion model (1) given by

$$\max_{d,f,x,\tau} \quad L(f) \quad \text{s.t.} \quad d \in \mathcal{U},\ (6). \tag{7}$$

Problem (11) is an MPEC because of the complementary constraints in (6c). Nevertheless, we can exploit the disjunctive nature of these constraints to derive an MINLP reformulation of the overall congestion model. To this end, we introduce auxiliary binary variables $y^s_a \in \{0,1\}$ and sufficiently large constants $M^s_a, N^s_a \in \mathbb{R}_{\ge 0}$

for all arcs $a \in A$ and all origin nodes $s \in S$ so that we can equivalently re-write the constraints in (6c) as

$$0 \leq c_a(f_a) + \tau_i^s - \tau_j^s \leq M_a^s(1 - y_a^s), \qquad a = (i,j) \in A,\ s \in S, \tag{8a}$$

$$0 \leq x_a^s \leq N_a^s y_a^s, \qquad a \in A,\ s \in S. \tag{8b}$$

For $a \in A$ and $s \in S$, the binary variable y_a^s indicates whether any flow originating in s uses arc a. The MINLP reformulation of Problem (1) then reads

$$\max_{d,f,x,\tau,y} \quad L(f) \quad \text{s.t.} \quad d \in \mathcal{U},\ (6a),\ (6b),\ (8),\ y_a^s \in \{0,1\},\ a \in A,\ s \in S. \tag{9}$$

In general, Problem (9) is a nonconvex MINLP because the constraints in (8a) are nonlinear and nonconvex for general convex travel cost functions; cf. Assumption 4. Nevertheless, we emphasize that the structural properties of Problem (9) depend strongly on the specific choice of travel cost and latency functions. For instance, if both are linear, Problem (9) is a convex mixed-integer linear problem (MILP). We elaborate on problem variants in Section 5, where we study different choices of travel cost and latency functions in more detail.

3.2. **Congestion Model under the System Optimum.** For a given demand realization $d \in \mathcal{U}$, the KKT conditions of Problem (5) can be stated as

$$f = \sum_{s \in S} x^s, \tag{10a}$$

$$\sum_{a \in \delta^{\text{out}}(v)} x_a^s - \sum_{a \in \delta^{\text{in}}(v)} x_a^s = d_v^s, \qquad v \in V,\ s \in S, \tag{10b}$$

$$0 \leq c_a(f_a) + f_a c_a'(f_a) + \tau_i^s - \tau_j^s \perp x_a^s \geq 0, \qquad a = (i,j) \in A,\ s \in S. \tag{10c}$$

Here, $c_a'(f_a)$ denotes the derivative of the travel cost function $c_a(f_a)$ for $a \in A$. Because the UE formulation (4) and the SO model (5) only differ in their objective functions, the corresponding KKT conditions only differ in (6c) and (10c). As before, we obtain a single-level reformulation of the congestion model (1) by replacing Problem (5) with its necessary and sufficient KKT conditions. This yields

$$\max_{d,f,x,\tau} \quad L(f) \quad \text{s.t.} \quad d \in \mathcal{U},\ (10). \tag{11}$$

Again, we introduce auxiliary binary variables $y_a^s \in \{0,1\}$ and sufficiently large constants $M_a^s, N_a^s \in \mathbb{R}_{\geq 0}$ for all $a \in A$ and $s \in S$ to obtain an equivalent MINLP reformulation of (11), which is given by

$$\max_{d,f,x,\tau,y} \quad L(f) \tag{12a}$$

$$\text{s.t.} \quad d \in \mathcal{U},\ (10a),\ (10b), \tag{12b}$$

$$0 \leq c_a(f_a) + f_a c_a'(f_a) + \tau_i^s - \tau_j^s \leq M_a^s(1 - y_a^s), \qquad a \in A,\ s \in S, \tag{12c}$$

$$0 \leq x_a^s \leq N_a^s y_a^s, \qquad a \in A,\ s \in S, \tag{12d}$$

$$y_a^s \in \{0,1\}, \qquad a \in A,\ s \in S. \tag{12e}$$

3.3. **Variable Bounds and Big-Ms.** Note that the equivalence of Problems (9) and (12) to the original congestion models under the UE and the SO relies on choosing appropriate values for the constants M_a^s, N_a^s for all $a \in A$ and $s \in S$. In this section, we show that such sufficiently large constants exist by proving valid bounds for the variables f, x, and τ.

Proposition 1. *Let $d \in \mathcal{U}$ be given arbitrarily. Then, under Assumptions 1, 2, and 4, Problems (4) and (5) each admit an optimal solution (f, x) that satisfies*

$$0 \leq x_a^s \leq \sum_{k \in K_s} d_k, \quad a \in A,\ s \in S,$$

as well as

$$0 \leq f_a \leq \sum_{s \in S} \sum_{k \in K_s} d_k, \quad a \in A.$$

Proof. Because the travel cost functions are positive and non-decreasing under Assumption 4, any positive flow on a cycle would increase the objective function value and therefore cannot occur in an optimal solution; cf., e.g., Theorem 3.8 in Ahuja et al. (1993). Hence, we can, w.l.o.g., add the constraints $0 \leq x_a^s \leq \sum_{k \in K_s} d_k$ for all $a \in A$ and $s \in S$ to Problems (4) and (5) without affecting their sets of optimal solutions. As the variables x are linearly coupled to the arc flows f via (2), this directly implies finite bounds for f, i.e., all flow variables are bounded. The feasibility of Problems (4) and (5) follows from Assumptions 1 and 2. Moreover, as the feasible sets are described by finitely many linear constraints, they are compact. Finally, because the objective functions of the UE and SO problems are continuous, the Weierstrass theorem yields the existence of optimal solutions to Problems (4) and (5); see, e.g., Theorem 2.4 in Patriksson (2015) for similar arguments. □

Proposition 2. *Let $d \in \mathcal{U}$ be given arbitrarily and suppose that Assumptions 1, 2, and 4 hold. Then, for every $(f, x) \in F(d)$, there exists τ such that $0 \leq \tau_v^s < \infty$ holds for all $v \in V$ and $s \in S$, and (f, x, τ) solves (6) or (10) under the user equilibrium or the system optimum, respectively.*

Proof. The claim can be shown in analogy to Remark 1 and the proof of Proposition 3 in Beck et al. (2024). □

Finally, we mention that sufficiently large constants M_a^s and N_a^s, $a \in A$, $s \in S$, to be used in Problems (9) and (12) can be obtained by exploiting Propositions 1 and 2 together with Assumption 4 and the specific structure of the travel cost functions; see, e.g., Beck et al. (2024) for further details.

3.4. **Existence of Solutions.** To conclude this section, we now prove the existence of optimal solutions to the congestion model (1) under the UE and the SO.

Theorem 1. *Under Assumptions 1–4, the congestion model (1) admits an optimal solution under both the user equilibrium and the system optimum.*

Proof. The function L is continuous because of Assumption 3. Moreover, the set $\mathcal{U}$ is non-empty and compact due to Assumption 2 and, by Propositions 1 and 2, the set $F(d)$ is non-empty and bounded for all $d \in \mathcal{U}$. In addition, the graph of the set-valued map F is closed as it is described by finitely many continuous equality and inequality constraints. Applying the Weierstraß theorem completes the proof. □

4. Strengthened Formulations

We now describe the main strategies used to improve the computational tractability of the congestion model (1). These techniques reduce the problem size and strengthen the formulation by exploiting the network topology and structural properties of optimal flows. In Section 4.1, we discuss graph partitioning techniques based on cut nodes and blocks that can be used to derive tighter bounds for the flow variables f and x, which is what we do in Section 4.2. In Section 4.3, we present valid inequalities that eliminate cycle flows in the traffic network.

4.1. **Graph Partitioning.** Given the directed graph $G = (V, A)$, we denote its underlying simple undirected graph as $\underline{G} = (V, E)$ in which each arc (i, j) is replaced by edge $\{i, j\}$ and no multiple edges are allowed. Before proceeding, we introduce some basic graph-theoretic concepts that will be used throughout this section; see, e.g., Chapter 4 in West (2001) for further details.

Definition 1 (Connected Component)**.** *A* connected component *of* $\underline{G}$ *is a maximal subset* $C \subseteq V$ *such that, for all* $u, v \in C$*, there exists a path connecting* u *and* v*.*

Definition 2 (Articulation Node)**.** *An* articulation node *(or* cut node*) is a node whose removal increases the number of connected components of* $\underline{G}$*.*

Definition 3 (Bridge)**.** *An edge* $e \in E$ *is a* bridge *(or* cut edge*) if the graph* $(V, E \setminus \{e\})$ *has more connected components than* $\underline{G}$*.*

Definition 4 (2-Connected Graph)**.** *An undirected graph* $\underline{G}$ *is said to be* 2-connected *(or* node-biconnected*) if it is connected, has at least three nodes, and contains no cut node.*

Definition 5 (Block)**.** *A block* $B = (V_B, E_B)$ *of* $\underline{G}$ *is a maximal connected subgraph of* $\underline{G}$ *that has no cut node. If a block has a single edge, it is called a* trivial block *or* bridge*, otherwise it is called a* non-trivial block*.*

Note that $\underline{G}$ itself is a block if it is connected and has no cut node. Moreover, we emphasize that every edge of $\underline{G}$ belongs to exactly one block and that two blocks share at most one node, which must be an articulation node.

Definition 6 (Block-Cut Tree)**.** *The* block-cut tree *of a connected graph* $\underline{G}$ *is a tree* T *whose node set consists of the blocks and articulation nodes of* $\underline{G}$*, with an edge between a block* B *and an articulation node* v *whenever* $v \in B$*.*

For a set of nodes $S \subseteq V$, we denote by $\underline{G}[S]$ and $G[S]$ the subgraphs of $\underline{G}$ and G induced by S, respectively. In the following, we consider graph decomposition techniques based on blocks and articulation nodes. Such decompositions are not only theoretically appealing but also relevant for real-world networks. Empirical evidence indicates that many practical networks contain a non-negligible fraction of articulation points, e.g., Tian et al. (2017) document substantial articulation structure in U.S. road networks, underscoring the practical relevance of such decomposition techniques.

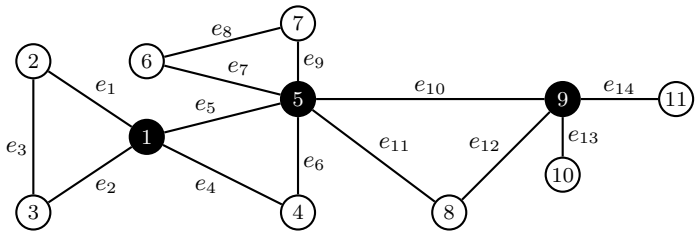
Example 2. *We consider the undirected graph* $\underline{G} = (V, E)$ *with node set* $V = \{1, \dots, 11\}$ *and edge set* $E = \{e_1, \dots, e_{14}\}$ *illustrated in Figure 2 (a). This graph has three articulation nodes,* $c_1 = 1$*,* $c_2 = 5$*, and* $c_3 = 9$*, and six blocks:*

$$B_1 = (\{1,2,3\}, \{e_1, e_2, e_3\}),\ B_2 = (\{1,4,5\}, \{e_4, e_5, e_6\}),$$
$$B_3 = (\{5,6,7\}, \{e_7, e_8, e_9\}),\ B_4 = (\{5,8,9\}, \{e_{10}, e_{11}, e_{12}\}),$$
$$B_5 = (\{9,10\}, \{e_{13}\}),\ B_6 = (\{9,11\}, \{e_{14}\}).$$

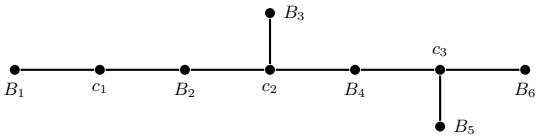
Here, B_1*,* B_2*,* B_3*, and* B_4 *are non-trivial blocks (2-connected subgraphs), whereas* B_5 *and* B_6 *are trivial blocks (bridges). The corresponding block-cut tree is shown in Figure 2 (b).*

Under Assumption 4, UE and SO solutions cannot have positive flow on directed cycles; see, e.g., Theorem 3.8 in Ahuja et al. (1993). This property also holds for the source-aggregated formulation in Problems (4) and (5); see, e.g., Section 5.2.3 in Boyles et al. (2025) for further details. The absence of directed cycle flows allows us to strengthen the formulation of the congestion models by fixing the values of certain variables in a preprocessing step. Specifically, any dipath that exits a block through an articulation node and re-enters it through the same node forms a directed cycle. Hence, all optimal flows with origin and destination nodes in the same block can be restricted to dipaths entirely contained in that block.

Proposition 3. *Suppose that Assumption 4 holds. Let* $B = (V_B, E_B) \subseteq \underline{G}$ *be a block and let* $u, v \in V_B$ *be such that there exists a dipath from* u *to* v *in* G*. Then, the following statements hold.*



(a) Graph $\underline{G}$



(b) Block-cut tree T

FIGURE 2. The graph considered in Example 2 with articulation nodes shown in black (a) and its block-cut tree (b).

(i) There exists at least one dipath from u to v in G that is entirely contained in $G[V_B]$.
(ii) Any dipath from u to v used in an optimal UE or SO solution is entirely contained in $G[V_B]$.

Proof. We first recall a key property of block-cut trees that will be used in the proof. To this end, let T be a block-cut tree of $\underline{G}$, i.e., a block B is a node of T and its neighbors in T are precisely the cut nodes of $\underline{G}$ contained in B. If a path in $\underline{G}$ leaves B through a node $w \in V_B$, then w must be a cut node. This is because the edge connecting w to a node outside of V_B belongs to a different block B' and $w \in B \cap B'$ implies that w is a cut node. Once the path leaves B through w, it continues within the subtree of $T - B$ that contains w. Because T is a tree, the only way for the path to return to B is through node w. Because every arc in G corresponds to an edge in $\underline{G}$, the same property holds for dipaths in G: any dipath that exits $G[V_B]$ must re-enter through the same cut node w, thereby forming a directed cycle through w.

We now prove the two statements. By assumption, there exists a dipath from u to v in G. As shown above, each time this dipath leaves $G[V_B]$, it must re-enter through the same cut node, forming a directed cycle. Removing all such cycles yields a dipath from u to v that is entirely contained in $G[V_B]$, which proves (i). Under Assumption 4, optimal UE and SO solutions do not carry flow on directed cycles. Due to (i), any dipath that leaves and later re-enters $G[V_B]$ forms a directed cycle. Hence, all dipaths carrying positive flow in an optimal solution must be entirely contained in $G[V_B]$, which proves (ii). □

By Part (ii) of Proposition 3, we can a priori fix certain variables of the congestion model (1) to zero. More formally, we have the following result.

Corollary 1. *Suppose that Assumptions 1–4 hold. Let (d, f, x, τ, y) be an optimal solution to either Problem (9) or Problem (12). Further, let $s \in S$ be given arbitrarily and let B_s denote a block of $\underline{G}$ containing s. For each $k = (s, t_k) \in K_s$, let B_{t_k} be a block of $\underline{G}$ containing t_k and let $\mathcal{B}_k$ be the set of blocks on the unique*

path from B_s to B_{t_k} in the block-cut tree T of $\underline{G}$. Moreover, let

$$\mathcal{B}_s = \bigcup_{k \in K_s} \mathcal{B}_k \quad \text{and} \quad A_{\mathcal{B}_s} = \{(i,j) \in A \colon \{i,j\} \in E_B \text{ for some } B \in \mathcal{B}_s\}.$$

Then, it holds

$$x_a^s = 0,\ y_a^s = 0, \quad a \notin A_{\mathcal{B}_s}.$$

We can also fix additional variables that would take the value of zero in any optimal solution. For instance, we can a priori fix the variables x_a^s and y_a^s for all arcs $a \in A$ that are not reachable from an origin $s \in S$.

Lemma 1. *Let $d \in \mathcal{U}$ be given arbitrarily. For each $s \in S$, let the set of nodes that are reachable from source s be given by*

$$V_s = \{v \in V \colon \text{there exists a } s\text{–}v \text{ dipath}\}.$$

Then, for any $(f,x) \geq 0$ that satisfies (2) and (3), it holds $x_a^s = 0$ for all $s \in S$ and $a = (i,j) \in A$ with $i \notin V_s$.

Proof. The claim immediately follows from the definition of the set V_s, $s \in S$. □

We use Lemma 1 to fix some of the binary variables y_a^s in Problems (9) and (12), which indicate whether arc $a \in A$ is used by any commodity with origin $s \in S$.

Corollary 2. *Suppose that Assumptions 1–4 hold. Let (d,f,x,τ,y) be feasible for either Problem (9) or Problem (12). For each $s \in S$, let the set of nodes V_s reachable from source s be defined as in Lemma 1. Then, $y_a^s = 0$ holds for all $s \in S$ and $a = (i,j) \in A$ with $i \notin V_s$.*

Lemma 1 also allows us to reduce the number of constraints in the model by removing the redundant flow conservation constraints

$$\sum_{a \in \delta^{\text{out}}(v)} x_a^s - \sum_{a \in \delta^{\text{in}}(v)} x_a^s = 0, \quad s \in S,\ v \notin V_s,$$

from the original congestion model.

4.2. **Improved Variable Bounds.** Building on the results of Section 4.1, we now compute tighter bounds for the flow variables f and x, which can be used to obtain smaller values for the big-M constants in Problems (9) and (12).

Proposition 4. *Suppose that Assumptions 1, 2, and 4 hold, and let $d \in \mathcal{U}$ and $(f,x) \in F(d)$ be given arbitrarily. For each $s \in S$ and $k \in K_s$, let $\mathcal{B}_k$ be defined as in Corollary 1 and*

$$A_{\mathcal{B}_k} := \{(i,j) \in A \colon \{i,j\} \in E_B \text{ for some } B \in \mathcal{B}_k\}.$$

Moreover, let the set of nodes V_s reachable from source s be defined as in Lemma 1. Then, for all $a = (i,j) \in A$ and $s \in S$, the flows satisfy

$$0 \leq x_a^s \leq \bar{x}_a^s := \sum_{k \in K_s} d_k \mathbf{1}_{\{a \in A_{\mathcal{B}_k}\}},$$

as well as

$$0 \leq f_a \leq \bar{f}_a := \sum_{s \in S} \sum_{k \in K_s} d_k \mathbf{1}_{\{a \in A_{\mathcal{B}_k}\}},$$

where $\mathbf{1}_{\{a \in A_{\mathcal{B}_k}\}}$ is an indicator function that equals 1 if $a \in A_{\mathcal{B}_k}$ and 0 otherwise.

Proof. The claim immediately follows from Proposition 1, Corollary 1, Lemma 1, and the constraints in (2). □

4.3. **Cycle-Elimination Constraints.** Because UE and SO solutions cannot carry positive flow on cycles, we can leverage this property to further strengthen the MINLP reformulations of the congestion models by adding cycle-elimination constraints. To this end, let $\mathcal{R}$ denote the set of all simple directed cycles in the network. Then, for all $r \in \mathcal{R}$ and $s \in S$, the associated cycle-elimination constraint is given by

$$\sum_{a \in r} y_a^s \leq |r| - 1. \tag{13}$$

Here, y_a^s is the binary variable indicating whether arc $a \in r$ is used by any commodity originating in $s \in S$ and $|r|$ is the number of arcs in (or the length of) the cycle $r \in \mathcal{R}$. Constraint (13) ensures that at most $|r| - 1$ arcs can carry positive flow, thereby preventing any directed cycles in an optimal solution. To reduce the number of cycle-elimination constraints added to the model, we exploit Corollary 1 so that Constraint (13) only needs to be imposed for cycles contained within the blocks associated with a source node s.

Proposition 5. *Suppose that Assumptions 1–4 hold. Let $s \in S$ be given arbitrarily and let the set $\mathcal{B}_s$ be the set of blocks associated with s as defined in Corollary 1. For each $B \in \mathcal{B}_s$, let V_B denote the set of nodes in block B. Further, let*

$$\mathcal{R}_s := \{r \in \mathcal{R} \colon V(r) \subseteq V_B \text{ for some } B \in \mathcal{B}_s\}, \tag{14}$$

where $V(r)$ denotes the set of nodes visited by cycle $r \in \mathcal{R}$. Then, the inequalities

$$\sum_{a \in r} y_a^s \leq |r| - 1, \quad r \in \mathcal{R}_s \tag{15}$$

are valid for Problems (9) and (12).

Proof. We prove the claim for the UE as the SO case can be shown analogously. Let (d, f, x, τ, y) be feasible for Problem (9). We prove the claim by contradiction. To this end, suppose that there exist $s \in S$ and $r \in \mathcal{R}_s$ with

$$\sum_{a \in r} y_a^s > |r| - 1.$$

Because $y^s \in \{0,1\}^{|A|}$, this implies $y_a^s = 1$ for all $a \in r$. By Constraints (8a), we then have $c_a(f_a) + \tau_i^s - \tau_j^s = 0$ for all $a = (i,j) \in r$. Summing these equalities over all arcs $a = (i,j) \in r$ yields

$$\sum_{a \in r} c_a(f_a) = 0,$$

which contradicts Assumption 4. This concludes the proof. □

5. Congestion Measures, Travel Cost Functions and Demand Uncertainty Modeling

In the previous sections, we have presented a general framework for modeling and solving the congestion model under demand uncertainty, which is applicable to a wide range of congestion measures, travel cost functions, and uncertainty sets. In this section, we now specify the modeling choices that we will adopt in our computational study in Sections 6 and 7. Specifically, we consider three congestion measures (Section 5.1), three travel cost functions (Section 5.2), and three uncertainty sets (Section 5.3), which are commonly used modeling choices in the transportation literature. We also discuss practical implementation details if relevant.

Throughout this section, we use $0 < u_a < \infty$ to denote the practical capacity of an arc $a \in A$, which represents the maximum flow that an arc can accommodate under ideal operating conditions without causing congestion; see, e.g., Section 1.5 in Patriksson (2015) for further details.

5.1. **Latency Functions.** We measure congestion using one of the following three latency functions. We emphasize that all three considered congestion measures are continuous in the flows f, i.e., they satisfy Assumption 3.

Maximum Utilization Function. We measure congestion by focusing on the most heavily used arc in the network. To this end, we consider the function

$$L(f) = \max_{a \in A} \frac{f_a}{u_a}. \tag{16}$$

In particular, (16) can be used to identify the network's most severe bottleneck.

Total Utilization Function. Whereas the previous function focuses on bottlenecks in the network, we also study the overall traffic load in the network by considering the linear function

$$L(f) = \sum_{a \in A} \frac{f_a}{u_a}. \tag{17}$$

Bureau of Public Roads (BPR) Function. Another congestion measure that is widely used in the literature is the BPR function (U.S. Bureau of Public Roads 1964), which is given by

$$L(f) = \sum_{a \in A} \mathrm{BPR}(f_a) \tag{18}$$

with

$$\mathrm{BPR}(f_a) = c_a^{\mathrm{fix}} \left(1 + \alpha \left(\frac{f_a}{u_a}\right)^{\beta}\right), \quad a \in A. \tag{19}$$

Here, $c_a^{\mathrm{fix}} \in \mathbb{R}_{>0}$ is the free-flow time of arc $a \in A$, i.e., the travel time experienced if the arc is used under ideal operating conditions. Moreover, the parameters $\alpha, \beta \in \mathbb{R}_{>0}$ can be adjusted to account for how quickly travel times increase as flows approach the practical capacity of an arc. Commonly used values are $\alpha = 0.15$ and $\beta = 4$, which reflect typical empirical observations of congestion effects and will also be used in our computational study in Sections 6 and 7. Note that the BPR function (19) is convex for $f \geq 0$ in this case.

We point out that the maximum utilization function (16) and degree-4 polynomials such as those in the BPR function are not directly supported by general-purpose solvers such as Gurobi. To handle these cases, we thus apply the following reformulation techniques.

5.1.1. *Reformulation of the Maximum Utilization Function.* When measuring congestion using the maximum utilization function, we consider the problem

$$\max_{d,f} \max_{a \in A} \frac{f_a}{u_a} \quad \text{s.t.} \quad d \in \mathcal{U},\ f \in F(d). \tag{20}$$

Because the outer maximization corresponds to an optimization problem and the inner maximization is taken over a discrete set of arcs, the resulting max–max structure cannot be reformulated using a standard epigraph reformulation. Nevertheless, an equivalent MINLP reformulation of Problem (20) can be obtained by introducing auxiliary binary variables $w_a \in \{0, 1\}$ for all $a \in A$ as well as a sufficiently large

constant $M \in \mathbb{R}_{>0}$. Specifically, Problem (20) is equivalent to

$$\max_{d,f,x,\lambda} \quad \lambda \tag{21a}$$

$$\text{s.t.} \quad \lambda \leq \frac{f_a}{u_a} + M(1 - w_a), \quad a \in A, \tag{21b}$$

$$\sum_{a \in A} w_a = 1, \tag{21c}$$

$$w_a \in \{0, 1\}, \quad a \in A, \tag{21d}$$

$$d \in \mathcal{U},\ (f, x) \in F(d). \tag{21e}$$

In Problem (21), exactly one of the w variables takes the value 1. If $w_a = 1$ holds for some arc $a \in A$, the corresponding constraint in (21b) enforces $\lambda \leq f_a/u_a$. Because the objective is to maximize λ, this inequality is satisfied with equality for the arc attaining the maximum utilization ratio. All other w variables take the value 0. Nevertheless, to ensure the correctness of the formulation, it is essential to select a sufficiently large constant M. A valid choice is

$$M := \max_{a \in A} \left\{ \frac{\bar{f}_a}{u_a} \right\}, \tag{22}$$

where $\bar{f}_a$ denotes an upper bound on the arc flows; see Proposition 4.

5.1.2. *Handling Degree-4 Polynomials.* Nowadays, general-purpose MILP solvers such as Gurobi can handle convex quadratic programs (MIQPs) but higher-degree polynomials are not directly supported. To address the congestion model with the BPR latency function, we thus introduce auxiliary variables $\phi \in \mathbb{R}^{|A|}_{\geq 0}$ modeling

$$\phi_a = f_a \cdot f_a, \quad a \in A,$$

so that degree-4 polynomials are expressed using only quadratic terms.

5.2. **Travel Cost Functions.** We also use the BPR function presented in (18) for the travel costs. In this paper, we focus on settings with $\alpha = 0.15$ and $\beta \in \{1, 2, 4\}$. For $\beta = 1$, the cost functions are linear; for $\beta = 2$, they are quadratic; and for $\beta = 4$, we obtain the classic BPR function that is commonly studied in the literature and also used as the latency function in this paper. In all three cases, the travel cost functions are convex, non-decreasing, and continuously differentiable in $f \in \mathbb{R}^{|A|}_{\geq 0}$, i.e., they satisfy Assumption 4. To handle the degree-4 polynomials in the travel cost functions with $\beta = 4$, we apply the reformulation described in Section 5.1.2.

Remark 1. *In Problem (12), the derivatives $c'_a(f_a)$ are polynomials of one degree lower than the original travel cost functions $c_a(f_a)$. Hence, the terms $f_a c'_a(f_a)$, $a \in A$, in (12c) are polynomials of the same degree as $c_a(f_a)$. As a result, considering the SO instead of the UE does not affect the structural properties of the MINLP reformulation of the congestion model.*

Given the above reformulation techniques and modeling choices, we now summarize the problem variants arising from different combinations of the latency function and travel cost functions in Table 1. Note that the MINLP reformulations of the congestion model under the user equilibrium (9) and the system optimum (12) have the same structural properties, as discussed in Remark 1.

Finally, let us briefly comment on the monotonicity of the latency function with respect to the travel demand. We note that increasing the travel demand of a single commodity does not necessarily lead to higher congestion levels, even if the travel cost functions are linear. The following example illustrates that higher travel demand may in fact reduce congestion levels due to route reallocation effects. For

TABLE 1. Properties of the MINLP reformulations of the congestion model depending on the choice of $\beta \in \{1, 2, 4\}$ in the BPR travel cost functions and the choice of the latency function $L(f)$, i.e., maximum utilization (16), total utilization (17), or BPR (18).

β	$L(f)$	convex feasible set	problem type
1	(16), (17)	✓	MILP
1	(18)	✓	MIQP
2	(16), (17), (18)	✗	MIQP
4	(16), (17), (18)	✗	MIQP

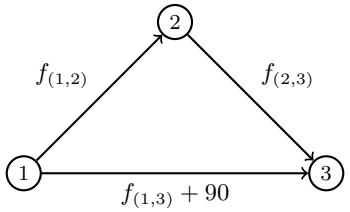


FIGURE 3. The network considered in Example 3. Arc labels represent the travel cost functions associated with each arc.

related monotonicity phenomena and paradoxes in congestion games, we also refer to Cominetti et al. (2024).

Example 3. *We consider the network shown in Figure 3, which is taken from Example 3 in Cominetti et al. (2024). We consider three OD pairs given by* $(1, 2)$, $(2, 3)$, *and* $(1, 3)$ *with demand* 1, 100, *and* 20, *respectively. The resulting equilibrium flows under the UE are* $f_{(1,2)} = 4$, $f_{(1,3)} = 17$, *and* $f_{(2,3)} = 103$. *Let us now assume that all arcs have a practical capacity of* 1, *i.e.,* $u_a = 1$ *for all* $a \in A = \{(1, 2), (2, 3), (1, 3)\}$. *Considering the maximum utilization function* (16) *as the congestion measure, the worst-case congestion is attained on arc* $(2, 3)$, *namely* $f_{(2,3)}/u_{(2,3)} = 103$. *As observed in Cominetti et al. (2024), if we now increase the travel demand of commodity* $(1, 2)$ *from* 1 *to* 4, *some of its flow is reallocated. The new equilibrium flows are given by* $f_{(1,2)} = 6$, $f_{(1,3)} = 18$, *and* $f_{(2,3)} = 102$, *which leads to the worst-case congestion level of* $102 < 103$. *Hence, increasing the travel demand of a commodity can in fact lead to reduced congestion.*

The previous observation implies that the demand realization causing the worst-case congestion in our framework does not necessarily lie on the boundary of the uncertainty set. Due to route reallocation effects, intermediate demand realizations can produce higher congestion, which contrasts with classic robust optimization settings in which worst cases typically occur at the boundary of the uncertainty set.

5.3. **Uncertainty Set.** We assume that the travel demand $d = (d_k)_{k \in K}$ is uncertain. In this paper, we consider three approaches to model this uncertainty: budgeted uncertainty, ellipsoidal uncertainty, and the hose model.

5.3.1. *Budgeted Uncertainty.* As it is unlikely that all demands simultaneously realize in a worst-case manner, a common approach is to adopt a budgeted (or Γ-robust) uncertainty modeling (Bertsimas and Sim 2003, 2004; Sim 2004). This model has been widely used in transportation contexts, including network design (Mattia 2019; Mattia and Poss 2018), traffic assignment (Ito 2011; Ordóñez and Stier-Moses 2007, 2010), and network pricing (Beck et al. 2024).

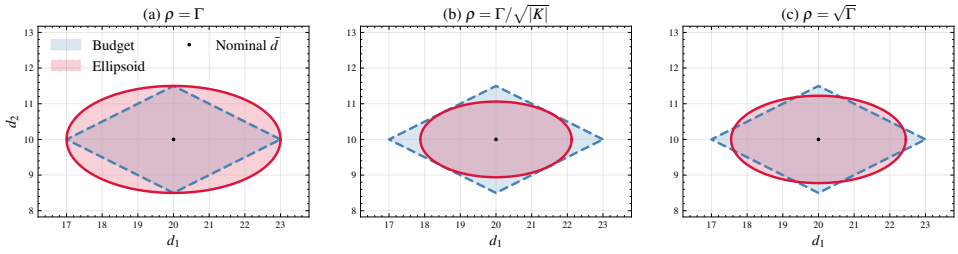


FIGURE 4. Comparison of budgeted and ellipsoidal uncertainty sets.

In the budgeted uncertainty model, each commodity k has a nominal demand $\bar{d}_k \in \mathbb{R}_{\geq 0}$, which corresponds to the demand value provided in the original instance data, and a maximum deviation $\delta_k \in \mathbb{R}_{\geq 0}$ from the nominal value so that $d_k \in [\bar{d}_k - \delta_k, \bar{d}_k + \delta_k]$ holds for all $k \in K$. Each commodity may deviate by an arbitrary fraction of its maximum deviation δ_k from the nominal value as long as the budget $\Gamma \in \{0, \dots, |K|\}$ for the aggregated percentage deviation is not exceeded. Overall, this can be modeled using the uncertainty set

$$\mathcal{U}_{\text{budget}} = \left\{ d \in \mathbb{R}^{|K|} \colon d_k = \bar{d}_k + \delta_k z_k,\ z_k \in [-1, 1],\ k \in K,\ \sum_{k \in K} |z_k| \leq \Gamma \right\}. \quad (23)$$

Here, the parameter Γ controls the size of the uncertainty set. For $\Gamma = 0$, no uncertainty is taken into account, i.e., we consider the deterministic case, whereas for $\Gamma = |K|$, all commodities may simultaneously experience their maximum deviation. Note that $\mathcal{U}_{\text{budget}}$ is non-empty, convex and compact. By assuming $\delta_k \leq \bar{d}_k$ for all $k \in K$, we ensure $d \in \mathbb{R}^{|K|}_{\geq 0}$, i.e., Assumption 2 is satisfied.

5.3.2. *Ellipsoidal Uncertainty.* Using an ellipsoidal uncertainty modeling, we replace the cardinality constraint in the budgeted uncertainty set with a quadratic constraint on the vector of deviations, also called perturbation vector in the following. The ellipsoidal uncertainty set is thus given by

$$\mathcal{U}_{\text{ellipsoid}} = \left\{ d \in \mathbb{R}^{|K|} \colon d_k = \bar{d}_k + \delta_k z_k,\ k \in K,\ \|z\|_2 \leq \rho \right\}, \quad (24)$$

where $\rho > 0$ is the radius of the ellipsoid and $z = (z_k)_{k \in K}$ is the perturbation vector. To have a fair comparison between the ellipsoidal and the budgeted uncertainty models, we select ρ such that the uncertainty sets $\mathcal{U}_{\text{budget}}$ and $\mathcal{U}_{\text{ellipsoid}}$ cover a comparable range of demand realizations. As shown in Li et al. (2011), there exists a geometric relationship between these two sets, which depends on the choice of the parameters ρ and Γ. In their analysis, they describe two cases:

(i) If $\rho = \Gamma$, the ellipsoid is the smallest one containing the polyhedron of the budgeted uncertainty model.
(ii) If $\rho = \Gamma/\sqrt{|K|}$, the ellipsoid is the largest one contained in the polyhedron of the budgeted uncertainty model.

In this paper, we introduce a third intermediate case with $\rho = \sqrt{\Gamma}$, which yields uncertainty sets of comparable size. In Figure 4, we illustrate all three aforementioned cases. As before, we assume that $\delta_k \leq \bar{d}_k$ holds for all $k \in K$ so that the set $\mathcal{U}_{\text{ellipsoid}}$ satisfies Assumption 2.

5.3.3. *Hose Model.* The hose model, originally introduced in the telecommunications network design literature (Duffield et al. 1999; Fingerhut et al. 1997), aggregates demand uncertainty at the node level rather than at the commodity level. This model has been applied in various network design contexts; see, e.g., Altın

et al. (2007), Chekuri et al. (2007), and Mattia (2013). For each node $v \in V$, a bound b_v for the total traffic generated by that node is specified. In this paper, we consider the so-called symmetric hose uncertainty set given by

$$\mathcal{U}_{\text{hose}} = \left\{ d \in \mathbb{R}^{|K|}_{\geq 0} : \sum_{\{k \in K :\, s_k = v \text{ or } t_k = v\}} d_k \leq b_v,\ v \in V \right\}, \tag{25}$$

where s_k and t_k denote the source and the destination node of commodity k, respectively. To ensure a fair comparison between the hose model and the budgeted and ellipsoidal uncertainty sets, we determine the node bounds as follows

$$b_v = \sum_{k \in K_v} \bar{d}_k + \sum_{j=1}^{\min\{\Gamma, |K_v|\}} \delta_k^{(j)},$$

where K_v is the set of commodities with source or destination in node $v \in V$ and $\delta_k^{(j)}$ are the deviations of the commodities in K_v sorted in non-increasing order. Hence, we consider the Γ largest deviations among the commodities in K_v. By construction, we have $\mathcal{U}_{\text{budget}} \subseteq \mathcal{U}_{\text{hose}}$ and $\mathcal{U}_{\text{hose}}$ satisfies Assumption 2.

6. Computational Setup and Experimental Design

In this section, we discuss the setup and design of our computational study. In Section 6.1, we elaborate on the hardware, the software, and the solver used in our computational study. Afterward, in Section 6.2, we describe our test instances.

6.1. Computational Setup.

All experiments were performed on a 64-bit Linux system with an Intel Core i9-13900K (3 GHz) and 128 GB RAM using 4 threads. The congestion models were implemented in Python 3.10.15 and solved with Gurobi 12.0.1. The time limit for each instance was set to 2 h. Moreover, we set the parameter MIPGap to 10^{-3} and disabled presolve, as preliminary experiments indicated numerical instabilities when using Gurobi's default settings. All other parameters were left at their default settings. For each instance, we first solve the deterministic version of the problem and use the solution to this problem as a warm start for solving the model under demand uncertainty, which considerably improved runtime in preliminary computational tests.

6.2. Test Instances.

We consider instances of the Sioux Falls network (LeBlanc et al. 1975), whose data is publicly available at https://github.com/bstabler/TransportationNetworks, and instances from the SNDlib (Orlowski et al. 2007, 2010), which can be accessed at https://sndlib.put.poznan.pl.

6.2.1. *Sioux Falls Instances.*

We consider subnetworks of the Sioux Falls network obtained as subgraphs induced by 50 % or 75 % of the nodes with a varying number of arcs. Specifically, we analyze five small instances with 12 nodes and five medium instances with 18 nodes, which are illustrated in Figure 5. For each subnetwork, we vary the number of commodities $|K| \in \{20, 30, 50, 75, 100\}$ and randomly select the desired number of commodities from the set of all origin-destination pairs. Because the original Sioux Falls instance does not account for demand uncertainty, we do the following. For the Γ-robust formulation, the nominal demand $\bar{d}_k$ for each commodity $k \in K$ is set to the demand value d_k reported in the original instance data. We then set $\delta_k = 0.25\bar{d}_k$ for the maximum deviation. The parameters of the ellipsoidal and hose uncertainty sets are derived from $\bar{d}_k$ and δ_k as described in Section 5.3. All values are scaled by a common factor to avoid numerical instabilities. Finally, we note that all generated subnetworks are strongly connected, whereas

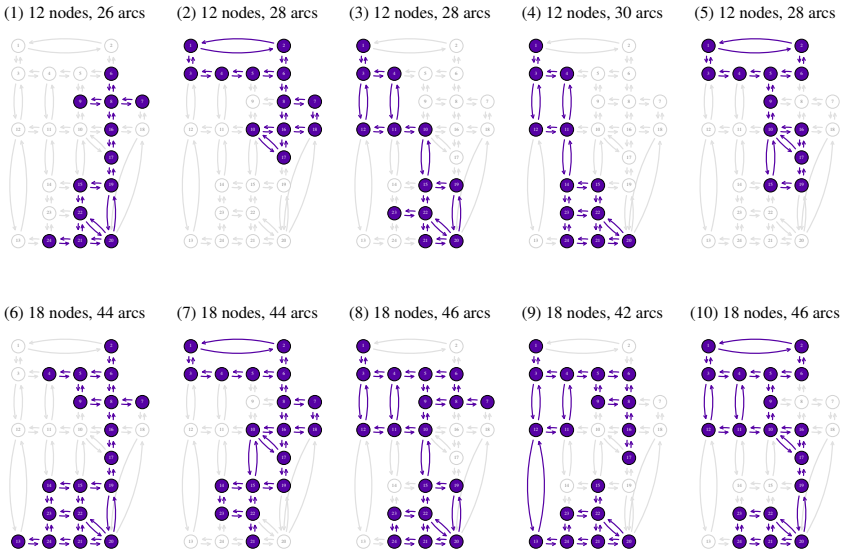


FIGURE 5. The subnetworks of the Sioux Falls network considered in the computational study. Nodes and arcs included in the subnetworks are shown in purple.

their underlying undirected graphs are not biconnected. This structure allows us to assess the effectiveness of the enhancement techniques introduced in Section 4.

6.2.2. *SNDlib Instances.* As the SNDlib instances have originally been created for telecommunication network design, they do not include all parameters required for traffic assignment problems, such as arc capacities and free-flow travel times. Therefore, we proceed as follows. From each SNDlib instance, we take the set of nodes V with their geographical coordinates, the set of links, and the traffic demands. We then parse the graph topology from the SNDlib native format, duplicating each undirected link to create a directed bidirectional graph. Arc capacities are set proportionally to the total demand $D = \sum_{k \in K} d_k$ across all commodities and we introduce heterogeneity through a random perturbation. More specifically, we consider $u_a = 0.1 D \xi_a$ for all $a \in A$, where $\xi_a \sim U[0.1, 2.0]$ is a uniformly distributed random value that introduces variability across arcs. Finally, to set the free-flow travel time c_a^{fix} for each arc $a = (i, j)$, we consider the Euclidean distance d_{ij} between its endpoints, which is computed using the geographical coordinates provided in the SNDlib data, and normalize it using the maximum distance $d_{\max}$ in the network. The free-flow time is then obtained as

$$c_a^{\text{fix}} = c_{\min} + \frac{d_{ij}}{d_{\max}} \left(c_{\max} - c_{\min} \right), \quad a \in A,$$

with $c_{\min} = 1$ and $c_{\max} = 50$. Hence, longer arcs have higher free-flow costs, whereas shorter arcs have lower costs. As with the Sioux Falls instances, we vary the number of commodities for each network. We consider five cases by selecting 10 %, 25 %, 50 %, 75 %, and 100 % of the number of commodities $|K|$ of the original instance. Commodities are randomly selected from the set of all OD pairs and the demand parameters are determined following the same procedure as for the Sioux Falls instances. We summarize the main characteristics of the considered SNDlib instances in Table 2 and illustrate their network topologies in Figure 6.

TABLE 2. The number of nodes, arcs, and commodities of the considered SNDlib instances.

Instance	$\|V\|$	$\|A\|$	$\lceil 10\,\%\|K\|\rceil$	$\lceil 25\,\%\|K\|\rceil$	$\lceil 50\,\%\|K\|\rceil$	$\lceil 75\,\%\|K\|\rceil$	$\|K\|$
abilene	12	30	13	33	66	99	132
atlanta	15	44	21	52	105	158	210
nobel-us	28	42	9	22	45	68	91
pdh	11	68	2	6	12	18	24

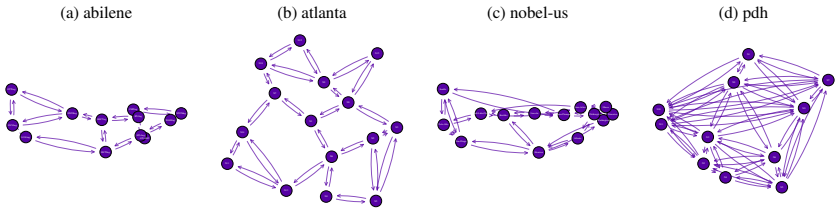


FIGURE 6. The network topologies of the considered SNDlib instances.

7. Discussion of the Computational Results

We now present and discuss the results of our computational study. Unless stated otherwise, the results refer to the Sioux Falls subnetworks under budgeted uncertainty using the user equilibrium formulation (9) and the enhancement techniques described in Section 4.

In Section 7.1, we discuss the computational performance and scalability of our models. In Section 7.2, we analyze the sensitivity of optimal solutions to different choices of travel cost and latency functions. In Section 7.3, we compare the UE and SO formulations in terms of congestion levels and computational performance, whereas the impact of different uncertainty models is discussed in Section 7.4. Finally, in Section 7.5, we present the results for the SNDlib instances.

For the ease of presentation, we abbreviate the maximum utilization and total utilization latency functions as max_ratio and sum_ratio, respectively. Moreover, we refer to the travel cost functions with $\beta = 1$, $\beta = 2$, and $\beta = 4$ as linear, quadratic, and BPR, respectively.

7.1. Performance analysis.

We first evaluate the effectiveness of the enhancement techniques from Section 4, along with the scalability and tractability of the congestion models across different combinations of cost and congestion functions.

7.1.1. *Impact of Enhancement Techniques.*

In what follows, we use *tightened formulation* to denote the formulation obtained after applying the enhancement techniques described in Section 4, which reduce the model size and strengthen the formulation by exploiting the network topology and structural properties of optimal flows. We use the term *standard formulation* to denote the original formulation without any enhancements.

Table 3 reports the impact of the enhancement techniques on the computation time and the number of solved instances for $|K| \in \{20, 30, 50, 75, 100\}$. Here, we aggregate the results over the three considered latency functions and report them by travel cost function, as the travel cost function has the largest impact on computation time. The latter is discussed in more detail in Section 7.1.2. To compare the tightened and standard formulations, we further report the speed-up ratio in Table 3, defined as the execution time of the tightened formulation divided by that

TABLE 3. The number of instances solved to optimality ("#opt") and the average execution time ("time", in s) for the tightened and standard formulations of the congestion model for different values of $|K|$ and travel cost functions. Additionally, the speed-up ratio ("speed-up"), the time saved (in %), and the difference in the number of solved instances ("Δ#opt") are shown.

		Tightened		**Standard**		**Comparison**		
$\|K\|$	$c(f)$	#opt	time	#opt	time	speed-up	time saved	Δ#opt
20		**90**	**0.08**	**90**	**1.77**	**21.34**	**95**	**0**
	linear	30	0.03	30	1.12	32.56	97	0
	quadratic	30	0.06	30	1.95	34.37	97	0
	BPR	30	0.16	30	2.25	14.22	93	0
30		**90**	**0.29**	**90**	**9.97**	**34.53**	**97**	**0**
	linear	30	0.07	30	2.95	40.55	98	0
	quadratic	30	0.17	30	6.68	38.85	97	0
	BPR	30	0.62	30	20.30	32.63	97	0
50		**90**	**2.14**	**89**	**59.06**	**27.61**	**96**	**1**
	linear	30	0.23	30	4.68	20.56	95	0
	quadratic	30	1.04	30	17.89	17.16	94	0
	BPR	30	5.25	29	157.90	30.08	97	1
75		**90**	**16.56**	**77**	**125.97**	**7.61**	**87**	**13**
	linear	30	0.96	30	12.61	13.16	92	0
	quadratic	30	5.14	28	69.60	13.55	93	2
	BPR	30	58.02	19	388.03	6.69	85	11
100		**81**	**27.44**	**64**	**438.14**	**15.96**	**94**	**17**
	linear	30	4.86	30	407.82	83.84	99	0
	quadratic	30	26.43	26	345.65	13.08	92	4
	BPR	21	127.98	8	911.62	7.12	86	13
Total		**441**	**7.89**	**410**	**106.64**	**13.51**	**93**	**31**

of the standard formulation. For a fair comparison, we only present runtime results for instances solved to optimality under both formulations.

The results show that, for $|K| \in \{20, 30\}$, all instances are solved to optimality within the time limit of 2 h under both formulations, with the tightened formulation reducing execution time by more than 93 %. As $|K|$ increases, the enhancement techniques also affect solvability. When incorporating our enhancement techniques, the number of additional instances solved to optimality increases from 1 at $|K| = 50$ to 13 at $|K| = 75$ to 17 at $|K| = 100$. In summary, the tightened formulation leads to faster execution times and more instances solved to optimality, especially when considering many commodities and the BPR travel cost function.

7.1.2. *Scalability Analysis.* We now assess the scalability of the congestion models across different combinations of travel cost functions and congestion measures. Figure 7 illustrates the results by comparing execution times with the size of the congestion model (Figure (a)) and with the number of commodities (Figure (b)). Model size refers to the number of variables multiplied by the number of constraints.

Both plots in Figure 7 show a considerable dispersion in execution times and, for comparable model sizes and identical values of $|K|$, execution times vary by several orders of magnitude. This indicates that the problem size alone does not fully explain the computational behavior and that structural properties of the formulation

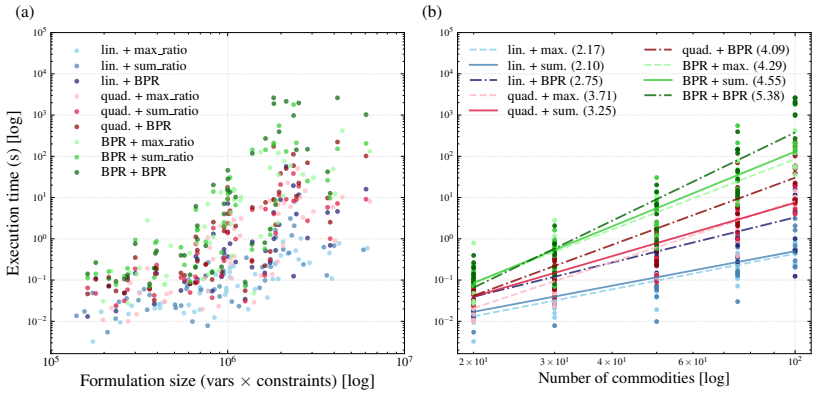


FIGURE 7. Scalability analysis for different combinations of travel cost and latency functions. Figure (a): Execution time versus problem size. Figure (b): Execution time versus number of commodities $|K|$ with fitted power-law models $t = a \cdot |K|^{\alpha}$ and α being the growth exponent reported in brackets in the legend.

play a key role in scalability. In this context, we observe that the dominant factor affecting scalability is the choice of the travel cost function. Instances with BPR costs exhibit the highest execution times and the steepest growth, whereas linear costs lead to the lowest execution times across all combinations. Quadratic costs reveal an intermediate behavior. These trends are visible in Figure 7 (a) through the vertical separation of color groups and in Figure 7 (b) via larger scaling exponents for BPR costs.

In contrast, the impact of the latency function is less pronounced. For a given cost function, differences among sum_ratio, max_ratio, and BPR congestion models result in smaller variations in execution time, as illustrated by the overlap of points of the same color but different shades in Figure 7 (a) and the proximity of curves with the same color but different line styles in Figure 7 (b). Nevertheless, for all three cost functions, we observe that the BPR congestion measure yields the highest execution times, followed by max_ratio, whereas sum_ratio provides the fastest solution times. Overall, these results indicates that the computational scalability is primarily determined by the travel cost function, whereas the latency function plays a secondary, yet still noticeable, role in execution times.

7.1.3. *Performance Across Function Combinations.* As observed in Section 7.1.2, the choice of the travel cost function $c(f)$ has a stronger impact on the solution time of the congestion models than the choice of latency function. To quantify these differences, we now aggregate performance metrics across all instances solved to optimality for each combination of travel cost and latency function. In Table 4, we summarize the average execution times and success rates for each combination. Here and in what follows, success rate refers to the percentage of instances solved to optimality out of all considered instances. The results confirm that the travel cost function is the primary factor affecting computational performance. Although latency functions also influence performance, there is no consistent trend that holds across all configurations. For example, with linear or quadratic costs, all instances are solved to optimality and execution times for max_ratio are comparable to those of sum_ratio and faster than BPR. However, for BPR travel costs, max_ratio leads to more challenging optimization problems, increasing the fraction of instances that cannot be solved to optimality within the time limit.

TABLE 4. The average execution time (in s) over all instances solved to optimality and the success rate (in %).

	Execution time			Success rate		
$c(f)$	max_ratio	sum_ratio	BPR	max_ratio	sum_ratio	BPR
linear	0.33	0.32	3.05	100.0	100.0	100.0
quadratic	3.72	3.65	19.23	100.0	100.0	100.0
BPR	27.52	83.42	303.24	92.0	94.0	96.0

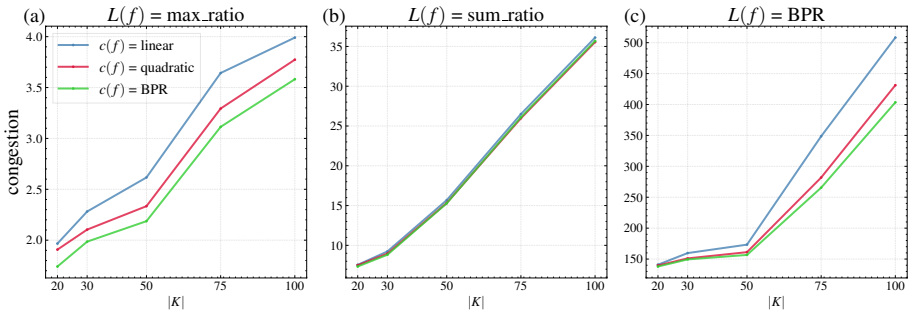


FIGURE 8. The congestion estimates for the models using the max_ratio (a), sum_ratio (b), and BPR (c) congestion measures. Each panel shows how the average congestion level varies with the number of commodities for different cost functions.

7.2. **Sensitivity of Network Congestion to Travel Cost and Latency Functions.** We now examine the impact of travel cost and latency functions on congestion estimates. To ensure a meaningful comparison, we focus on those instances that were solved to optimality for all nine combinations of travel cost and latency functions for each value of $|K|$.

7.2.1. *Impact of the Travel Cost Function on Congestion Estimates.* In Figure 8, we show the evolution of congestion levels as a function of the number of commodities $|K|$ for each combination of travel cost and latency function. For a given latency function $L(f)$, the choice of cost function $c(f)$ noticeably affects the estimated congestion levels. In particular, the BPR cost function consistently yields lower congestion estimates than linear or quadratic costs. Figure 8 (a) illustrates this effect for the max_ratio latency function. Here, BPR costs produce congestion estimates that are about 15 %–20 % lower than those obtained with linear costs. This indicates that nonlinear travel cost functions, which penalize congestion more strongly, are more effective at limiting extreme arc saturation. Figure 8 (c) shows the same qualitative behavior for the BPR latency function. Again, BPR costs achieve the lowest congestion values and the effect is particularly pronounced for instances with a large number of commodities $|K|$.

In contrast, Figure 8 (b) shows that the sum_ratio metric is less sensitive to the choice of travel cost function as all three of them yield similar congestion estimates across the range of $|K|$. This behavior is explained by the aggregate nature of the metric. Different cost functions lead to different flow distributions at the arc level. However, local variations tend to compensate when summed over all arcs of the network as higher utilization of some arcs is balanced by lower utilization of others. The small size of the network further amplifies this effect as the limited number

TABLE 5. The relative gap (in %) on L_{eval} when optimizing for L_{opt} by the number of commodities $|K|$. Values are averaged over all considered instances and cost functions.

		$\|K\|$				
L_{eval}	L_{opt}	20	30	50	75	100
max_ratio	sum_ratio	3.1	2.3	3.8	5.0	7.8
	BPR	1.6	0.5	0.7	1.1	2.9
sum_ratio	max_ratio	2.3	3.3	3.2	3.1	6.0
	BPR	1.3	1.3	1.2	1.6	1.4
BPR	max_ratio	0.1	0.3	0.9	2.7	10.2
	sum_ratio	0.2	0.7	2.1	5.6	7.8

of alternative paths constrains the degree to which flows can be redistributed in response to changes in the cost function.

7.2.2. *Impact of the Latency Function on Congestion Estimates.* To assess how the choice of the latency function in the objective of the congestion model affects congestion estimates, we now analyze how solutions obtained by optimizing one latency function perform under another. Specifically, for a given latency function L_{opt}, i.e., either max_ratio, sum_ratio, or BPR, we compute an optimal flow $f^*_{L_{\text{opt}}}$ by solving the congestion model (9) and then evaluate the resulting congestion level using a different latency function L_{eval}. For this purpose, we define the relative gap

$$\text{gap}(L_{\text{opt}}, L_{\text{eval}}) = \frac{L_{\text{eval}}(f^*_{L_{\text{eval}}}) - L_{\text{eval}}(f^*_{L_{\text{opt}}})}{L_{\text{eval}}(f^*_{L_{\text{eval}}})} \tag{26}$$

for each pair $(L_{\text{opt}}, L_{\text{eval}})$. Here, $f^*_{L_{\text{opt}}}$ and $f^*_{L_{\text{eval}}}$ denote the flows that optimize the congestion measures L_{opt} and L_{eval}, respectively. By construction, we have $\text{gap}(L_{\text{opt}}, L_{\text{eval}}) \geq 0$. Gaps close to zero indicate that optimizing for L_{opt} yields solutions that are near-optimal when evaluated under L_{eval}, whereas larger values indicate a greater discrepancy between the two formulations. We note that this analysis does not account for the multiplicity of equilibrium flows. Hence, other equilibria optimal for L_{opt} may perform better under L_{eval}. Here, we restrict ourselves to evaluating the solution returned by Gurobi.

Table 5 reports the gaps for all distinct pairs of $(L_{\text{eval}}, L_{\text{opt}})$ and for each value of $|K|$. For instances with $|K| \in \{20, 30\}$, the discrepancies between latency functions are limited, with gap values remaining within 3.3 % for all pairs. However, as $|K|$ increases, larger differences emerge. For instance, optimizing with respect to sum_ratio instead of max_ratio results in a gap of approximately 8 % for $|K| = 100$. Similarly, optimizing for max_ratio or sum_ratio instead of BPR leads to gaps of up to 10 %. Conversely, flows obtained by optimizing for BPR exhibit smaller discrepancies when evaluated under other latency functions, with gaps below 3 % in most cases. These results indicate that BPR-optimized flows lead to maximum utilization and total utilization levels that are close to their respective worst case. Overall, the magnitude of the discrepancy between latency functions increases with the number of commodities.

7.3. **User equilibrium vs. system optimum.** We now compare the UE and SO formulations of the congestion model in terms of both computational performance and obtained congestion levels. As before, we focus on instances that were solved to optimality by both formulations to ensure a fair comparison.

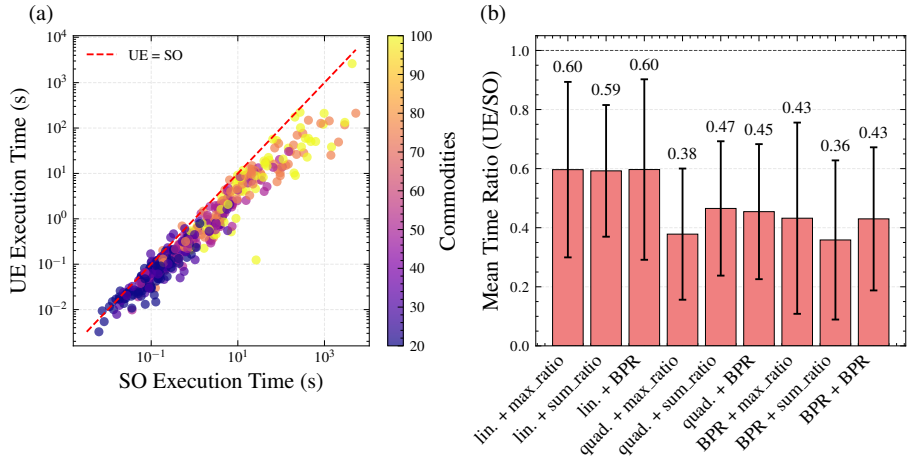


FIGURE 9. Computation time comparison between UE and SO formulations. Figure (a): Scatter plot of execution times (in s, logarithmic scale). Each point represents a single instance, with colors indicating the number of commodities. The red dashed line marks instances for which solving the UE and SO formulation requires the same amount of time. Figure (b): Mean execution time ratio (red box) with standard deviation (error bars) for each combination of travel cost and latency function. Mean values are shown above each error bar.

7.3.1. *Runtime Comparison.* In Figure 9, we compare the computation times of the UE and SO formulations. The scatter plot in Figure 9 (a) shows that solving the UE formulation consistently requires shorter execution times than solving the SO formulation. This advantage is particularly pronounced for instances with a large number of commodities. Figure 9 (b) presents the average execution time ratios for each combination of travel cost and latency function. Ratios are computed as the execution time of UE divided by that of SO. The mean ratio ranges from approximately 0.4 to 0.6, indicating that solving the UE formulation requires only 40 %–60 % of the time needed for solving the SO formulation. Linear cost functions exhibit the highest ratios (around 0.6), whereas quadratic and BPR cost functions yield lower ratios. This suggests that the computational advantage of solving the UE formulation increases for more involved travel cost functions. In contrast, no consistent pattern is observed with respect to the choice of the latency function.

7.3.2. *Comparison of Congestion Levels.* We now compare the congestion levels obtained by the UE and SO formulations by analyzing the distribution of the congestion ratio (CR, see Section 2.3) across all instances and function combinations. Figure 10 (a) shows a histogram of CR values. It can be seen that most instances achieve CR values close to 1, indicating that UE solutions typically yield congestion levels comparable to those of SO solutions. However, notable outliers exist, showing that UE solutions can lead to significantly higher congestion than SO solutions in some cases. At the same time, some instances have CR values below 1, which is consistent with the behavior observed in Example 1. Figure 10 (b) illustrates the distribution of CR values across different combinations of travel cost and latency functions. As in Figure 10 (a), the distributions are generally concentrated around 1. Combinations including the max_ratio congestion measure tend to exhibit higher congestion ratios, indicating that UE solutions under this measure can

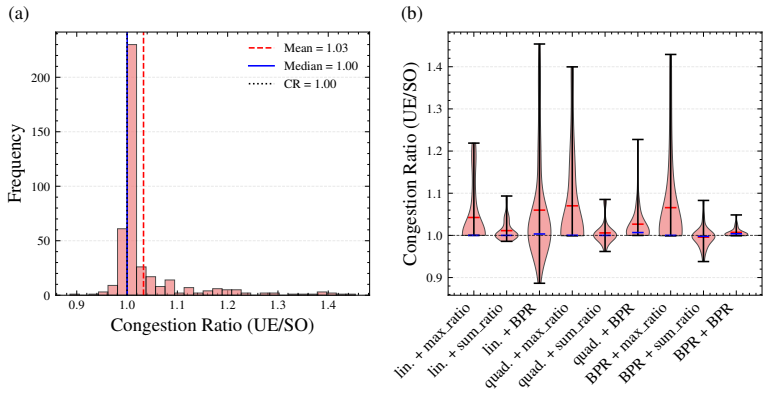


FIGURE 10. The distribution of the congestion ratio (CR) across all instances and function combinations. Figure (a): Histogram of CR values, with mean and median indicated by dashed red and solid blue lines, respectively. Figure (b): Violin plots for the CR distribution for each combination of travel cost and latency functions, with mean (red) and median (blue) markers.

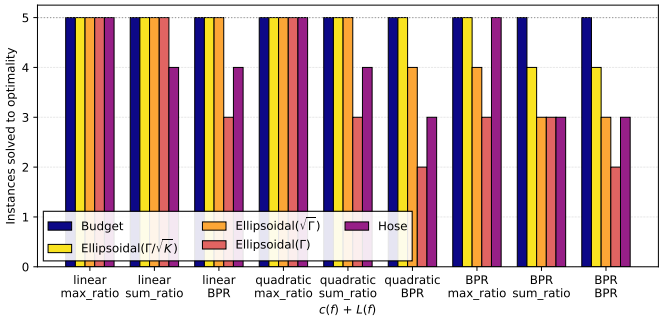


FIGURE 11. The number of instances solved to optimality by uncertainty set and function combination.

produce more congestion relative to SO. In contrast, the sum_ratio measure yields the smallest CR values, suggesting that UE solutions under this measure can sometimes lead to more favorable outcomes. Finally, combinations with BPR costs show a wider spread of CR values, reflecting greater variability in congestion outcomes between the UE and SO formulations.

7.4. **Uncertainty Set Comparison.** We now compare the five uncertainty sets introduced in Section 5.3: the budgeted (Γ-robust) uncertainty set $\mathcal{U}_{\text{budget}}$, the ellipsoidal uncertainty set $\mathcal{U}_{\text{ellipsoid}}$ with $\rho \in \{\sqrt{\Gamma}, \Gamma, \Gamma/\sqrt{|K|}\}$, and the symmetric hose uncertainty set $\mathcal{U}_{\text{hose}}$. For the ease of presentation, we restrict the analysis to instances derived from the Sioux Falls network with 12 nodes and 26 arcs (cf. Figure 5) for $|K| \in \{20, 30, 50, 75, 100\}$ and all nine combinations of travel cost and latency functions. This results in 45 considered instances per uncertainty set.

7.4.1. *Computational Tractability.* Figure 11 illustrates the number of instances solved to optimality within the time limit of 2 h for each uncertainty set and combination of travel cost and latency functions. The results indicate a clear dominance in terms of computational tractability of the resulting congestion models: the budget uncertainty model is the most computationally tractable, whereas the ellipsoidal

TABLE 6. The average computation times (in s) by uncertainty set and function combination. Additionally, the number of instances solved to optimality ("#opt") is shown.

				$\mathcal{U}_{\text{ellipsoid}}$			
$c(f)$	$L(f)$	#opt	$\mathcal{U}_{\text{budget}}$	$\rho = \frac{\Gamma}{\sqrt{\lvert K \rvert}}$	$\rho = \sqrt{\Gamma}$	$\rho = \Gamma$	$\mathcal{U}_{\text{hose}}$
linear		**12**	**0.05**	**0.60**	**1.55**	**150.72**	**1.14**
	sum_ratio	4	0.01	0.11	1.29	3.90	1.62
	max_ratio	5	0.07	0.83	1.97	24.34	0.87
	BPR	3	0.06	0.87	1.18	557.09	0.94
quadratic		**10**	**1.20**	**9.93**	**33.30**	**446.52**	**5.54**
	sum_ratio	3	0.05	1.75	1.99	4.36	1.16
	max_ratio	5	2.34	18.65	65.19	884.85	9.99
	BPR	2	0.08	0.38	0.53	13.95	0.52
BPR		**8**	**0.05**	**1.76**	**1.68**	**42.07**	**21.34**
	sum_ratio	3	0.65	3.17	2.85	98.01	55.90
	max_ratio	3	0.72	1.12	1.11	12.92	0.47
	BPR	2	0.12	0.59	0.78	1.86	0.81

model with $\rho = \Gamma$ seems to be the most challenging formulation. In particular, the number of solved instances per formulation reflects the computational challenges associated with both the geometric properties of the uncertainty sets (e.g., their size) and the nature of the resulting optimization problems (MILP vs. MIQP).

Table 6 further summarizes the average execution times for each uncertainty model over the 30 instances (out of 45) that were solved to optimality under all five uncertainty sets. We observe that the budget uncertainty set consistently yields the shortest solution times across all function combinations. In contrast, the ellipsoidal model with $\rho = \Gamma$ requires the longest computation times, often exceeding the other formulations by two to three orders of magnitude.

7.4.2. *Impact on Worst-Case Congestion.* We now assess the impact of the uncertainty model on worst-case congestion by comparing the congestion levels of the models with demand uncertainty to those of the deterministic model. In Figure 12, we show heatmaps of the percentage increase in congestion relative to the deterministic case for each uncertainty model, disaggregated by travel cost and latency function. We observe that all uncertainty models lead to higher worst-case congestion compared to the deterministic setting. The lowest worst-case congestion is observed for the ellipsoidal uncertainty set with $\rho = \Gamma/\sqrt{|K|}$, followed by the budgeted uncertainty set, the ellipsoidal sets with $\rho = \sqrt{\Gamma}$ and $\rho = \Gamma$, and finally the hose model. This ordering aligns with the set inclusions discussed in Section 5.3. Distinct patterns also emerge with respect to the travel cost and latency functions. Across all uncertainty sets, the max_ratio congestion measure produces the largest increases in worst-case congestion, whereas sum_ratio yields the smallest. The hose model and the ellipsoidal uncertainty set with $\rho = \Gamma$ consistently result in higher worst-case congestion. We further observe that, for the budgeted model and ellipsoidal uncertainty sets with $\rho \in \{\Gamma/\sqrt{|K|}, \sqrt{\Gamma}\}$, the choice of travel cost function has a limited effect on congestion levels. In contrast, congestion levels are highly sensitive to the travel cost function for the ellipsoidal set with $\rho = \Gamma$ and the hose model. In particular, the hose model produces increases exceeding 50 % in most cases, with peaks above 90 % for linear and quadratic cost functions.

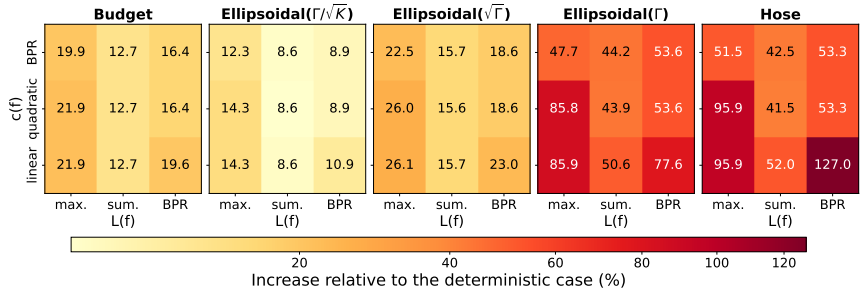


FIGURE 12. The percentage increase in worst-case congestion relative to the deterministic case, by uncertainty set and function combination. All panels share a common color scale for comparability. Darker colors indicate larger deviations from the deterministic solution.

TABLE 7. The number of SNDlib instances solved to optimality ("#opt", out of 9) and the average execution time (in s) for instances with 10 %, 25 %, 50 %, 75 %, and 100 % of the number of commodities from the original network.

	10 %		25 %		50 %		75 %		100 %	
	#opt	time	#opt	time	#opt	time	#opt	time	#opt	time
abilene	9	2.2	9	7.0	9	11.9	6	24.9	6	28.0
atlanta	7	73.5	9	637.4	5	1993.4	3	29.9	4	1320.1
nobel-us	9	1.5	9	67.1	4	1094.1	3	88.3	3	67.7
pdh	9	0.6	9	7.2	6	48.5	3	17.0	3	43.8

7.5. **Results for SNDlib Instances.** We now discuss the computational results for the SNDlib instances. These instances exhibit different network topologies and demand structures compared to those of the Sioux Falls subnetworks, allowing us to evaluate the congestion models from a broader perspective.

Table 7 reports the number of solved instances and the average execution times for each network across different numbers of commodities. For instances with 10 % and 25 % of the commodities from the original network, almost all instances are solved to optimality. However, as the number of commodities increases, solving the congestion models becomes increasingly challenging. Among the considered networks, atlanta is particularly challenging. Only 5 out of 9 instances can be solved to optimality for 50 % of the commodities from the original network and we observe that execution times are high even for smaller instances. The same qualitative behavior can be observed for the nobel-us and pdh networks: as the number of commodities increases, the number of instances solved to optimality decreases. Overall, abilene is the most computationally tractable network, with all instances solved to optimality for up to 50 % of the original commodities, and 6 out of 9 instances solved for larger numbers of commodities.

In Table 8, we additionally report results disaggregated by travel cost and latency function. For each function combination, we consider a total of 20 SNDlib instances, aggregated across all networks. Overall, 125 out of 180 instances (69.4 %) are solved to optimality. Moreover, the results for the SNDlib instances support the previous observations made for the Sioux Falls subnetworks. All 20 instances with linear travel costs are solved to optimality, whereas the success rate decreases to 65 % for

TABLE 8. The number of instances solved to optimality ("#opt", out of 20 instances) and the average execution time ("time", in s) for different combinations of travel cost and latency functions.

$c(f)$	$L(f)$	#opt	time
linear	sum_ratio	20	39.3
	max_ratio	20	24.8
	BPR	20	22.2
quadratic	sum_ratio	14	916.9
	max_ratio	12	29.0
	BPR	13	538.7
BPR	sum_ratio	8	70.0
	max_ratio	9	577.7
	BPR	9	47.5

quadratic and 43.3 % for BPR costs. Among the latency functions, sum_ratio tends to require longer execution times when combined with quadratic costs (916.9 s on average), whereas max_ratio remains faster across all cost functions.

Overall, the SNDlib instances are more challenging to solve than the Sioux Falls subnetworks, especially for nonlinear travel cost functions. This is likely due to the more complex topologies of the SNDlib networks, as the absence of articulation nodes limits the effectiveness of certain enhancement techniques described in Section 4.

7.6. **Managerial Insights.** To conclude this computational study, we highlight insights relevant for practitioners and network operators.

When selecting travel cost functions, congestion measures, routing policies, and uncertainty sets for traffic networks under demand uncertainty, practitioners must balance model realism, computational tractability, and solution quality. Our study shows that the choice of travel cost function has a stronger impact on computational tractability than the choice of congestion measure. Linear cost functions yield the fastest solution times and highest success rates. BPR cost functions, although more realistic in capturing congestion effects, require longer solution times and may fail to reach optimality for large instances. For operational settings that require frequent re-optimization, linear or quadratic cost functions provide a practical compromise between realism and computational tractability. Nevertheless, flows obtained using BPR costs produce congestion levels that are close to the worst case using linear or quadratic cost functions across all latency measures, suggesting that BPR costs may serve as a prudent choice if the appropriate measure is not known exactly.

Among the congestion measures, sum_ratio seems to provide the best trade-off between computational tractability and solution quality. The max_ratio measure, although effective at limiting extreme arc utilization, leads to more challenging optimization problems if combined with nonlinear travel cost functions.

Comparing UE and SO formulations, we find that solving the UE formulation requires only 40 %–60 % of the time needed for the SO formulation. Hence, the UE formulation is preferable for repeated re-optimization and when a more realistic representation of travelers' behavior is desired.

Our study further underscores the importance of accounting for demand uncertainty. Even the smallest uncertainty set considered in our evaluations increases worst-case congestion by at least 4.7 % compared to the deterministic case. On average, solutions under uncertainty exhibit 11 %–72 % higher congestion than their

deterministic counterparts, depending on the chosen uncertainty set. Ignoring uncertainty can thus lead to substantial underestimation of network congestion. Regarding the choice of the uncertainty set, we observe a trade-off between worst-case estimates and computational tractability. The budgeted uncertainty set offers the most favorable balance, providing protection against demand fluctuations with moderate increases in worst-case congestion, while achieving high success rates and the fastest computation times. The ellipsoidal uncertainty set with $\rho = \sqrt{\Gamma}$ yields similar worst-case congestion levels as the budgeted model but is computationally slightly more expensive. Thus, the budgeted model should be preferred. Overall, the hose model and the ellipsoidal uncertainty set with $\rho = \Gamma$ produce the highest worst-case congestion levels. Hence, these models are suitable if strong protection against fluctuating demand is critical, but practitioners should be aware of the reduced computational tractability of these formulations.

To sum up, for typical operational planning, the combination of a budgeted uncertainty model with the sum_ratio congestion measure and linear or quadratic cost functions appears preferable, as it provides resilient solutions within reasonable computation times for networks with up to 100 commodities. If peak congestion is the main concern, the max_ratio measure should be adopted, keeping in mind that this may lead to more challenging problems to solve. For long-term planning, where computation time is less critical, BPR cost functions offer a more realistic representation of congestion.

8. Conclusion

We study the problem of determining the worst-case congestion in a multi-commodity traffic network under demand uncertainty. Our aim is to stress-test a given network by identifying demand realizations and corresponding travelers' route choices that maximize congestion. To this end, we propose a novel bilevel formulation in which a traffic planner acts as the leader and the users of the traffic network act as the followers. In the leader's problem, we account for demand uncertainty using ideas from robust optimization, whereas the follower's problem models a traffic equilibrium in which travelers follow one of the two Wardrop principles—the user equilibrium (UE) or the system optimum (SO). We develop single-level mixed-integer nonlinear reformulations for the resulting congestion models that exploit binary variables and big-M constants, prove the existence of optimal solutions, derive valid big-Ms, and propose several enhancement techniques to strengthen the formulations. Moreover, we perform an extensive computational study on instances of the Sioux Falls network and the SNDlib, which provides several key insights.

First, we observe that the proposed enhancement techniques significantly improve computational performance, reducing solution times by over 90 % and enabling the solution of larger network instances that would otherwise remain unsolved within the time limit of 2 h. Second, the choice of travel cost and latency functions has a substantial impact on both the computational tractability of the congestion models and the estimated worst-case congestion. In particular, we observe that the choice of the cost function has a stronger impact on the tractability of the model than the congestion measure. Third, solving the UE formulation can be done considerably faster than solving the SO formulation, whereas congestion levels under the UE are typically very close to those under the SO. Indeed, by studying the so-called congestion ratio, we show that the two equilibrium concepts yield similar worst-case congestion levels. This indicates that, under demand uncertainty, centralizing routing decisions as in the SO does not provide a clear advantage over the UE in terms of reducing congestion. In particular, although the SO aims at minimizing the total travel time in the system, it does not always lead to less

congestion compared to the UE in our framework. Fourth and finally, our analysis of budgeted, ellipsoidal, and hose uncertainty sets reveals a trade-off between worst-case congestion and computational cost. The budgeted model can be solved the fastest, whereas the ellipsoidal and hose models require more computation time but generally yield higher congestion estimates. Based on these findings, we derive managerial guidelines for choosing the most appropriate modeling of congestion, travel costs, and uncertainty.

This work provides a first step towards designing traffic networks that are more resilient to congestion under demand uncertainty. Building on our stress-testing framework, a natural next step is to embed it into a network design model, resulting in a robust bilevel optimization problem in which infrastructure decisions are made while anticipating traffic equilibrium responses under demand uncertainty.

Acknowledgment

This research was funded in part by the French National Research Agency (ANR) under project ANR-24-CE48-3565-03. Sara Mattia is a member of GNAMPA-INdAM. The authors contributed equally to this work (are co-lead authors) and are listed in alphabetical order.

References

Ahuja, R. K., T. L. Magnanti, and J. B. Orlin (1993). *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall.

Altın, A., E. Amaldi, P. Belotti, and M. Pınar (2007). "Provisioning virtual private networks under traffic uncertainty." In: *Networks* 49.1, pp. 100–115. DOI: 10.1002/net.20145.

Bard, J. F. (1998). *Practical Bilevel Optimization: Algorithms and Applications*. Nonconvex Optimization and Its Applications. Springer New York, NY. DOI: 10.1007/978-1-4757-2836-1.

Beck, Y., M. Labbé, and M. Schmidt (2024). *Toll Setting with Robust Wardrop Equilibrium Conditions Under Budgeted Uncertainty*. Tech. rep. URL: https://optimization-online.org/?p=26949.

Ben-Ayed, O., D. E. Boyce, and C. E. Blair (1988). "A general bilevel linear programming formulation of the network design problem." In: *Transportation Research Part B: Methodological* 22.4, pp. 311–318. DOI: 10.1016/0191-2615(88)90006-9.

Ben-Tal, A., L. Ghaoui, and A. Nemirovski (2009). *Robust Optimization*. DOI: 10.1515/9781400831050.

Bertsimas, D., D. B. Brown, and C. Caramanis (2011). "Theory and Applications of Robust Optimization." In: *SIAM Review* 53.3, pp. 464–501. DOI: 10.1137/080734510.

Bertsimas, D. and M. Sim (2003). "Robust discrete optimization and network flows." In: *Mathematical Programming* 98, pp. 49–71. DOI: 10.1007/s10107-003-0396-4.

– (2004). "The Price of Robustness." In: *Operations Research* 52.1, pp. 35–53. DOI: 10.1287/opre.1030.0065.

Bienstock, D., S. Chopra, O. Günlük, and C.-Y. Tsai (1998). "Minimum cost capacity installation for multicommodity network flows." In: *Mathematical Programming* 81.2, pp. 177–199. DOI: 10.1007/BF01581104.

Boyles, S. D., N. E. Lownes, and A. Unnikrishnan (2025). *Transportation Network Analysis, Volume I: Static and Dynamic Traffic Assignment*. URL: https://arxiv.org/abs/2502.05182.

Brotcorne, L., M. Labbé, P. Marcotte, and G. Savard (2001). "A Bilevel Model for Toll Optimization on a Multicommodity Transportation Network." In: *Transportation Science* 35.4. DOI: 10.1287/trsc.35.4.345.10433.

Cerulli, M., C. Archetti, E. Fernández, and I. Ljubić (2024). "A Bilevel Approach for Compensation and Routing Decisions in Last-Mile Delivery." In: *Transportation Science* 58.5, pp. 1076–1100. DOI: 10.1287/TRSC.2023.0129.

Chekuri, C., G. Oriolo, M. Scutellà, and B. Shepherd (2007). "Hardness of Robust Network Design." In: *Networks* 50.1, pp. 50–154. DOI: 10.1002/net.20165.

Chouman, M., T. G. Crainic, and B. Gendron (2017). "Commodity Representations and Cut-Set-Based Inequalities for Multicommodity Capacitated Fixed-Charge Network Design." In: *Transportation Science* 51.2, pp. 650–667. DOI: 10.1287/trsc.2015.0665.

Colson, B., P. Marcotte, and G. Savard (2005). "Bilevel programming: A survey." In: *4OR*, pp. 87–107. DOI: 10.1007/s10288-005-0071-0.

– (2007). "An overview of bilevel optimization." In: *Annals of Operations Research* 153, pp. 235–256. DOI: 10.1007/s10479-007-0176-2.

Cominetti, R., V. Dose, and M. Scarsini (2024). "Monotonicity of equilibria in nonatomic congestion games." In: *European Journal of Operational Research* 316.2, pp. 754–766. DOI: 10.1016/j.ejor.2024.01.050.

Correa, J. R. and N. E. Stier-Moses (2011). "Wardrop Equilibria." In: *Wiley Encyclopedia of Operations Research and Management Science*. Wiley. DOI: 10.1002/9780470400531.eorms0962.

Dempe, S. (2002). *Foundations of Bilevel Programming*. Springer US. DOI: 10.1007/b101970.

Dempe, S., V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnykova (2015). *Bilevel Programming Problems: Theory, Algorithms and Applications to Energy Networks*. Energy Systems. Springer Berlin, Heidelberg. DOI: 10.1007/978-3-662-45827-3.

Dempe, S. and A. Zemkoho (2012). "Bilevel road pricing: theoretical analysis and optimality conditions." In: *Annals of Operations Research* 196, pp. 223–240. DOI: 10.1007/s10479-011-1023-z.

Dewez, S., M. Labbé, P. Marcotte, and G. Savard (2008). "New formulations and valid inequalities for a bilevel pricing problem." In: *Operations Research Letters* 36.2, pp. 141–149. DOI: 10.1016/j.orl.2007.03.005.

Duffield, N. G., P. Goyal, A. Greenberg, P. Mishra, K. K. Ramakrishnan, and J. E. van der Merive (1999). "A flexible model for resource management in virtual private networks." In: *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*. SIGCOMM '99. Cambridge, Massachusetts, USA: Association for Computing Machinery, 95–108. DOI: 10.1145/316188.316209.

Ferris, M. C. and J. S. Pang (1997). "Engineering and Economic Applications of Complementarity Problems." In: *SIAM Review* 39.4, pp. 669–713. DOI: 10.1137/S0036144595285963.

Fingerhut, J., S. Suri, and J. S. Turner (1997). "Designing Least-Cost Nonblocking Broadband Networks." In: *Journal of Algorithms* 24.2, pp. 287–309. DOI: 10.1006/jagm.1997.0866.

Fontaine, P. and S. Minner (2014). "Benders Decomposition for Discrete–Continuous Linear Bilevel Problems with application to traffic network design." In: *Transportation Research Part B: Methodological* 70, pp. 163–172. DOI: 10.1016/j.trb.2014.09.007.

Fortuny-Amat, J. and B. McCarl (1981). "A Representation and Economic Interpretation of a Two-Level Programming Problem." In: *The Journal of the Operational Research Society* 32.9, pp. 783–792. DOI: 10.1057/jors.1981.156.

Hansen, P., B. Jaumard, and G. Savard (1992). "New branch-and-bound rules for linear bilevel programming." In: *SIAM Journal on Scientific and Statistical Computing* 13.5, pp. 1194–1217. DOI: 10.1137/0913069.

Ito, Y. (2011). "Robust Wardrop Equilibria in the Traffic Assignment Problem with Uncertain Data." MA thesis. Graduate School of Informatics, Kyoto University. URL: http://www-optima.amp.i.kyoto-u.ac.jp/papers/master/2011_master_ito.pdf.

Kalashnikov, V., J. G. Flores Muñiz, and N. Kalashnykova (2020). "Bilevel Optimal Tolls Problems with Nonlinear Costs: A Heuristic Solution Method." In: *Beyond Traditional Probabilistic Data Processing Techniques: Interval, Fuzzy etc. Methods and Their Applications*. Ed. by O. Kosheleva, S. P. Shary, G. Xiang, and R. Zapatrin. Cham: Springer International Publishing, pp. 481–516. DOI: 10.1007/978-3-030-31041-7_28.

Kleinert, T., M. Labbé, I. Ljubić, and M. Schmidt (2021). "A Survey on Mixed-Integer Programming Techniques in Bilevel Optimization." In: *EURO Journal on Computational Optimization* 9. DOI: 10.1016/j.ejco.2021.100007.

Koutsoupias, E. and C. Papadimitriou (1999). "Worst-case equilibria." In: *Proceedings of the 16th Annual Conference on Theoretical Aspects of Computer Science*. STACS'99. Trier, Germany: Springer-Verlag, pp. 404–413. DOI: 10.1007/3-540-49116-3_38.

Labbé, M., P. Marcotte, and G. Savard (1998). "A Bilevel Model of Taxation and Its Application to Optimal Highway Pricing." In: *Management Science* 44.12.1. DOI: 10.1287/mnsc.44.12.1608.

LeBlanc, L. J. and D. E. Boyce (1986). "A bilevel programming algorithm for exact solution of the network design problem with user-optimal flows." In: *Transportation Research Part B: Methodological* 20.3, pp. 259–265. DOI: 10.1016/0191-2615(86)90021-4.

LeBlanc, L. J., E. K. Morlok, and W. P. Pierskalla (1975). "An efficient approach to solving the road network equilibrium traffic assignment problem." In: *Transportation Research* 9.5, pp. 309–318. DOI: 10.1016/0041-1647(75)90030-1.

Li, Z., R. Ding, and C. A. Floudas (2011). "A Comparative Theoretical and Computational Study on Robust Counterpart Optimization: I. Robust Linear Optimization and Robust Mixed Integer Linear Optimization." In: *Industrial & Engineering Chemistry Research* 50.18, pp. 10567–10603. DOI: 10.1021/ie200150p.

Luo, Z.-Q., J.-S. Pang, and D. Ralph (1996). *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press. DOI: 10.1017/CBO9780511983658.

Marcotte, P. (1986). "Network design problem with congestion effects: A case of bilevel programming." In: *Mathematical Programming* 34.2, pp. 142–162. DOI: 10.1007/BF01580580.

Mattia, S. (2013). "The Robust Network Loading Problem with Dynamic Routing." In: *Computational Optimization and Applications* 54.3, pp. 619–643. DOI: 10.1007/s10589-012-9500-0.

Mattia, S. (2019). "A polyhedral analysis of the capacitated edge activation problem with uncertain demands." In: *Networks* 74.2, pp. 190–204. DOI: 10.1002/net.21888.

Mattia, S. and M. Poss (2018). "A comparison of different routing schemes for the robust network loading problem: polyhedral results and computation." In: *Computational Optimization and Applications* 69.3, pp. 753–800. DOI: 10.1007/s10589-017-9956-z.

Migdalas, A. (1995). "Bilevel programming in traffic planning: Models, methods and challenge." In: *Journal of Global Optimization* 7.4, pp. 381–405. DOI: 10.1007/BF01099649.

Ordóñez, F. and N. E. Stier-Moses (2007). "Robust Wardrop Equilibrium." In: *Network Control and Optimization*. Ed. by T. Chahed and B. Tuffin. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 247–256. DOI: 10.1007/978-3-540-72709-5_26.

– (2010). "Wardrop Equilibria with Risk-Averse Users." In: *Transportation Science* 44.1, pp. 63–86. DOI: 10.1287/trsc.1090.0292.

Orlowski, S., M. Pióro, A. Tomaszewski, and R. Wessäly (2007). "SNDlib 1.0–Survivable Network Design Library." In: *Proceedings of the 3rd International Network Optimization Conference (INOC 2007), Spa, Belgium*. URL: https://opus4.kobv.de/opus4-zib/frontdoor/deliver/index/docId/958/file/ZR-07-15.pdf.

– (2010). "SNDlib 1.0–Survivable Network Design Library." In: *Networks* 55.3, pp. 276–286. DOI: 10.1002/net.20371.

Patriksson, M. (2015). *The Traffic Assignment Problem: Models and Methods*. Courier Dover Publications.

Rey, D. and M. W. Levin (2024). *A Branch-and-Price-and-Cut Algorithm for Discrete Network Design Problems Under Traffic Equilibrium*. Tech. rep. URL: https://optimization-online.org/?p=28420.

Sim, M. (2004). "Robust Optimization." PhD thesis. Massachusetts Institute of Technology, Sloan School of Management. URL: https://dspace.mit.edu/handle/1721.1/17725.

Soyster, A. L. (1973). "Technical Note—Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming." In: *Operations Research* 21.5, pp. 1154–1157. DOI: 10.1287/opre.21.5.1154.

Tian, L., A. Bashan, D.-N. Shi, and Y.-Y. Liu (2017). "Articulation points in complex networks." In: *Nature Communications* 8.14223. DOI: 10.1038/ncomms14223.

U.S. Bureau of Public Roads (1964). *Traffic assignment manual*. Washington, D.C.: U.S. Government Printing Office.

Wardrop, J. G. (1952). "Some Theoretical Aspects of Road Traffic Research." In: *Proceedings of the Institution of Civil Engineers* 1.3, 325—362. DOI: 10.1680/ipeds.1952.11259.

Wardrop, J. G. and J. I. Whitehead (1952). "Correspondence. Some Theoretical Aspects of Road Traffic Research." In: *ICE Proceedings: Engineering Divisions* 1.5, p. 767. DOI: 10.1680/ipeds.1952.11362.

West, D. B. (2001). *Introduction to Graph Theory*. 2nd Edition. Pearson Education, Inc.

(Y. Beck) Eindhoven University of Technology, Department of Industrial Engineering and Innovation Sciences, PO Box 513, 5600 MB Eindhoven, the Netherlands
Email address: y.beck@tue.nl

(F. Giancola) (A) Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", Consiglio Nazionale delle Ricerche, Via dei Taurini, 19, 00185 Rome, Italy; (B) Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma, Via Ariosto, 25, 00185 Rome, Italy
Email address: francesca.giancola@iasi.cnr.it

(I. Ljubić) ESSEC Business School, Department of Information Systems, Data Analytics and Operations, 3 Avenue Bernard Hirsch, 95021 Cergy-Pontoise Cedex, France
Email address: ljubic@essec.edu

(S. Mattia) Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", Consiglio Nazionale delle Ricerche, Via dei Taurini, 19, 00185 Rome, Italy
Email address: sara.mattia@iasi.cnr.it