The Burden of Risk Aversion in Mean-Risk Selfish Routing

EVDOKIA NIKOLOVA, University of Texas at Austin
NICOLAS E. STIER-MOSES, Facebook Inc., Universidad Torcuato Di Tella, and CONICET

Considering congestion games with uncertain delays, we compute the inefficiency introduced in network routing by risk-averse agents. At equilibrium, agents may select paths that do not minimize the expected latency so as to obtain lower variability. A social planner, who is likely to be more risk neutral than agents because it operates at a longer time-scale, quantifies social cost with the total expected delay along routes. From that perspective, agents may make suboptimal decisions that degrade long-term quality. We define the price of risk aversion (PRA) as the worst-case ratio of the social cost at a risk-averse Wardrop equilibrium to that where agents are risk-neutral. For networks with general delay functions and a single source-sink pair, we show that the PRA depends linearly on the agents’ risk tolerance and on the degree of variability present in the network. In contrast to the price of anarchy, in general the PRA increases when the network gets larger but it does not depend on the shape of the delay functions. To get this result we rely on a combinatorial proof that employs alternating paths that are reminiscent of those used in max-flow algorithms. For series-parallel (SP) graphs, the PRA becomes independent of the network topology and its size. As a result of independent interest, we prove that for SP networks with deterministic delays, Wardrop equilibria maximize the shortest-path objective among all feasible flows.

Categories and Subject Descriptors: J.4 [Computer Applications]: Social and Behavioral Sciences; F.2 [Theory of Computation]: Analysis of Algorithms and Problem Complexity

General Terms: Algorithms; Economics; Theory

Additional Key Words and Phrases: Congestion Game; Stochastic Networks; Risk-averse Wardrop Equilibrium; Mean-Stdev Risk Measure; Mean-Var Risk Measure; Nash Equilibrium; Stochastic Selfish Routing.

1. INTRODUCTION

A central question in decision making is how to make good decisions under uncertainty, in particular when decision makers are risk averse. Applications of crucial national importance, including alleviating congestion in transportation networks, as well as improving telecommunications, robotics, security and others, all face pervasive uncertainty and often require finding reliable or risk-minimizing solutions. Those applications have motivated the development of algorithms that incorporate risk primitives, and the inclusion of risk aversion in questions related to algorithmic game theory. While risk has been extensively studied in the fields of finance and operations, among others, in comparison there is relatively little literature in the theoretical computer science community devoted to this issue. One of the goals of this paper is to inspire more work devoted to understanding and mitigating risk in networked systems.

Author's addresses: E. Nikolova, Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX, USA; nikolova@austin.utexas.edu. N. E. Stier-Moses, Facebook Core Data Science, Menlo Park, CA, USA; Universidad Torcuato Di Tella School of Business, Buenos Aires, Argentina; and CONICET Argentina; nstier@utdt.edu. All work for this publication was performed prior to the second author joining Facebook.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EC’15, June 15–19, 2015, Portland, OR, USA.
ACM 978-1-4503-3410-5/15/06 ...$15.00.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
http://dx.doi.org/10.1145/2764468.2764485
Capturing uncertainty and risk aversion in traditional combinatorial problems studied by theoretical computer science often reduces to nonlinear or nonconvex optimization over combinatorial feasible sets, for which no efficient algorithms are known. Possibly due to the difficulty in writing the ensuing problems in simple terms, at present we lack a systematic understanding of how risk considerations can be successfully incorporated into classic combinatorial problems. Doing so would necessitate new techniques for analyzing risk-minimizing combinatorial structures rigorously.

Within the fields of algorithms and algorithmic game theory, routing has proved to be a pervasive source of important questions. Indeed, many fundamental questions on risk-averse routing are still open, including several intriguing cases where the complexity is unknown. For example, in a network with uncertain edge delays, what is the complexity of finding the path with highest probability of reaching the destination within a given deadline? What is the path that minimizes the mean delay plus the standard deviation along that path? The best known algorithms for both of these questions have a polynomial average and a polynomial smoothed complexity, but a subexponential worst-case running time [Nikolova et al. 2006b; Nikolova 2010]. Hence, these problems are unlikely to be NP-hard, yet polynomial algorithms have so far been elusive. There are also a myriad of other possible risk objectives, yielding nonlinear and nonconvex optimization problems over the path polytope, that are open from a complexity, algorithmic and approximation points of view.

Consequently, we are at the very beginning of understanding risk and designing appropriate models in routing games, which have been instrumental in the development of algorithmic game theory in the past two decades. Routing games capture decision-making by multiple agents in a networking context, internalizing the congestion externalities generated by self-minded agents. Externalities are traditionally captured by considering that edge delays are functions of the edge flow. In addition to the technical challenges involving the characterization and computation of equilibria, a key insight brought by the study of these games was that equilibria are not extremely inefficient. The price of anarchy—by now a widely studied concept used in a variety of problems and applications—represents the worst-case ratio between the cost of a Nash equilibrium and the cost of the socially-optimal solution. This ratio, which quantifies the degradation of system performance due to selfish behavior, was first defined in the context of routing and applied to a network with parallel links [Koutsoupias and Papadimitriou 2009]. It was subsequently analyzed for general networks and different types of players [Roughgarden and Tardos 2002; Roughgarden 2003; Correa et al. 2004, 2008]. With few exceptions, this stream of work has assumed that delays are deterministic, while almost every practical situation in which such games could be useful presents uncertainty. For instance, there are uncertain delays in a transportation network due to weather, accidents, traffic lights, etc., and in telecommunications networks due to changing demand, hardware failures, interference, packet retransmissions, etc.

**Risk Model.** A generalization of the classic selfish routing model [Beckmann et al. 1956] to the case of uncertain delays is to associate every edge to a random variable whose distribution depends on the edge flow. The problem faced by a risk-averse agent becomes more involved since it is not merely finding the shortest path with respect to delays. To find the best route, agents must consider both the expected delay and the variability along all possible choices, leading them to solve stochastic shortest path problems. For instance, it is common that commuters add a buffer to the expected travel time for their trip to maximize the chance of arriving on-time to an important meeting or to a flight at the airport. A classic model in finance that captures the tradeoff between mean and variability is Markowitz’ mean-risk framework [Markowitz 1952]. It considers an agent that optimizes a linear combination of mean and risk,
weighted by a risk-aversion coefficient that quantifies the degree of risk aversion of that agent. In the context of routing, Nikolova and Stier-Moses [2014a] adapted that framework to Wardrop equilibria, showed the existence of equilibria, commented on their uniqueness, and computed their worst-case inefficiency as captured by the price of anarchy. Henceforth, we now focus on studying the impact of risk aversion.

Of course, there are multiple ways to capture risk. The expected utility theory [Von Neumann and Morgenstern 1944], which is prevalent in economics, captures risk-averse preferences using concave utility functions. This theory has been criticized due to unrealistic assumptions such as independence of irrelevant alternatives so other theories have been proposed [Tversky and Kahneman 1981]. The theory of coherent risk measures, proposed in the late nineties, takes an axiomatic approach to risk [Rockafellar 2007; Krokhmal et al. 2011]. Cominetti and Torrico [2013] adapted these ideas to the context of network routing and concluded that the mean-variance objective has benefits over other risk measures in being additively consistent. In finance, the mean-risk and other traditional risk measures have been criticized for leading to paradoxes such as preferring stochastically dominated solutions. Nevertheless, different risk postulates may be relevant in the context of transportation and telecommunications. For instance, stochastically-dominated routes may be admissible if certainty is more valued than a stochastically-dominant solution with large variance. Indeed, one may choose a larger-latency path rather than routing along variable paths that introduce jitter in real-time communications. Following on our previous work on risk-averse congestion games [Nikolova and Stier-Moses 2014a], in this paper we consider both the mean-variance and mean-standard deviation objectives for risk-averse routing.

It is important to note that risk aversion may induce agents to choose longer routes to reduce risk, effectively trading off mean with variance. Hence, a natural question to ask for a network game with uncertain delays and risk-averse agents is how much of the degradation in system performance can be attributed to the agents’ risk-aversion. We refer to this degradation by the price of risk-aversion (PRA), which we formally define as the worst-case ratio of the social cost of the equilibrium (with risk-averse agents) to that of an equilibrium if agents were risk neutral. The rationale for choosing this particular ratio is that we want to disentangle the effects caused by selfish behavior, captured by the price of anarchy, from those caused by risk aversion per se. The social cost is considered with respect to average delays because a central planner would typically care about a long-term perspective and minimize average agent delays and average pollutant emissions.

Using the variance of delay along a route as a risk indicator, although not directly intuitive since it is not expressed in the same units as the mean delay, leads to models that satisfy natural and intuitive optimality conditions for routes; namely, a subpath of an optimal path remains optimal (called the additive consistency property). Indeed, the mean-variance objective is additive along paths (the cost of a path is the sum of the cost of its edges). It thus lends itself to tractable algorithms in terms of computing equilibria, at least as long as delays are pairwise independent across edges. On the other hand, comparing a Wardrop equilibrium with risk-averse players to a standard Wardrop equilibrium is far from straightforward and requires new techniques for understanding how the two differ.

Alternatively, one could set the risk indicator to be the standard deviation of delays. A big advantage is that the mean-standard deviation objective can be thought of as a quantile of delay, easily justifying the buffer time that commuters consider when selecting when to start the trip. The disadvantages are that additive consistency is lost and, technically, that to compute the standard deviation one must take a square root, which makes the objective non-separable and nonconvex. For more details, we
refer the reader to Nikolova and Stier-Moses [2014a], where these pros and cons are discussed in further detail.

**Our results**

We define a new concept, the price of risk aversion (PRA), as the worst-case ratio of the social cost (total expected delay) of a risk-averse Wardrop equilibrium (RAWE) to that of a risk-neutral Wardrop equilibrium (RNWE). Our main result, presented in Section 4, is a bound on the price of risk aversion for arbitrary graphs with a single origin-destination (OD) pair and symmetric players who minimize their mean-variance objective. We provide a bound of $1 + \gamma \kappa \eta$, where $\gamma$ is the risk-aversion coefficient, $\kappa$ is the maximum possible variability (variance-to-mean ratio) of all edges when the prevailing traffic conditions are those under the equilibrium, and $\eta$ is a topological parameter that captures how many flow-bearing paths are needed to cover a special structure called an alternating path. The parameter $\eta$ strongly depends on the topology, and is at most half the number of nodes in the network, $\lceil (n - 1)/2 \rceil$. The resulting bound is appealing in that it depends on the three factors that one would have expected (risk aversion, variability and network size), but perhaps unexpectedly does so in a linear way and for arbitrary delay functions. The proof of this result is based on constructing a type of **alternating path** that switches between forward edges for which the flow under a RAWE is less than or equal to the flow under a RNWE, and backward edges for which the opposite inequality holds. This construction is the key that allows us to compare both equilibrium flows and derive our main result.

The proof of the main result consists of three key lemmas that show that (a) an alternating path always exists, (b) the cost of a RAWE is upper-bounded by an inflated total mean delay along forward edges minus the total mean delay along backward edges (Lemma 4.4), and (c) the cost of a RNWE is lower-bounded by the total mean delay along forward edges minus the total mean delay along backward edges (Lemma 4.5). Steps (a) and (c) are proved independently of the choice of risk model. Step (b) is more subtle: it constructs a series of subpaths that connect different parts of the alternating path to the source and the sink, and uses the equilibrium conditions to provide partial bounds for subpaths of the alternating path. The lemma then exploits the linearity of the mean-variance objective to get a telescopic sum that simplifies precisely to the total delay along the alternating path.

Theorem 4.6 puts the lemmas together and upper bounds the total mean delay of the forward subpaths in the alternating path by the cost of the RNWE times the number of such forward paths, obtaining the factor $\eta \leq \lceil (n - 1)/2 \rceil$ in the worst-case, as mentioned above. We prove that this bound is tight for two families of graphs. For series-parallel (SP) graphs, it turns out that there must exist an alternating path that consists of only forward edges (that is, $\eta = 1$), which implies that the price of risk-aversion for those topologies is exactly $1 + \gamma \kappa$ (Corollary 4.8 and lower bound in Section 4.1). For Braess graphs, we establish that the price of risk aversion is bounded by $1 + 2\gamma \kappa$ and this bound is tight, as well.

As mentioned above, many of the results for the mean-variance risk model extend to the mean-stdev objective. In particular, the only piece missing to prove a general theorem is an equivalent of Lemma 4.4, which bounds the cost of the RAWE by an expression of the edge delays along the alternating path. The difficulty in extending our current proof to general graphs is the nonlinearity of the mean-stdev cost function, which in turn puts a restriction on the equilibrium flow in that its edge-flow representation cannot be decomposed arbitrarily to a path-flow representation. (The latter leads to an interesting open problem, posed by Nikolova and Stier-Moses [2014a]: is there an efficient algorithm that converts a given equilibrium edge-flow vector into an equilibrium path-flow decomposition? That reference shows that a succinct path flow...
decomposition that uses polynomially-many paths exists.) Circumventing the nonlinearity challenge, we are able to prove the equivalent of Lemma 4.4 on the Braess graph with a more involved case analysis (Lemma 5.6). Henceforth, we establish that the exact value of the price of risk aversion for Braess graphs in the mean-stdev case match those in the mean-variance case.

The independence of the network topology property for SP graphs also extends to the mean-stdev case. To obtain that extension, we provide a result for SP graphs that is interesting in its own right. As mentioned earlier, it has already been established that RNWE are typically not extremely inefficient because the price of anarchy is bounded. We prove that considering SP graphs with deterministic delays, the equilibrium maximizes the shortest path objective among all feasible flows (Theorem 5.7).

Missing proofs and additional information can be found in the full version of this paper [Nikolova and Stier-Moses 2014b].

2. RELATED WORK

In this work we consider how having stochastic delays and risk-averse users influence the traditional competitive network game introduced by Wardrop [1952]. He postulated that the prevailing traffic conditions can be determined from the assumption that users jointly select shortest routes, and the mathematics that go with this idea were formalized in Beckmann et al. [1956]. These models find applications in various application domains such as in transportation [Sheffi 1985] and telecommunications [Altman et al. 2006]. In the last decade, these types of models have received renewed attention with many studies aimed at understanding existence, uniqueness, computation, and efficiency of equilibria [Nisan et al. 2007; Correa and Stier-Moses 2011].

The variance terms on the edges in our mean-variance model can be interpreted as tolls and thus our work can be seen as related to the vast literature on tolls, of which the most closely related work is that by [Cole et al. 2006], [Bonifaci et al. 2011] and [Karakostas and Kolliopoulos 2005]. Our alternating path construction is similar to an alternating cycle concept used in the context of tolls by Bonifaci et al. [2011] and an alternating path concept used in the proof of Braess phenomena by Roughgarden [2006]. Despite the notion of alternations, the proofs for these three different contexts seem unrelated. It is an interesting open question to discover deeper connections between the three models and results.

There is a growing literature on stochastic congestion games with risk-averse players. Ordóñez and Stier-Moses [2010] introduce a game with uncertain delays and risk-averse users and study the relations between its solutions and percentile equilibria, which are flows under which percentiles of delays along flow-bearing paths are equal. Similarly to the present work, they compare risk-averse equilibria to those with risk-neutral players. Nie [2011] presents additional results on percentile equilibria. More closely related to the model considered here, Nikolova and Stier-Moses [2014a] prove existence and POA results, when the variability is captured by the standard deviations of delays. Piliouras et al. [2013] consider the sensitivity of the price of anarchy to several risk-averse user objectives, in a different routing game model with atomic players and affine delay functions. Angelidakis et al. [2013] also focus on atomic congestion games with uncertainty induced by stochastic players or stochastic delays, and char-
acterize when equilibria can be computed efficiently. Meir and Parkes [2015] study a congestion game where agents have uncertainty over the routes used by other agents, which leads to the consideration of a range of users choosing each edge.

For general congestion games, a series of papers in the last decade have studied the inefficiency introduced by self-minded behavior. To quantify that inefficiency, Koutsoupias and Papadimitriou [2009] computed the supremum over all problem instances of the ratio of the equilibrium cost to the social optimum cost, which has been called the price of anarchy (POA) [Papadimitriou 2001]. The POA has been characterized for increasingly more general assumptions [Roughgarden and Tardos 2002; Roughgarden 2003; Correa et al. 2004; Chau and Sim 2003; Perakis 2007; Correa et al. 2008]. Nikolova and Stier-Moses [2014a] extended that notion to the case of stochastic delays with risk-averse players. A different concept, the price of uncertainty, was considered in congestion games in reference to how best response dynamics change under randomness introduced by an adversary and random ordering of players [Balcan et al. 2009]. Risk aversion in the algorithmic game theory literature has been considered recently in the context of general games (e.g., Fiat and Papadimitriou [2010]), and mechanism design (e.g., Dughmi and Peres [2012]; Fu et al. [2013]; Dughmi [2014]). In transportation, the mean-variance model has been considered in the context of shortest paths [Khani and Boyles 2015].

The mean-standard deviation model has been considered in the context of shortest paths [Khan and Boyles 2015].

3. THE MODEL

We consider a directed graph $G = (V, E)$ with a single source-sink pair $(s, t)$ and an aggregate demand of $d$ units of flow that need to be routed from $s$ to $t$. We let $P$ be the set of all feasible paths between $s$ and $t$. We encode the decisions of the symmetric players as a flow vector $f = (f_p)_{p \in P} \in \mathbb{R}^{|P|}$ over all paths. Such a flow is feasible when demand is satisfied, as given by the constraint $\sum_{p \in P} f_p = d$. For notational simplicity, we denote the flow on a directed edge $e$ by $f_e = \sum_{p \in \ell} f_p$. When we need multiple flow variables, we use the analogous notation $x, x_p, x_e$ and $z, z_p, z_e$.

The network is subject to congestion, modeled with stochastic delay functions $\ell_e(f_e) + \xi_e(f_e)$ for each edge $e \in E$. Here, the deterministic function $\ell_e(f_e)$ measures the expected delay when the edge has flow $f_e$, and $\xi_e(f_e)$ is a random variable that represents a noise term on the delay, encoding the error that $\ell_e(\cdot)$ makes. Functions $\ell_e(\cdot)$, generally referred to as latency functions, are assumed continuous and non-decreasing.

The expected latency along a path $p$ is given by $\ell_p(f) := \sum_{e \in p} \ell_e(f_e)$. $\ell_p(f)$ is a random variable that represents a noise term on the delay, encoding the error that $\ell_p(\cdot)$ makes. Functions $\ell_p(\cdot)$, generally referred to as latency functions, are assumed continuous and non-decreasing.

Random variables $\xi_e(f_e)$ have expectation equal to zero and standard deviation equal to $\sigma_e(f_e)$, for arbitrary continuous functions $\sigma_e(\cdot)$. We assume that these random variables are pairwise independent. From there, the variance along a path equals $\sigma_p^2(f) = \sum_{e \in p} \sigma_e^2(f_e)$, and the standard deviation (stddev) is $\sigma_p(f) = \sigma_p^2(f)^{1/2}$. We will initially work with variances and then extend the model to standard deviations, which have the complicating square roots. (For details on the complications, we refer the reader to Nikolova and Stier-Moses [2014a]).

We will consider the nonatomic version of the routing game where infinitely many players control an infinitesimal amount of flow each so that the path choice of a single player does not unilaterally affect the costs experienced by other players (even though the joint actions of players affect other players).

As explained in the introduction, players are risk-averse and hence choose paths taking into account the variability of delays. We follow the literature and perturb the mean delay of path $p$ with a factor of the variance: $Q_p(f) = \ell_p(f) + \gamma \sigma_p(f)$. This objective function will be referred to as the mean-var objective, and frequently simply as the
**path cost** (as opposed to latency). Here, $\gamma \geq 0$ is a constant that quantifies the risk-aversion of the players, which we assume homogeneous. The special case of $\gamma = 0$ corresponds to risk-neutrality.

The variability of delays is usually not too large with respect to the expected latency. It is common to consider the **coefficient of variation** $CV_e(f_e) \coloneqq \sigma_e(f_e)/\ell_e(f_e)$ given by the ratio of the standard deviation to the expectation as a relative measure of variability [McAuliffe 2015]. In this case, we consider the **variance-to-mean ratio** $\nu_e(f_e)/\ell_e(f_e)$ as a relative measure of variability. Consequently, we assume that $\nu_e(x_e)/\ell_e(x_e)$ is bounded from above by a fixed constant $\kappa$ for all $e \in E$ at the equilibrium flow of interest $x_e \in \mathbb{R}_+$, which is less restrictive than requiring such a bound for all feasible flows. This means that the variance cannot be larger than $\kappa$ times the expected latency in any edge at the equilibrium flow.\(^1\)

In summary, an instance of the problem is given by the tuple $(G, d, \ell, v, \gamma)$, which represents the topology, the demand, the latency functions, the variability functions, and the degree of player risk-aversion.

The following definition captures that at equilibrium players route flow along paths with minimum cost $Q^*_p()$. In essence, users will switch routes until at equilibrium costs are equal along all used paths. This is the natural extension of the traditional Wardrop Equilibrium to risk-averse users.

**Definition 3.1 (Equilibrium).** A $\gamma$-equilibrium of a stochastic nonatomic routing game is a flow $f$ such that for every path $p \in \mathcal{P}$ with positive flow, the path cost $Q^*_p(f) \leq Q^*_q(f)$ for any other path $q \in \mathcal{P}$. For a fixed risk-aversion parameter $\gamma$, we refer to a $\gamma$-equilibrium as a **risk-averse Wardrop equilibrium** (RAWE), denoted by $f$.

Notice that since the variance decomposes as a sum over all the edges that form the path, the previous definition represents a standard Wardrop equilibrium with respect to modified costs $\ell_e(f_e) + \gamma \nu_e(f_e)$. For the existence of the equilibrium, it is sufficient that the modified cost functions are increasing.

Our goal is to investigate the effect that risk-averse players have on the quality of equilibria. The quality of a solution that represents collective decisions can be quantified by the cost of equilibria with respect to expected delays since, over time, different realizations of delays average out to the mean by the law of large numbers. For this reason, a social planner, who is concerned about the long term, is typically risk neutral, as opposed to users who tend to be more emotional about decisions. Furthermore, the social planner may aim to reduce long-term emissions, which would be better captured by the total expected delay of all users. These arguments justify the difference between the risk aversion coefficient that characterizes user behavior at equilibrium and the behavior of the social planner.

**Definition 3.2.** The **social cost** of a flow $f$ is defined as the sum of the expected latencies of all players: $C(f) \coloneqq \sum_{p \in \mathcal{P}} f_p \ell_p(f) = \sum_{e \in E} f_e \ell_e(f_e)$.

Although one could have measured total cost as the weighted sum of the costs $Q^*_p(f)$ of all users, this captures users’ utilities but not the system’s benefit. Nikolova and Stier-Moses [2014a] considered such a cost function to compute the price of anarchy; in the current paper, our goal is to compare across different values of risk aversion so we want the various flow costs to be measured with the same units.

---

\(^1\) Since we provide bounds that depend on the product $\gamma \kappa$, we could have defined the variance-to-mean ratio as the risk-aversion coefficient $\gamma$ times the variance over the mean instead, making the units of the numerator and denominator in the ratio the same. Although appealing, that would have imposed limits on the degree of risk-aversion of a user. In any case, both definitions lead to the same mathematical result, subject to a slightly different interpretation of the meaning of $\gamma \kappa$ in the final bound.
The next definition captures the increase in social cost at equilibrium introduced by user risk-aversion, compared to the cost one would have if users were risk-neutral. Hence, we use a risk-neutral Wardrop equilibrium (RNWE), defined as a 0-equilibrium according to Definition 3.1, as the yardstick to determine the inefficiency caused by risk-aversion. We define the price of risk aversion as the worst-case ratio among all possible instances of expected costs of the risk-averse and risk-neutral equilibria.

**Definition 3.3.** Considering a family of instances \( \mathcal{F} \) of a routing game with uncertain delays, the price of risk aversion (PRA) associated with \( \gamma \kappa \) (the risk-aversion coefficient times the variance-to-mean ratio) is defined by

\[
PRA(\mathcal{F}, \gamma, \kappa) := \sup_{G, d, \ell, v} \left\{ \frac{C(x)}{C(z)} : (G, d, \ell, v) \in \mathcal{F} \text{ and } v(x) \leq \kappa \ell(x) \right\},
\]

where \( x \) and \( z \) are the RAWE and the RNWE of the corresponding instance.

This supremum depends on \( \mathcal{F} \), which may be defined in terms of the network topology (as, e.g., general, series-parallel, or Braess networks) and the number of vertices. We do not impose restrictions to functions \( \ell \) or \( v \), except indirectly through the \( \kappa \) bound.

We present the following example to motivate the form of the bound to the PRA, which is linear in \( \gamma \kappa \). The example is based on a simple network with two edges, usually referred to as the Pigou network [Pigou 1920; Roughgarden and Tardos 2002].

**Example 3.4.** Consider an instance with two nodes connected by two parallel edges with latencies equal to \((1 + \gamma \kappa)x\) and 1, respectively, variances equal to \(v_1(\cdot) = 0\) and \(v_2(\cdot) = \kappa\), and \(d = 1\). Computing equilibria, the RNWE flow routes \(\frac{1}{1 + \gamma \kappa}\) units along the first edge and \(\frac{\gamma \kappa}{1 + \gamma \kappa}\) along the second. This gives a total cost of 1. Instead, the RAWE flow routes all the flow along the first edge, which gives a total cost of \(1 + \gamma \kappa\). Dividing, we get that \(PRA \geq 1 + \gamma \kappa\).

The previous example motivates the need of imposing an upper bound on the variability of delays.

**Remark 3.5.** Taking \( \kappa \to \infty \) in the previous example, it follows that if one does not constrain variability of delays, the price of risk aversion is unbounded.

Having bounded variability is a reasonable assumption in real-life networks since the variability is never too many times larger than the expected latency of an edge. Moreover, the more congested the network is, the less variable delays are since speeds approach zero and hence the possibilities of variation are minimal. In the following section, we shall prove that \(1 + \gamma \kappa\) is a matching upper bound for instances with the topology of Pigou networks. Indeed, we will see that this will be a special case of a result for general topologies.

### 4. PRA IN GENERAL GRAPHS

We first generalize the lower bound given in the previous section to suggest what one can expect in general, and then find an upper bound to the price of risk aversion.

#### 4.1. A lower bound to the PRA

We exhibit a family of Braess instances [Braess 1968; Roughgarden and Tardos 2002] that shows that \(1 + 2\gamma \kappa\) is a valid lower bound.

**Example 4.1.** Consider the symmetric Braess network instance shown in Figure 1 with arbitrary values for \(\gamma\) and \(v\) such that \(\gamma v \leq 1\). Latencies and variances are as shown in the figure. In this case, the value of \(\kappa\) is \(v\). We refer to the top path as \(p\), to the bottom path as \(q\), and to the zigzag path as \(r\). The RNWE is \(z_p = z_q = 1/2\)
with cost \( C(z) = \ell_p(z) = 1 + \gamma v \), and the RAWE is \( x_r = 1 \) with cost \( C(x) = \ell_r(x) = 1 + 3\gamma v \). To find the tightest lower bound of the form \( 1 + h\gamma \kappa \) for some \( h \), we compute \( \lim_{v \to 0} (C(x)/C(z) - 1)/(\gamma v) = 2 \) for any value of \( \gamma \) since \( \gamma v \leq 1 \) in the limit.

We next derive an upper bound for general networks, which confirms that the lower bound is tight for Braess instances.

### 4.2. An upper bound on the PRA for general graphs

We start by introducing bounds on the latency of the risk-averse Wardrop equilibrium (RAWE), which we will use to find the PRA. As before, we let \( z \) denote the risk-neutral Wardrop equilibrium (RNWE) and let \( x \) denote the RAWE. It is well-known that, by definition, the social cost \( C(z) \) of a RNWE can be upper-bounded by the latency \( \ell_p(z) \) of an arbitrary path \( p \in \mathcal{P} \), and the bound is tight if the path carries flow. We now extend that argument to a RAWE. We prove that its social cost is bounded by the cost \( Q_{\gamma}p(z) \) of an arbitrary path \( p \in \mathcal{P} \). As a corollary, \( C(z) \) is also bounded by the expected latency of an arbitrary path, blown up by a constant that depends on the risk-aversion coefficient \( \gamma \) and the maximum coefficient of variation \( \kappa \).

**Lemma 4.2.** Letting \( p \in \mathcal{P} \) denote an arbitrary path (potentially not carrying flow at equilibrium), the social cost of a RAWE \( C(x) \) is upper bounded by the path cost \( Q_{\gamma}p(z) \).

In addition, if variance functions satisfy that the variance-to-mean ratio at equilibrium is bounded by \( \kappa \), then \( C(x) \leq (1 + \gamma \kappa) \ell_p(x) \).

We proceed to bound the price of risk aversion on a general graph by an appropriate construction of an alternating path that contains edges from the following two sets, which form a partition of the edges in \( E \):

\[
A = \{ e \in E \mid z_e \geq x_e \text{ and } z_e > 0 \} \quad \text{and} \quad B = \{ e \in E \mid z_e < x_e \text{ or } z_e = x_e = 0 \}
\]

We will assume from here on that edges in \( B \) satisfy \( z_e < x_e \) since edges that carry no flow in both equilibria (\( z_e = x_e = 0 \)) can be removed from the graph without loss of generality. If there is a full \( s-t \) path \( \pi \) contained in the set \( A \), then it is not too hard to prove that \( C(x) \leq (1 + \gamma \kappa)C(z) \). In other words, this would give the lowest possible PRA bound of \( 1 + \gamma \kappa \). We now prove that this bound can be extended to alternating paths in \( G \), which are \( s-t \) paths consisting of edges in \( A \) plus reversed edges in \( B \). We shall refer to edges on the alternating path that belong to \( A \) as forward edges and those in \( B \) as backward edges.

**Definition 4.3.** We say that a path \( \pi := (e_1, \ldots, e_r) \) from \( s \) to \( t \) is alternating if reversing the direction of edges in \( B \) makes it a feasible \( s-t \) path.

Figure 2 provides an illustration of the alternating path definition (in bold) where reversing edges in \( B \) creates a feasible path. The existence of an alternating path follows from flow conservation and the definitions of sets \( A \) and \( B \) (see Lemma 4.5 in the full version of this paper [Nikolova and Stier-Moses 2014b]). We use the alternating
path to provide an upper bound on the PRA that depends on the number of times the alternating path switches from A to B. To get there, we need two lemmas. The first lemma extends Lemma 4.2, which applies to (standard) paths, to the case of alternating paths. Note that it allows us to tighten the previous bound by subtracting the latencies of the backward edges in the alternating path. The lemma provides an upper bound on the social cost of the RAWE $x$ by exploiting the equilibrium conditions on the subpaths $B_i$ on the alternating path with respect to the risk-averse objective.

**Lemma 4.4.** Consider a graph with variance functions that satisfy that the variance-to-mean ratio at equilibrium is bounded by $\kappa$. Letting $\pi$ be an alternating path, the social cost of a risk-averse Wardrop equilibrium $x$ satisfies $C(x) \leq (1 + \gamma\kappa) \sum_{e \in A \cap \pi} \ell_e(x_e) - \sum_{e \in B \cap \pi} \ell_e(x_e)$.

**Proof.** Let us assume that the alternating path consists of subpaths $A_1B_1A_2\ldots A_{\eta-1}B_{\eta-1}A_{\eta}$, where each subpath is in the corresponding set $A$ or $B$. Since by definition each edge $e$ in $B_k$ carries flow ($x_e > 0$) for any $k$, $e$ must belong to a flow-carrying $s$-$t$ path. Selecting a decomposition where the whole subpath $B_k$ is on the same path (we have the freedom to do that since this is a standard Wardrop model with respect to the mean-variance objective), there must be a flow-carrying path that consists of subpaths $C_kB_kD_k$ where $C_k$ originates at the source node and $D_k$ terminates at the destination node (see Figure 2 for an illustration). We define $C_0 = D_0 = \emptyset$.

To simplify notation, only for the proof of this lemma, we are going to refer to the mean-variance cost of subpath $P$ also by $P = \sum_{e \in P}(\ell_e(x_e) + \gamma v_e(x_e))$.

We next use the equilibrium conditions to derive bounds on $C_k$ and $D_k$. Since the subpath $C_kB_k$ carries flow and the subpath $C_{k-1}A_k$ is an alternative route between the endpoints of $C_kB_k$, we have that $C_k + B_k \leq C_{k-1} + A_k$ for all $k$. Note that here and in what follows we critically use the additivity of the mean-variance cost. Therefore,

$$C_k \leq C_{k-1} + A_k - B_k \leq \ldots \leq (A_1 + A_2 + \ldots + A_k) - (B_1 + B_2 + \ldots + B_k). \quad (2)$$

Similarly, since $B_kD_k$ carries flow and $A_{k+1}D_{k+1}$ is an alternative route between the same endpoints, we have that

$$D_k \leq (A_{k+1} + A_{k+2} + \ldots + A_{\eta}) - (B_{k} + B_{k+1} + \ldots + B_{\eta-1}). \quad (3)$$

Then, for path $q = C_kB_kD_k$ for any $k$, we have that

$$C(x) = \sum_{p} x_p \ell_p(x)$$

$$\leq \sum_{p} x_p(\ell_q(x) + \gamma v_q(x) - \gamma v_p(x))$$

$$\leq C_kB_kD_k$$

$$\leq (A_1 + \ldots + A_{\eta}) - (B_1 + \ldots + B_{\eta-1})$$

since either $x_p = 0$ or $Q^\gamma_p(x) \leq Q^\gamma_q(x)$ after neglecting the negative term using inequalities (2) and (3).
\[
\leq \sum_{i=1}^{\eta} \sum_{e \in A_i} (f_e(x_e) + \gamma v_e(x_e)) - \sum_{i=1}^{\eta-1} \sum_{e \in B_i} f_e(x_e) \text{ neglecting variances in negative term}
\]

\[
\leq (1 + \gamma \kappa) \sum_{e \in A \cap \pi} f_e(x_e) - \sum_{e \in B \cap \pi} f_e(x_e).
\]

The last inequality follows by applying the variability bound on the variances. □

The previous result provided an upper bound for the RAWE \( x \). Now, we complement it with a lower bound for the RNWE \( z \). Again, to get the result we exploit the equilibrium conditions, now with respect to \( \ell(\cdot) \).

**Lemma 4.5.** Letting \( \pi \) be an alternating path, the social cost of a risk-neutral Wardrop equilibrium \( z \) satisfies \( C(z) \geq \sum_{e \in A \cap \pi} f_e(x_e) - \sum_{e \in B \cap \pi} f_e(x_e) \).

**Proof.** Since \( z_e > 0 \) for any \( e \in A_k \), there must be a subpath \( C_{k-1} \) that brings flow to \( A_k \) (this \( C_{k-1} \) need not be the same as that used in the proof of Lemma 4.4). Then, there is a flow decomposition in which the subpath \( C_k A_k \) is used by \( z \). Because subpath \( C_k B_k \) is an alternative route from \( x \) to the node at the end of \( A_k \), we must have that \( \ell_{C_{k-1}}(z_e) + \ell_{A_k}(z_e) \leq \ell_{C_k}(z_e) + \ell_{B_k}(z_e) \). Summing the previous inequalities for all \( k \) (where \( C_0 \) is defined as an empty path), we get \( \ell_{C_{k-1}}(z_e) \geq \sum_{k=1}^{\eta-1} (\ell_{A_k}(z_e) - \ell_{B_k}(z_e)) \). This proves the lemma because \( C(z) = \ell_{C_{\eta-1}}(z) + \ell_{A_\eta}(z) \), since \( C_{\eta-1} A_\eta \) is a flow-carrying \( s-t \) path for \( z \). □

With the previous two lemmas that provided bounds for \( x \) and \( z \), and the sets \( A \) and \( B \) that allow us to compare both flows, the proof of the main result consists of just chaining the inequalities.

**Theorem 4.6.** Consider a general instance with variance functions that satisfy that the variance-to-mean ratio at equilibrium is bounded by \( \kappa \). Letting \( \pi \) be an alternating path, the price of risk aversion is upper bounded by \( 1 + \gamma \kappa \eta \), where \( \eta \) is the number of disjoint forward subpaths in the alternating path \( \pi \).

**Proof.**

\[
C(x) \leq (1 + \gamma \kappa) \sum_{e \in A \cap \pi} f_e(x_e) - \sum_{e \in B \cap \pi} f_e(x_e) \ 	ext{by Lemma 4.4}
\]

\[
\leq (1 + \gamma \kappa) \sum_{e \in A \cap \pi} f_e(z_e) - \sum_{e \in B \cap \pi} f_e(z_e) \ 	ext{by definition of \( A \) and \( B \)}
\]

\[
\leq C(z) + \gamma \kappa \sum_{e \in A \cap \pi} f_e(z_e) \ 	ext{by Lemma 4.5}
\]

\[
\leq C(z) + \gamma \kappa \eta C(z) = (1 + \gamma \kappa \eta) C(z).
\]

In the last inequality, we have used that \( \sum_{e \in A \cap \pi} f_e(z_e) \leq \eta C(z) \). This holds because for all forward subpaths \( A_k \in \pi \), their edges satisfy \( z_e > 0 \) so \( f_{A_k}(z) \leq f_q(z) = C(z) \) for some path \( q \) with \( z_q > 0 \) that includes the subpath \( A_k \). (The latter holds because any path flow decomposition is valid for the risk-neutral equilibrium.) □

The parameter \( \eta \), referred to in the introduction, is the maximum possible number of disjoint forward subpaths. By way of construction, an alternating path goes through every node at most once and the number of forward subpaths is maximized when the path consists of alternating forward and backward edges, for a total of at most \( n - 1 \) edges. Therefore \( \eta \leq \lceil (n - 1)/2 \rceil \).
The price of risk aversion in a general graph is upper bounded by \(1 + \gamma \kappa \lceil \frac{n - 1}{2} \rceil\).

The bound depends on the three factors that one would expect (risk aversion, variability and network size), but perhaps unexpectedly does so in a linear way and for arbitrary delay and variance functions. The tightness of the bound can be seen immediately for the family of all instances with a Braess graph topology, by applying Corollary 4.7 to the Braess graph used in Example 4.1 (see Figure 1) and noting that \(n = 4\) for the Braess graph.

Next, we derive that the price of risk aversion in series-parallel graphs is at most \(1 + \gamma \kappa\), independently of the size of the network. Given the lower bound provided by Example 3.4 (a Pigou graph is series-parallel), this bound must be tight. Series-parallel graphs are those formed recursively by subdividing an edge in two sub-edges, or replacing an edge by two parallel edges (see Figure 3). A noteworthy alternative characterization is that a graph is series-parallel if and only if it does not contain a Braess subgraph as an induced minor [Valdes et al. 1982].

Corollary 4.8. The price of risk aversion among all series-parallel instances is exactly \(1 + \gamma \kappa\).

Proof. We are going to prove that there exists an alternating path \(\pi\) consisting only of forward edges, so \(\pi \subseteq A\). Let us consider a minimal (cardinality-wise) alternating path with a backward edge. The key property of series-parallel graphs is that after taking a reverse edge \(e^-\), where \(e = (i, j) \in E\), \(\pi\) has to either come back to node \(j\) or close a loop with itself. If that did not happen, it would imply that a Braess graph is embedded in the instance, which is not possible. Hence, there is an alternating path \(\pi'\) without the reverse edge \(e^-\), which is a contradiction to the minimality of \(\pi\). □

5. Representing Risk as the Standard Deviation

We now consider a related risk measure based on the standard deviation rather than the variance. The mean-stdev objective that each user seeks to minimize is a linear combination of the expectation and the standard deviation along a route \(p\), parametrized as before by the risk-averse coefficient \(\gamma\):

\[
Q_\gamma^p(f) = \sum_{e \in p} \ell_e(f_e) + \gamma \sqrt{\sum_{e \in p} \sigma_e(f_e)^2}. \tag{4}
\]

For this objective, equilibrium existence follows from a variational inequality formulation if standard deviation functions \(\sigma(x)\) are continuous, as we have assumed here [Nikolova and Stier-Moses 2014a].

Example 3.4 and Remark 3.5 for the mean-var model can be adapted here, replacing the variances with standard deviations in the example specification. Since for arbitrary instances the PRA is unbounded, we assume that \(\sigma_e(x_e)/\ell_e(x_e)\) is no more than a fixed constant \(\kappa_\sigma\) for all \(e \in E\) at the RAWE \(x_e \in \mathbb{R}_+\). This means that the standard deviation cannot be larger than \(\kappa_\sigma\) times the expected latency in any edge at the equilibrium flow.
We start by identifying which results extend from the mean-var model to the mean-stdev model here. Essentially all lemmas extend, except for Lemma 4.4. Proving this lemma is thus the only remaining roadblock to proving the equivalent of Theorem 4.6 in the case of the mean-stdev cost, namely establishing a price of risk aversion bound for general graphs.

For completeness, we restate the lemmas and some of the proofs that require a slight modification. By the definition of equilibrium, the cost $C(z)$ of a RNWE can be bounded by the latency $\ell_p(z)$ of an arbitrary path $p$, and both are equal if $z_p > 0$. We now extend that argument to a RAWE $x$. We prove that its total cost is bounded by the expected latency of an arbitrary path, blown up by a constant that depends on the risk-aversion coefficient and the maximum coefficient of variation.

**Lemma 5.1.** Consider a general instance with a single source-sink pair and general latencies and standard deviation functions. Letting $p\in\mathcal{P}$ denote an arbitrary path (potentially not carrying flow at equilibrium), the social cost of a RAWE $C(x)$ is bounded by the path cost $Q^\gamma_p(x)$.

**Corollary 5.2.** Consider a general instance with a single source-sink pair, general latencies and general standard deviation functions that satisfy that the coefficient of variation at equilibrium is bounded by $\kappa_\sigma$. Letting $p\in\mathcal{P}$ denote an arbitrary path (potentially not carrying flow at equilibrium), the social cost $C(x)$ of a RAWE $x$ is bounded by $(1 + \gamma\kappa_\sigma)\ell_p(x)$.

**Proof.** From Lemma 5.1,

$$C(x) \leq Q^\gamma_p(x) = \ell_p(x) + \gamma \sqrt{\sum_{a\in p} \sigma_a^2(x)} \leq \ell_p(x) + \gamma \sum_{a\in p} \kappa_\sigma \ell_a(x) = \ell_p(x)(1 + \gamma\kappa_\sigma).$$

Here, we have used that $\|x\|_2 \leq \|x\|_1$ for an arbitrary nonnegative vector, and applied the bound of $\kappa_\sigma$ on the coefficient of variation. □

As in the mean-var case, to get the tightest possible upper bound, one would consider the path with smallest standard deviation induced by the RAWE. Selecting that path, we can get rid of the factor and bound the total cost by the expected latency of the path.

**Lemma 5.3.** Consider a general instance with a single source-sink pair, general latencies and general standard deviation functions. Letting $p\in\mathcal{P}$ denote the path that minimizes the standard deviation under a RAWE $x$ (where the path $p$ may or may not carry flow), the social cost $C(x)$ is bounded by $\ell_p(x)$.

An alternating path exists as in the mean-var model, even though the corresponding sets $A$ and $B$ will be different reflecting that $x$ is now a RAWE for the mean-stdev cost. We now state a special case of Theorem 4.6 for the mean-stdev model, namely that if we can find an alternating path in the set $A$, then the PRA in the mean-stdev model is $1 + \gamma\kappa_\sigma$.

**Theorem 5.4.** Consider a general instance with standard deviation functions that satisfy that the coefficient of variation (stdev-to-mean ratio) at equilibrium is bounded by $\kappa_\sigma$. Suppose there exists an alternating path $p$ consisting of forward edges only (namely, edges in the set $A = \{e\in E \mid z_e \geq x_e \text{ and } z_e > 0\}$). Then, the cost of a risk-averse Wardrop equilibrium $x$ satisfies $C(x) \leq C(z) + (1 + \gamma\kappa_\sigma)C(z)$ for a risk-neutral Wardrop equilibrium $z$.

Extending Lemma 4.4 to the mean-stdev objective for general graphs remains elusive. The proof for the mean-var objective relies on the equilibrium conditions on sub-paths of the RAWE $x$. Although the mean-var objective leads to a separable model, the
mean-stdev one does not (for details, we refer the reader to Nikolova and Stier-Moses [2014a]), and that complicates a general proof. Moreover, as an additional complication, in the variance case we use a flow decomposition that suits our needs but for the standard deviation case, the decompositions cannot be arbitrary [Nikolova and Stier-Moses 2014a] so we cannot guarantee that the one we need is valid. As preliminary steps to a general proof, the following results yield tight bounds for PRA in two well-studied families of graphs: Braess networks and series-parallel networks.

5.1. Price of risk aversion in the Braess Network

In what follows, we will prove a version of Lemma 4.4 for the mean-stdev in the Braess graph and consequently we will get that the PRA is $1 + 2\gamma \kappa_r$. We start with an auxiliary lemma. As before, we denote the top path by $p$, consisting of edges $(a, b)$; the bottom path by $q$, consisting of edges $(c, d)$, and the zigzag path by $r$, consisting of edges $(a, c, d)$.

**Lemma 5.5.** For an arbitrary flow $f$ in a Braess network, $\sigma_p(f) + \sigma_q(f) - \sigma_r(f) \leq \sigma_b(f) + \sigma_c(f)$ when $\sigma_r(f) \leq \max(\sigma_p(f), \sigma_q(f))$.

**Proof.** To simplify notation and since the flow $f$ does not play a role in the proof, we are going to suppress the dependence on $f$ of the standard deviation functions. The inequality we wish to prove is equivalent to

$$\sqrt{\sigma_a^2 + \sigma_b^2} + \sqrt{\sigma_d^2 + \sigma_a^2} \leq \sigma_b + \sigma_c + \sqrt{\sigma_a^2 + \sigma_c^2 + \sigma_d^2} \Leftrightarrow 2\sqrt{(\sigma_a^2 + \sigma_b^2)(\sigma_d^2 + \sigma_a^2)} \leq 2\sigma_b \sigma_c + \sigma_c^2 + 2(\sigma_b + \sigma_c) \sqrt{\sigma_a^2 + \sigma_c^2 + \sigma_d^2}.$$

Finally, squaring once more, we get

$$\sigma_a^2 \sigma_d^2 \leq \frac{\sigma_a^4}{4} + \sigma_b \sigma_c \sigma_d^2 + \sigma_a^2 \sigma_b^2 + \sigma_c^2 \sigma_d^2 + 2\sigma_b \sigma_c (\sigma_a^2 + \sigma_c^2 + \sigma_d^2) + \sigma_a^2 (\sigma_b^2 + \sigma_c^2) + (2\sigma_b \sigma_c + \sigma_c^2) (\sigma_b + \sigma_c) \sqrt{\sigma_a^2 + \sigma_c^2 + \sigma_d^2}.$$

If $\sigma_p \geq \sigma_r$, the last inequality holds because $\sigma_a^2 + \sigma_b^2 \geq \sigma_a^2 + \sigma_c^2 + \sigma_d^2$, from where $\sigma_b \geq \sigma_d$. The case of $\sigma_q \geq \sigma_r$ is similar. \hfill \Box

We now prove a variant of Lemma 4.4 for the mean-stdev cost on Braess graphs.

**Lemma 5.6.** Consider a Braess network with standard deviation functions that satisfy that the coefficient of variation at equilibrium is bounded by $\kappa_r$. Letting $\pi$ be an alternating path, the social cost $C(x)$ of a RAWE $x$ is upper bounded by

$$(1 + \gamma \kappa_r) \sum_{e \in A \cap x} \ell_e(x_e) - \sum_{e \in B \cap x} \ell_e(x_e).$$

**Proof.** If $\pi \subseteq A$, the result follows from Theorem 5.4. This can happen when $\pi \in \{p, q, r\}$. Otherwise $\pi$ is the alternating path that consists of edges $c, e^-$, and $b$. In that case, $e \in B$. Note that $z_e = x_e = 0$ cannot hold for $e$ because it is part of the alternating path, so we must have that $x_e > 0$ and consequently that $x_r > 0$. We must prove that $C(x) \leq (1 + \gamma \kappa_r)(\ell_a(x_e) + \ell_b(x_b)) - \ell_e(x_e)$. Since $x_r > 0$, we have that

$$\ell_a(x) + \gamma \sigma_r(x) \leq \ell_p(x) + \gamma \sigma_p(x) \Leftrightarrow$$

$$\ell_a(x) + \ell_e(x) + \ell_d(x) \leq \ell_a(x) + \ell_b(x) + \gamma (\sigma_p(x) - \sigma_r(x)) \Leftrightarrow \ell_d(x) \leq \ell_b(x) - \ell_e(x) + \gamma (\sigma_p(x) - \sigma_r(x)). \ (5)$$

502
Similarly, $\ell_r(x) + \gamma \sigma_r(x) \leq \ell_q(x) + \gamma \sigma_q(x) \iff$

\[\ell_a(x) + \ell_c(x) + \ell_d(x) \leq \ell_e(x) + \ell_d(x) + \gamma (\sigma_q(x) - \sigma_r(x)) \]

\[\iff \ell_a(x) \leq \ell_e(x) - \ell_c(x) + \gamma (\sigma_q(x) - \sigma_r(x)). \quad (6)\]

If $\sigma_r(x) \leq \max\{\sigma_p(x), \sigma_q(x)\}$, then we can apply Lemma 5.5:

$$C(x) \leq \ell_r(x) + \gamma \sigma_r(x) \quad \text{by Lemma 5.1}$$

$$= \ell_a(x) + \ell_c(x) + \ell_d(x) + \gamma \sigma_r(x)$$

$$\leq \ell_e(x) - \ell_c(x) + \gamma (\sigma_q(x) - \sigma_r(x)) + \ell_e(x) + \ell_b(x) - \ell_e(x) + \gamma (\sigma_p(x) - \sigma_r(x)) + \gamma \sigma_r(x)$$

$$= \ell_e(x) + \ell_b(x) - \ell_c(x) + \gamma (\sigma_p(x) + \sigma_q(x) - \sigma_r(x))$$

$$\leq \ell_e(x) + \ell_b(x) - \ell_c(x) + \gamma (\sigma_b(x) + \sigma_c(x_e)) \quad \text{by Lemma 5.5}$$

$$\leq (1 + \gamma \kappa_r) (\ell_e(x_e) + \ell_b(x_b)) - \ell_e(x_e),$$

where the second inequality follows by (5)-(6) and the last inequality follows from the coefficient of variation constraint $\sigma(x) \leq \kappa_r \ell(x)$ for the corresponding edges.

Otherwise, $\sigma_r(x) > \sigma_p(x)$ and $\sigma_r(x) > \sigma_q(x)$. Then, inequalities (5) and (6) imply that

$$\ell_b(x) \geq \ell_d(x) + \ell_e(x) + \gamma (\sigma_r(x) - \sigma_q(x)) \geq \ell_d(x) + \ell_e(x)$$

and

$$\ell_e(x) \geq \ell_e(x) + \ell_c(x) + \gamma (\sigma_r(x) - \sigma_q(x)) \geq \ell_a(x) + \ell_e(x).$$

Summing the above two inequalities, we have: $\ell_b(x) + \ell_e(x) \geq \ell_d(x) + \ell_c(x) + \ell_a(x) + \ell_e(x) = \ell_r(x) + \ell_e(x)$. Therefore,

$$C(x) \leq (1 + \gamma \kappa_r) \ell_r(x) \quad \text{by Corollary 5.2}$$

$$\leq (1 + \gamma \kappa_r) (\ell_b(x) + \ell_e(x) - \ell_c(x)) \quad \text{by inequality above}$$

$$\leq (1 + \gamma \kappa_r) (\ell_b(x) + \ell_e(x)) - \ell_e(x).$$

This completes the proof of this lemma. \qed

Using Lemma 5.6 in place of Lemma 4.4, we can apply the proof of Theorem 4.6 to get a tight PRA bound of $1 + \gamma \kappa_r \lceil (n-1)/2 \rceil = 1 + 2 \gamma \kappa_r$ as in the mean-var model. The extension to general networks seems much harder than in the case where the cost is the mean-var objective because not all flow decompositions are valid.

### 5.2. Series-Parallel Networks

For series-parallel (SP) networks, a tight bound of $1 + \gamma \kappa_r$ on the price of risk aversion follows from Theorem 5.4 combined with the proof of Corollary 4.8. An alternative way to derive this bound is via a result of independent interest that states that for SP networks, shortest paths are longest at a risk-neutral equilibrium, compared to any other flow.

**Theorem 5.7.** Consider a RNWE $z$ of a SP network with a single source-sink pair. Then, $z$ maximizes the shortest expected path delay $S(f) := \min_{P \in P} \ell_P(f)$ among all feasible flows $f$.

**Proof.** We use induction on the construction steps of the SP network. First, let us consider that the last composition in the construction of the network is series. Because it is a series composition, the restriction of $z$ to each component is a RNWE for the component. Considering an arbitrary feasible flow $f$, the induction implies that the restriction of $f$ to each component cannot have a longer shortest path than the restriction of $z$ in the same component. We get the desired inequality adding the inequalities corresponding to each component back together.

Second, let us consider that the last composition in the construction of the network is parallel and denote the subcomponents by $G_1, G_2, \ldots, G_k$. Let us denote by $z_i$ and $f_i$
the projection of each flow into a component. There must exist a component \( i \) such that 
\[ f_i \leq z_i \] and \( z_i > 0 \) because 
\[ \sum z_i = \sum f_i = 1. \] Then, 
\[ S(f) \leq S(f_i) \leq S(RNWE_{\mathcal{G}_i}(f_i)) \leq S(RNWE_{\mathcal{G}_i}(z_i)) = S(z). \] Here, we have denoted by \( RNWE_{\mathcal{G}_i}(d) \) the risk-neutral equilibrium in subgraph \( \mathcal{G}_i \) for demand \( d \). The first inequality is because \( S(f) \) is the minimal shortest path across components, the second is by the inductive hypothesis applied to the component \( \mathcal{G}_i \) with demand \( f_i \), the third is because increasing demand from \( f_i \) to \( z_i \) cannot reduce the shortest path at equilibrium, and the fourth is because some flow is routed through component \( i \) so the shortest path in that component is equal to the shortest path in the whole graph.

The bound for the price of risk-aversion of \( 1 + \gamma \kappa \sigma \) for SP networks follows by combining Theorem 5.7 with the following general bound, which also holds for the mean-variance model.

**Theorem 5.8.** Consider an arbitrary graph with standard deviations whose coefficients of variation are bounded by \( \kappa \sigma \). Then, the PRA is bounded by \( (1 + \gamma \kappa \sigma) \rho \), where 
\[ \rho := \min_{p \in \mathcal{P}} \ell_p(x) / \min_{q \in \mathcal{P}} \ell_q(z) \] for a RAWE \( x \) and a RNWE \( z \).

### 6. CONCLUSION

This paper marks a first step in understanding the consequences on the inefficiency of selfish routing caused by uncertain edge delays and risk-averse players. We have established an upper bound on the ratio of the cost of the risk-averse equilibrium to that of the risk-neutral one, for users that aim to minimize the mean-variance of their route in a general network. We have proved that the bound is tight on series-parallel and Braess networks. In addition, we have shown that both tight bounds extend to the case where users minimize the mean-standard deviation of a route instead and have elaborated on the challenges of extending the analysis to general graphs for users with such risk profiles. We leave open whether there is a deeper connection between the alternating paths in our upper bound on the price of risk aversion and those in the proof on the severity of Braess paradox [Roughgarden 2006], as well as the alternating (negative) cycles in the study of tolls of Bonifaci et al. [2011].

We remark that our model and results can extend to locally correlated edge delays via a polynomial graph transformation, as mentioned in Nikolova and Stier-Moses [2014a]. It is then possible to obtain a graph with independent edge delays (albeit with more vertices) where all our results carry through.

Some immediate open questions include if the general upper bound can be extended to mean-stdev and other risk objectives, and if the bounds or analysis can be extended to heterogeneous risk profiles, and multiple origin-destination pairs. Another interesting direction is to characterize how much risk can help rather than hurt the quality of equilibrium and social welfare. For a different model of uncertainty and user objectives, Meir and Parkes [2015] show that moderate uncertainty can improve the social welfare. More closely related to our model, Lianeas [2014] characterizes when a socially-optimal flow can be enforced as an equilibrium of risk-averse users. This is also related to the question of what flow can be enforced as an equilibrium under restricted tolls [Bonifaci et al. 2011]. In light of these recent results and our current work, it would be especially interesting to understand when risk-averse attitudes can be leveraged in mechanism design in the place of tolls to reduce congestion.

### ACKNOWLEDGMENTS

The authors gratefully acknowledge funding from the National Science Foundation grant numbers CCF-1216103, CCF-1350823, CCF-1331863, a Google Faculty Research Award, CONICET Argentina Grant PIP 112-201201-00450CO and ANPCyT Argentina Grant PICT-2012-1324.
REFERENCES

E. Altman, T. Boulogne, R. El-Azouzi, T. Jiménez, and L. Wynter. A survey on networking games in telecommunications. Computers and Operations Research, 33(2): 286–311, 2006.

H. Angelidakis, D. Fotakis, and T. Lianeas. Stochastic congestion games with risk-averse players. In SAGT, LNCS 8146, 86–98, 2013.

M.-F. Balcan, A. Blum, and Y. Mansour. The price of uncertainty. In EC, 285–294, 2009.

M. J. Beckmann, C. B. McGuire, and C. B. Winsten. Studies in the Economics of Transportation. Yale University Press, 1956.

V. Bonifaci, M. Salek, and G. Schafer. Efficiency of restricted tolls in non-atomic network routing games. In SAGT, LNCS 6982, 302–313, 2011.

S. D. Boyles and S. T. Waller. A mean-variance model for the minimum cost flow problem with stochastic arc costs. Networks, 56(3):215–227, 2010.

S. D. Boyles, K. M. Kockelman, and S. T. Waller. Congestion pricing under operational, supply-side uncertainty. Transportation Research C, 18(4):519–535, 2010.

D. Braess. Über ein Paradoxon aus der Verkehrsplanung. Unternehmensforschung, 12:258–268, 1968.

C. K. Chau and K. M. Sim. The price of anarchy for non-atomic congestion games with symmetric cost maps and elastic demands. Operations Research Letters, 31(5): 327–334, 2003.

R. Cole, Y. Dodis, and T. Roughgarden. How much can taxes help selfish routing? Journal of Computer and System Sciences, 72(3):444–467, 2006.

R. Cominetti and A. Torrico. Additive consistency of risk measures and its application to risk-averse routing in networks. Technical report, Universidad de Chile, 2013.

J. R. Correa and N. E. Stier-Moses. Wardrop equilibria. In J. J. Cochran, editor, Encyclopedia of Operations Research and Management Science. Wiley, 2011.

J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. Selfish routing in capacitated networks. Mathematics of Operations Research, 29(4):961–976, 2004.

J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. A geometric approach to the price of anarchy in nonatomic congestion games. Games and Economic Behavior, 64:457–469, 2008.

S. Dughmi. On the hardness of signaling. In FOCS, 354–363, 2014.

S. Dughmi and Y. Peres. Mechanisms for risk averse agents, without loss. Workshop on risk aversion in algorithmic game theory and mechanism design, 2012.

A. Fiat and C. Papadimitriou. When the players are not expectation maximizers. In SAGT, 1–14, 2010.

H. Fu, J. Hartline, and D. Hoy. Prior-independent auctions for risk-averse agents. In EC, 471–488, 2013.

G. Karakostas and S. G. Kolliopoulos. The efficiency of optimal taxes. In CAAN, LNCS 3405, 3–12, 2005.

A. Khani and S. D. Boyles. An exact algorithm for the mean-standard deviation shortest path problem. Transportation Research B, 2015. Forthcoming.

E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In STACS, 404–413, 1999.

P. Krokhmal, M. Zabarankin, and S. Uryasev. Modeling and optimization of risk. Surveys in Operations Research and Management Science, 16(2):49–66, 2011.

J. Li and A. Deshpande. Maximizing expected utility for stochastic combinatorial optimization problems. In FOCS, 797–806, 2011.

T. Lianeas. Congestion games: stochastic extensions and techniques for reducing the price of anarchy. Ph.D. thesis. National Technical University of Athens. 2014
H. Markowitz. Portfolio selection. The Journal of Finance, 7(1):77–91, 1952.
R. E. McAuliffe. Coefficient of Variation. Wiley Encyclopedia of Management, 8(1), 2015.
R. Meir and D. Parkes. Congestion games with distance-based strict uncertainty. In AAAI, 2015. Forthcoming.
Y. Nie. Multi-class percentile user equilibrium with flow-dependent stochasticity. Transportation Research B, 45:1641–1659, 2011.
E. Nikolova. Approximation algorithms for reliable stochastic combinatorial optimization. In APPROX/RANDOM, 338–351, 2010.
E. Nikolova and N. E. Stier-Moses. A mean-risk model for the traffic assignment problem with stochastic travel times. Operations Research, 62(2):366–382, 2014a.
E. Nikolova and N. E. Stier-Moses. The Burden of Risk Aversion in Mean-Risk Selfish Routing. Working paper, 2014b. http://arxiv.org/abs/1411.0059
E. Nikolova, M. Brand, and D. R. Karger. Optimal route planning under uncertainty. In ICAPS, 131–141, 2006a.
E. Nikolova, J.A. Kelner, M. Brand, and M. Mitzenmacher. Stochastic shortest paths via quasi-convex maximization. In ESA, LNCS 4168, 552–563, 2006b.
N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani. Algorithmic Game Theory. Cambridge University Press, Cambridge, UK, 2007.
F. Ordóñez and N. E. Stier-Moses. Wardrop equilibria with risk-averse users. Transportation Science, 44(1):63–86, 2010.
C. H. Papadimitriou. Algorithms, games, and the Internet. In STOC, 749–753, 2001.
G. Perakis. The “price of anarchy” under nonlinear and asymmetric costs. Mathematics of Operations Research, 32(3):614–628, 2007.
A. C. Pigou. The Economics of Welfare. Macmillan, London, 1920.
G. Piliouras, E. Nikolova, and J.S. Shamma. Risk sensitivity of price of anarchy under uncertainty. In EC, 715–732, 2013.
T. Rockafellar. Coherent approaches to risk in optimization under uncertainty. Tutorials in operations research, pages 38–61, 2007.
T. Roughgarden. The price of anarchy is independent of the network topology. Journal of Computer and System Sciences, 67(2):341–364, 2003.
T. Roughgarden and É. Tardos. How bad is selfish routing? Journal of the ACM, 49(2):236–259, 2002.
T. Roughgarden. On the severity of Braess’s paradox: Designing networks for selfish users is hard. Journal of Computer and System Sciences, 72(5):922–953, 2006.
Y. Sheffi. Urban Transportation Networks. Prentice-Hall, Englewood, NJ, 1985.
C. Swamy. Risk-averse stochastic optimization: probabilistically-constrained models and algorithms for black-box distributions. In SODA, 1627–1646, 2011.
A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. Science, New Series, 211:453–458, 1981.
J. Valdes, R. Tarjan, and E. Lawlers. The Recognition of Series Parallel Digraphs. SIAM J. Comput 11: 298–313, 1982.
J. Von Neumann and O. Morgenstern. Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ, 1944.
J. G. Wardrop. Some theoretical aspects of road traffic research. Proceedings of the Institute of Civil Engineers, Part II, Vol. 1, pages 325–378, 1952.